

LAMPIÃO

REVISTA DE FILOSOFIA

DOSSIÊ
FILOSOFIA
DA MENTE
& CIÊNCIAS
COGNITIVAS

PARTE 2

VOLUME

05

NÚMERO 1

2024

ISSN: 2675-9659





LAMPIÃO

Revista do Programa de Pós-graduação em Filosofia da UFAL

ISSN: 2675-9659

Publicação anual

Conselho editorial

André Leclerc (UNB)
Cristina Amaro Viana (UFAL)
Fernando Monegalha (UFAL)
Francisco Pereira de Sousa (UFAL)
Juliele Maria Sievers (UFAL)
Marcos Silva (UFPE)
Marcus José Alves de Souza (UFAL)
Maxwell Moraes de Lima Filho (UFCA)
Rodrigo Barros Gewehr (UFAL)
Taynam Santos Luz Bueno (UFAL)

Editores

Fernando Monegalha (UFAL)
Rodrigo Barros Gewehr (UFAL)

Editores convidados

Argus Romero Abreu de Moraes
Marcus José Alves de Souza
Maxwell Moraes de Lima Filho

Ilustrações

André Dias

Imagem da capa

Pedro Lucena

Endereço

Programa de Pós-graduação em Filosofia
Universidade federal de Alagoas – Campus A. C. Simões.
Instituto de Ciências Humanas, Comunicação e Artes
Av. Lourival Melo Mota, s/n. Tabuleiro dos Martins. Maceió, Al.
CEP: 57.072-970

E-mail

lampiaorevista@gmail.com



Licença: Creative Commons. Atribuição: Não Comercial-
Compartilha Igual 4.0

Sumário

PÁGINA

- 8** Apresentação
MARCUS JOSÉ ALVES DE SOUZA, ARGUS ROMERO ABREU DE MORAIS &
MAXWELL MORAIS DE LIMA FILHO
- ARTIGOS**
- 11** A tese da identidade mente-corpo e o fisicismo qualitativo
OSVALDO PESSOA JR.
- 35** A emergência da perspectiva da primeira pessoa
CLÁUDIA PASSOS-FERREIRA
- 56** Em busca de um sentido plausível de *self*
FABIENSE PEREIRA ROMÃO & LEONARDO FERREIRA ALMADA
- 80** Cores e mundo: uma discussão sobre a hipótese do espectro invertido
JULIANA DE ORIONE ARRAES FAGUNDES
- 91** Mente, mundo e corpo: responsabilidade e agenciamento estendido
MURILO KARASINSKI & LÉO PERUZZO JÚNIOR
- 110** Nos entrelaçamentos da cognição, a experiência estética
MONICA AIUB
- 124** Filosofia da psicologia de Ludwig Wittgenstein: as palavras no fluxo da vida
ARTURO FATTURI
- 142** Linguagem e mente: o problema da Inteligência Artificial de um ponto de vista
wittgensteiniano
MIRIAN DONAT
- 159** A teoria de Franz Brentano sobre a consciência: debates atuais
TÁRIK DE ATHAYDE PRATA
- 183** William James: pragmatismo, intencionalidade, 4ES
ARTHUR ARAÚJO
- 213** Uma análise crítica da proposta neomecanicista de William Bechtel para a ciência
cognitiva
DIEGO AZEVEDO LEITE
- 235** Análises investigativas dos seres vivos e a questão do princípio anímico em Aris-
tóteles
RODRIGO ROMÃO DE CARVALHO

253 O absurdo da matéria e a palavra extraviada: a crítica de Berkeley aos aspectos estruturantes da teoria do conhecimento

LUAMA SOCIO

RESENHA

267 Resenha do livro *Neurocognitive Mechanisms: explaining Biological Cognition* (Oxford University Press, 2020), de Gualtiero Piccinini

DIEGO AZEVEDO LEITE

TRADUÇÃO

277 Palestra 1 do livro *Emergent Evolution (The Gifford Lectures)*: Emergência
C. Lloyd Morgan

TRADUTORA: CRISTIANE XEREZ BARROSO

O ACENDEDOR DE LAMPIÕES

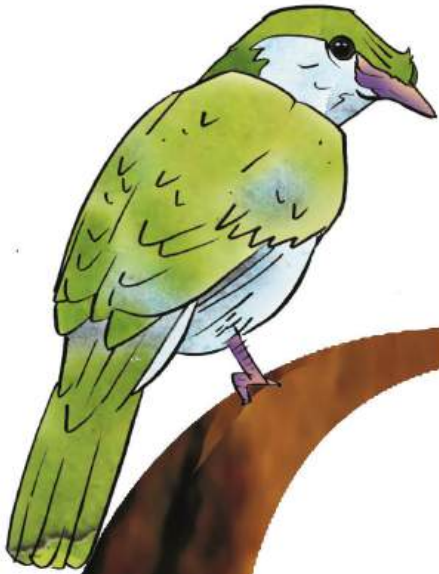
Jorge de Lima

Lá vem o acendedor de lampiões da rua!
Este mesmo que vem infatigavelmente,
Parodiar o sol e associar-se à lua
Quando a sombra da noite enegrece o poente!

Um, dois, três lampiões, acende e continua
Outros mais a acender imperturbavelmente,
À medida que a noite aos poucos se acentua
E a palidez da lua apenas se pressente.

Triste ironia atroz que o senso humano irrita: —
Ele que doira a noite e ilumina a cidade,
Talvez não tenha luz na choupana em que habita.

Tanta gente também nos outros insinua
Crenças, religiões, amor, felicidade,
Como este acendedor de lampiões da rua!



APRESENTAÇÃO



O homem é visivelmente feito para pensar.
Aí reside toda a sua dignidade e todo o seu mérito,
e todo o seu dever é pensar com acerto.

(Pascal, Pensamentos Escolhidos, 60)

É com grande alegria e satisfação que publicamos estas duas partes do **Dossiê Filosofia da Mente & Ciências Cognitivas** pela *Lmpião – Revista de Filosofia*, vinculada ao Programa de Pós-Graduação em Filosofia (PPGFil) da Universidade Federal de Alagoas (UFAL).

Publicar este Dossiê expressa o trabalho exitoso de dezenas de pesquisadores(as) de Filosofia do Brasil e de outros países em nosso jovem periódico. Expressa também a constância na investigação de tópicos específicos e importantes da linha de pesquisa Linguagem e Cognição do Programa, através do Grupo de Pesquisa Linguagem e Cognição. Este número especial mantém nosso trabalho de integrar estudiosos(as) de Filosofia publicando temáticas ligadas à filosofia da mente, às ciências cognitivas, à filosofia da linguagem, à lógica, à epistemologia etc.

As várias iniciativas do Grupo caminham na direção do significado da epígrafe de Pascal, trabalhar para ser digno do “pensamento com acerto”, especialmente em áreas com tanta heterogeneidade teórica como a filosofia da mente e as ciências cognitivas. Assim, promover encontros, manter grupos, ensinar e, sobretudo, divulgar o pensamento filosófico pela publicação acessível de temáticas ligadas à mente, à linguagem, à lógica e ao conhecimento são modos de ação de cumprir o “dever” de “pensar com acerto”.

Desde 2012, promovemos grupos de estudo, realizamos 10 encontros de pesquisadores(as), publicamos 6 livros de coletâneas, organizamos um número temático para *Prometheus – Journal of Philosophy* (UFS) e, agora, apresentamos estas duas partes da *Lmpião – Revista de Filosofia*.

Ilustrado pelo artista cearense André Dias, este número especial conta com mais de 30 textos – entre artigos, resenhas e traduções. Neles, o(a) leitor(a) terá a oportunidade de tomar contato e refletir sobre uma gama bastante variada de temas, problemas, autores(as) e teorias da filosofia da mente e das ciências cognitivas. Lendo e refletindo criticamente sobre estes trabalhos, pensamos que os(as) leitores(as) podem nos ajudar a cumprir o compromisso de “pensar com acerto”, pois somente a publicidade e a avaliação crítica compartilhada podem indicar o que esta expressão significa.

Por fim, agradecemos a todos(as) colaboradores(as), saibam que sua atividade fomenta diretamente o pensamento filosófico em Alagoas, no Nordeste, no Brasil e em muitos lugares do mundo. Em nome do Grupo de Pesquisa Linguagem e Cognição, agradecemos ao PPGFil-UFAL e aos Professores Fernando Monegalha e Rodrigo Barros Gewehr (Editores-Chefes da *Lampião*) pela confiança no trabalho desta publicação.

Dedicamos nosso número especial ao Jorginho e ao Rudá e desejamos a todos(as) uma ótima e proveitosa leitura!

Marcus José Alves de Souza

(Professor do PPGFil-UFAL)

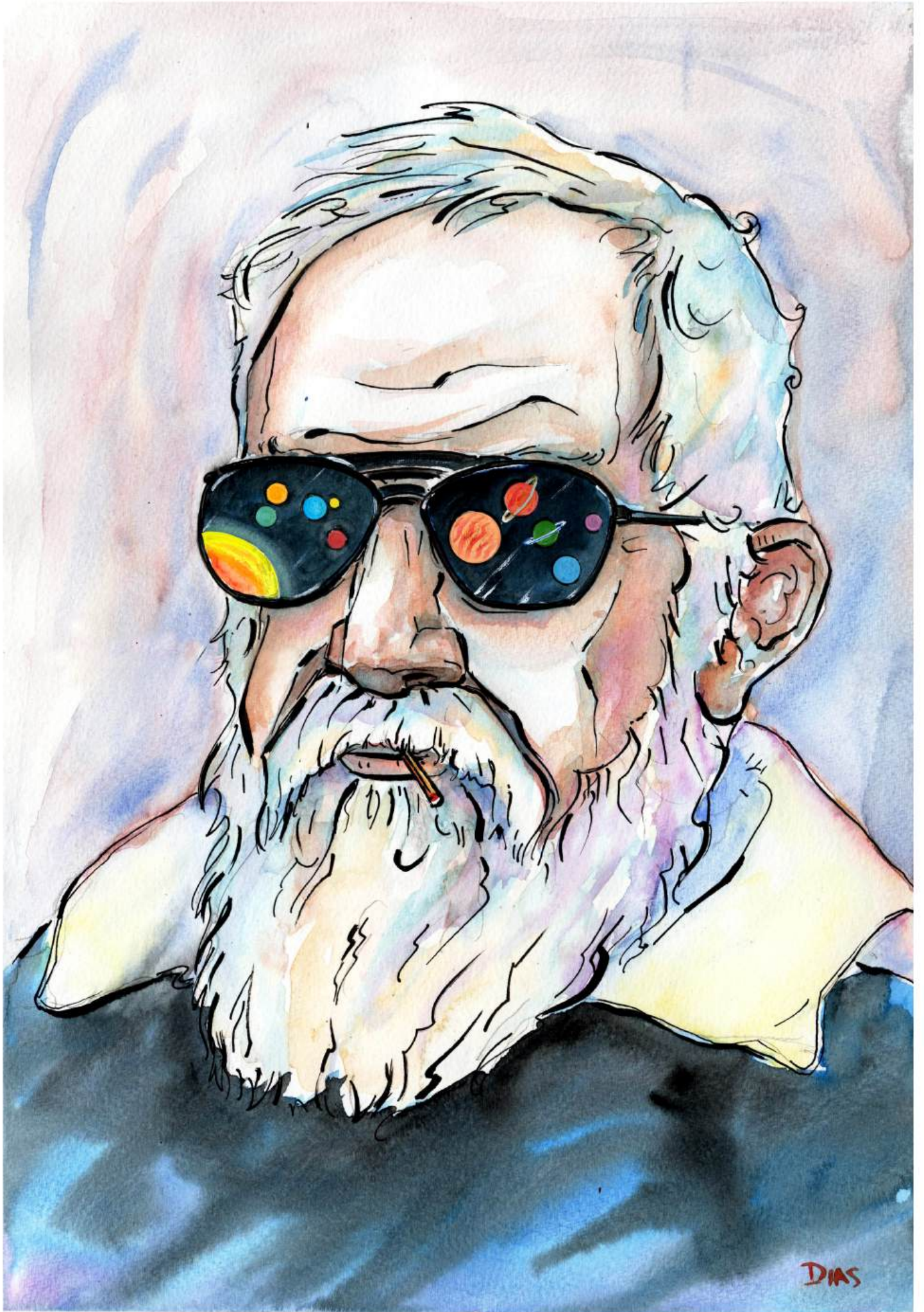
Argus Romero Abreu de Moraes

(Professor Visitante da FAALC/UFMS)

Maxwell Moraes de Lima Filho

(Professor da UFCA)





DIAS

A TESE DA IDENTIDADE MENTE-CORPO E O FISICISMO QUALITATIVO



Oswaldo Pessoa Jr.¹

Resumo:

A tese da identidade mente-corpo parece ser necessária para qualquer concepção materialista da consciência. Iniciamos com um apanhado histórico de sua presença na filosofia da mente, e um resumo de quais concepções aceitam ou negam a tese da identidade. Analisamos então diferentes tipos de teses de identidade, destacando-se a distinção entre identidade extensional e intensional. Apresenta-se então o fisicismo qualitativo, ou tese do encéfalo colorido, que identifica *qualia* e propriedades intrínsecas da matéria encefálica, comparando-o com as abordagens de Feigl e Pepper. Argumenta-se que U. T. Place não defende uma tese de identidade, mas sim um reducionismo. Desconstruímos também o argumento da múltipla realizabilidade de Putnam, e indicamos como a tese da identidade se liga com o isomorfismo psicofísico de Köhler. Na conclusão, re-

Abstract:

The mind-body identity thesis seems to be necessary for any materialist conception of consciousness. We begin with a historical overview of its presence in the philosophy of mind, and a summary of which conceptions accept or deny the identity thesis. We then analyze different types of identity theses, highlighting the distinction between extensional and intensional identity. We then present qualitative physicalism, or the colored-brain thesis, which identifies qualia and intrinsic properties of brain matter, comparing it with the approaches of Feigl and Pepper. It is argued that U.T. Place does not defend an identity thesis, but reductionism. We also deconstruct Putnam's multiple realizability argument, and indicate how the identity thesis is connected with Köhler's psychophysical isomorphism. In conclusion, we return to the distinction between structure and

¹ Professor Titular do Departamento de Filosofia, FFLCH, Universidade de São Paulo. E-mail: opessoa@usp.br

tomamos a distinção entre estrutura e qualidade e defendemos que propriedades mentais estruturais seriam idênticas a estruturas encefálicas, ao passo que os *qualia* seriam idênticos a propriedades intrínsecas, não estruturais e não mensuráveis com instrumentos físicos, da matéria encefálica.

Palavras-chave:

Identidade mente-corpo. Fisicismo qualitativo. Isomorfismo psicofísico. *Qualia*. Estrutura. Múltipla realizabilidade.

quality, and argue that structural mental properties are identical to brain structures, while qualia are identical to intrinsic properties, non-structural and not measurable with physical instruments, of brain matter.

Keywords:

Mind-body identity. Qualitative physicalism. Psychophysical isomorphism. Qualia. Structure. Multiple realizability.

1. Tese da identidade mente-corpo: visão geral

A tese da identidade mente-corpo considera que qualquer estado mental é *numericamente idêntico* a um estado corporal, restando explicar porque eles têm propriedades aparentemente distintas. O mais antigo registro da tese da identidade alma-corpo parece ter sido a concepção materialista dos atomistas Leucipo e Demócrito, que Aristóteles (*De Anima*, I.2, 403b33) expressou ao escrever que, para estes, “os [átomos] de forma esférica são fogo e alma”. Sendo um monismo, o dualismo de atributos de Espinosa é apontado por David Skrbina (2005, p. 8-9) como a primeira teoria da identidade da história.

Influenciado pela concepção de Espinosa, Gustav Fechner considerou em 1860 que mente e cérebro são duas perspectivas de algo único, e chamou assim sua teoria de “visão de identidade”. Além desta aceção metafísica, a Fechner é associada a origem do termo “paralelismo psicofísico” (HEIDELBERGER, 2004, cap. 5), que foi bastante usado no período 1860-1915 para designar a constatação empírica de que há uma correlação entre os eventos mentais e os eventos corporais, sem uma explicação adicional de porque isso acontece.

Aqui e ali encontramos afirmações da tese da identidade, como em George H. Lewes (1875, p. 412), para quem o paralelismo psicofísico desemboca em uma identidade:

“[...] há indícios contundentes de que a função varia com as variações nas condições físicas; em outras palavras, de que o fenômeno senciante é um fenômeno nervoso”. Em 1918, Moritz Schlick formulou uma versão da tese da identidade:

Mas todas essas complicações do retrato de mundo são desnecessárias. Elas podem ser facilmente evitadas se, no lugar da suposição dualista, introduzirmos a hipótese muito mais simples de que os conceitos das ciências naturais são adequados para designar toda realidade, incluindo aquela que é imediatamente vivenciada. A relação resultante entre a realidade imediatamente vivenciada e os processos do cérebro físico é então não mais uma de dependência causal, mas de simples *identidade*. O que temos é uma e a mesma realidade, não “vistas de dois lados diferentes” ou “manifestando-se de duas formas diferentes”, mas designadas por dois sistemas conceituais diferentes, o psicológico e o físico. (SCHLICK, [1918] 1974, p. 299).

Todo defensor da tese da identidade precisa explicar porque mente e corpo aparecem de maneiras diferentes, ou seja, precisa dar conta da chamada “lacuna explicativa”. Vemos que a explicação de Schlick substitui a doutrina das duas perspectivas de Fechner por “dois sistemas conceituais diferentes”. Retornaremos a isso, quando examinarmos a “tese da linguagem dupla” de Feigl, na seção 6.

Seria de se esperar que qualquer visão *monista* trabalhe com alguma identidade: o materialismo identifica a mente com parte do corpo, o idealismo identifica o corpo com parte da mente, o dualismo de atributos identifica as substâncias subjacentes à mente e ao corpo. Porém, para um monismo neutro tradicional² não há exatamente uma identidade mente-corpo. O que há é uma identidade entre os *elementos* que constituem a mente e os elementos que constituem o corpo, elementos esses que seriam “perceptos” ou “dados dos sentidos”. Mas o corpo envolve uma certa ordem de coextensão e sucessão destes elementos, que seria distinta da maneira com que os elementos se ordenam na mente consciente. Assim, não haveria uma identidade entre consciência e corpo, apesar de a concepção ser monista.

² Mais recentemente, o termo “monismo neutro” passou a se referir a qualquer monismo que seja neutro em relação aos polos mente e matéria, como o dualismo de atributos de Espinosa. Aqui nos restringimos ao monismo neutro “tradicional” de Mach, Avenarius, Petzoldt, James, Perry, Holt, Schlick e Russell (cf. STUBENBERG, 2016).

2. Concepções que negam a tese da identidade

Vale a pena mencionar as concepções que rejeitam a identidade mente-encéfalo. Em primeiro lugar, é claro, qualquer dualismo de substância ou espiritualismo rejeita a tese da identidade (DESCARTES, [1641] 1973, cap. VI, § 17). Dualistas de propriedades como Chalmers (1996, p. 162), que consideram que o corpo é físico e a mente não física, também a rejeitam. Mesmo materialismos emergentistas tendem a rejeitar a tese da identidade (SPERRY, 1980, p. 136-8). Um monismo materialista pode eliminar o conceito de mente, como no eliminativismo sugerido por Paul Feyerabend (1963), então neste caso não teríamos uma tese de identidade (pois um dos termos foi eliminado). Um materialismo reducionista tenderia a aceitar a tese da identidade, mas se o estatuto ontológico da consciência for diminuído (aproximando-se do eliminativismo), não haveria propriamente identidade (este é o caso de U. T. Place, 1956, como argumentarei na seção 5).

O “epifenomenismo” é a concepção segundo a qual a consciência existe, mas por si só não exerce nenhuma influência causal sobre o corpo. Esta concepção foi defendida pelo biólogo Thomas Huxley (1874), que a chamou de “teoria do autômato”. À primeira vista, diríamos que um epifenomenista não deveria aceitar a tese da identidade, pois se os estados mentais fossem idênticos a estados corporais, eles (os estados mentais) também teriam os poderes causais atribuídos aos estados corporais, negando a tese epifenomenista. Porém, veremos mais para frente uma concepção que considera que propriedades mentais qualitativas são idênticas a propriedades quimiofísicas qualitativas intrínsecas de uma certa região do encéfalo, propriedades intrínsecas estas que seriam distintas das relações causais (estruturais, extrínsecas) associadas a esta mesma região, sendo que apenas estas relações estruturais são passíveis de mensuração. Tal concepção seria epifenomenista com respeito às propriedades qualitativas, mas não com respeito às outras propriedades mentais que seriam idênticas às relações estruturais (aquelas descritas pelo paralelismo psicofísico, por exemplo na versão mais detalhada de KÖHLER, 1943, p. 61-2, o “isomorfismo psicofísico”, que veremos na seção 8).

O “pampsiquismo” é a tese de que todas as coisas estão imbuídas de mente (psiquismo) ou de alma, ou mesmo apenas de senciência ou de sentimento: animais, plantas, minerais, moléculas, átomos etc. Se tais propriedades presentes em todas as coisas não forem propriamente “mentais”, mas forem consideradas “protomentais” ou “precursores da consciência” (CHALMERS, 2015, p. 248), temos o “pamprotopsiquismo”.

Há diversas teorias da identidade que são pampsiquistas. Já mencionamos que Skrbina (2005) considera o dualismo de atributos de Spinoza como uma teoria da identidade, mas o filósofo sefardita também é considerado um pampsiquista. Se a dualista de atributos for uma pampsiquista, deve então considerar que a substância subjacente a qualquer parte da matéria, mesmo uma pedra, possui também um atributo de mentalidade. Tal conclusão contraintuitiva pode ser amenizada em um pamprotopsiquismo: neste caso, a substância da pedra não teria propriamente uma mente, mas propriedades “protomentais”. Tal noção de protopsiquismo aparece claramente em William Clifford (1878), com o conceito de *mind-stuff* (material-mental), mas sem defender explicitamente uma tese da identidade. Influenciado por Clifford, o neorrealista estadunidense Durand Drake (1933, p. 379) acabou abandonando o pampsiquismo e articulando um pamprotopsiquismo que chamou de “teoria da identidade”, afirmando que “mente e encéfalo são idênticos”.

Quanto ao funcionalismo, pode-se dizer que ele é uma espécie de tese de identidade, mas não é uma identidade mente-*encéfalo*, mas sim uma identidade entre mente e uma classe de estados de organização dinâmica complexa da matéria. Nas palavras de Skrbina (2005, p. 9), o funcionalismo

argumenta que estados mentais são reais e que eles são idênticos a um particular ‘estado de processo’, ou estado de informação. [...] Assim, o funcionalismo pode ser visto como uma espécie de generalização da teoria da identidade: não somente um encéfalo, não apenas um sistema nervoso, mas qualquer sistema físico é capaz de gerar um estado mental.

Mais adiante, analisaremos uma crítica feita pelo funcionalismo à tese de identidade mente-encéfalo, o argumento da múltipla realização (seção 7).

3. Tipos de teses da identidade

(a) Identidade numérica e identidade de propriedades

Como é bem sabido, há na Filosofia uma distinção fundamental entre duas acepções do termo “identidade”. *Identidade numérica* designa a situação em que se trata do mesmo indivíduo. No exemplo clássico de Gottlob Frege ([1892] 2011), a estrela d’alva e a estrela Vésper são idênticas, no sentido numérico, pois ambas denotam o mesmo indivíduo, o planeta Vênus, visto de manhã ou ao entardecer. Por outro lado, dois irmãos

gêmeos univitelinos são (quase) idênticos no sentido de que eles possuem as mesmas propriedades, apesar de serem indivíduos diferentes. Esta pode ser chamada *identidade de propriedades*³.

(b) Identidade de formas e qualidades

O debate filosófico clássico entre “realistas de universais” e “nominalistas” envolve uma discussão sobre o tipo de identidade que se aplica a propriedades. Por exemplo, suponha que dois gêmeos univitelinos tenham narizes “idênticos”. Claramente, isso deve ser entendido como identidade no sentido de propriedades. Mas surge aqui uma pergunta: pode-se dizer que a *forma* instanciada em ambos os narizes possui identidade numérica? É a mesma forma (no sentido numérico), ou seja, uma e a mesma forma? Em caso afirmativo, poderíamos concluir que esta forma tem existência independente da matéria? Um realista de universais, como Platão, diria que sim, ao passo que um nominalista (com relação à geometria) como Guilherme de Ockham, diria que não (cf. LOUX, 2006, caps. 1, 2).

Poderíamos tentar deixar de lado estas duas concepções a respeito de termos universais, mas ela é uma distinção que está por trás de alguns debates em Filosofia da Mente, por exemplo, o de se uma propriedade relacional pode ser considerada uma propriedade material. Introduziremos a seguir duas outras distinções e, com elas, tentaremos navegar ao largo da querela dos universais.

(c) Identidade de instância e identidade de espécie

Há uma distinção tradicional entre identidade de instância (*token*, em inglês) e identidade de espécie (ou tipo, do inglês *type*). No ataque funcionalista às teorias da identidade mente-encéfalo de Feigl, Place e Smart, argumentou-se que seria plausível supor que uma *instância* de um *quale* de azulidão vivenciado por alguém em um determinado momento é idêntico à instância do estado encefálico correspondente; porém, não seria plausível supor que a *espécie* de experiência subjetiva da azulidão vivenciada por seres humanos e marcianos (supondo que esses seres construídos de maneira distinta tenham a “mesma” experiência de azul) seja idêntica a uma espécie de estado material, justamente porque os estados materiais de humanos e marcianos seriam completamente distintos.

³ Tradicionalmente esta é chamada de “identidade qualitativa”, mas queremos reservar a palavra “qualitativo” para designar as qualidades subjetivas.

Na seção 7, propomos uma desconstrução de tal argumento, baseado na impossibilidade de saber “como é que é ser” um marciano. Mas a própria tese de identidade de instância precisa ser melhor elaborada, como argumentaremos a seguir.

(d) Identidade extensional e intensional

Há outra distinção tradicional importante, aquela entre identidade “extensional” e “intensional” (com “s”) (por exemplo, FEIGL, 1967, p. 73, 145). Um exemplo clássico é dizer que o ser humano é idêntico a um bípede implume: tal identidade seria meramente extensional, pois de fato (ignorando exceções) todo ser humano é um bípede implume e todo bípede implume é um ser humano; no entanto, o significado dos termos é distinto, de maneira que não há identidade intensional. Dizer que água é H_2O é considerado uma identidade extensional por Feigl, mas não intensional.

(e) Identidade extensional de instância

Combinando as distinções dos itens (c) e (d), obtemos a noção de uma “identidade extensional de instância”, que aplicada a dois termos indica que esses dois termos referem-se a um mesmo lugar do espaço e do tempo. Dizer que um elemento mental é extensionalmente idêntico a um elemento fisiológico é dizer que ambos ocupam o mesmo lugar do espaço e do tempo, apesar de poderem constituir propriedades diferentes. Por exemplo, uma materialista localizacionista poderia afirmar que um *quale* de azulidão em um ponto do campo visual é extensionalmente idêntico a um agrupamento de neurônios em nosso (hipotético) “sensório” visual.

Por outro lado, a *identidade intensional* de instância de dois termos exige que os dois termos tenham o mesmo significado, ou refiram-se a um único objeto ou propriedade do mundo. Este é um conceito mais difícil de definir. No exemplo de Frege, pode-se defender que “estrela d’alva” e “estrela Vésper” têm significados diferentes, pois referem-se a Vênus observada em diferentes momentos do dia. Mas se estivermos interessados no objeto visto em diferentes contextos, podemos considerar que o significado (ou sentido) dos dois termos é idêntico. Por trás dessas duas opiniões parece haver concepções distintas sobre a realidade: a primeira é mais holista (toda propriedade depende do contexto em que o objeto está inserido) e a segunda é mais separabilista (há propriedades de objetos individuais).

Simplificando ainda mais a discussão, *considerarei a identidade intensional apenas em objetos com identidade extensional*. Com isso, pretendo deixar de lado, por ora, a discussão envolvendo a identidade intensional de formas ou qualidades instanciadas em diferentes sistemas físicos (localizados em diferentes pontos do espaço ou do tempo). Voltemos ao exemplo da materialista localizacionista: aceitando que a instância do *quale* de azulidão em um ponto do campo visual é extensionalmente idêntico a um agrupamento de neurônios no sensorio visual, poderíamos defender que esses termos são intensionalmente idênticos?

A resposta negativa a esta última questão constitui uma defesa da *lacuna explicativa*, pelo menos da maneira que entendo a expressão “lacuna explicativa”. O argumento clássico que nega a identidade intensional, presente por exemplo em C. D. Broad (1925, p. 622-3), considera que uma expressão como “um agrupamento de neurônios” deve explicitar as propriedades fisicamente mensuráveis deste agrupamento, o que envolverá medições elétricas, químicas e de imageamento preciso. Claramente, tais medições, redutíveis a quantidades (números e geometria), não capturam a experiência de uma qualidade subjetiva, como o *quale* da azulidão, vivenciada em um certo momento por uma pessoa.

Por outro lado, aceitamos o isomorfismo psicofísico (KÖHLER, 1943, p. 61-2), que permite identificar posição, instante e intensidade de um *quale* com propriedades fisicamente mensuráveis em um agrupamento de neurônios (ver seção 8). No entanto, o *quale* em si mesmo não se apresenta em uma medição física convencional (cf. PESSOA, 2019). No entanto, defenderemos a seguir que o *quale* é intensionalmente idêntico a uma qualidade quimiofísica intrínseca da matéria encefálica.

4. A tese do encéfalo colorido

“A tese do encéfalo colorido” (*the colored-brain thesis*) é o nome dado por Leopold Stubenberg (1998, p. 169) para a concepção de que as qualidades fenomênicas subjetivas, ou *qualia*, são “propriedades do encéfalo”. H. H. Price (1932, p. 127) se referiu a esta tese como “a hipótese de que os dados dos sentidos [*sense data*] são cerebrais [*cerebral*]”. Price (1932) apontou que o filósofo hegeliano F. H. Bradley achou difícil acreditar que “quando cheiro um cheiro estou ciente de um estado fedido de meu próprio sistema nervoso”, criticando a teoria proposta pelo filósofo de Oxford Thomas Case (1888, p. 33), que caracterizara a percepção sensitiva como “a apreensão imediata de um objeto físico interno dentro do sistema nervoso de um ser senciente”. Sua posição, todavia, não era materialista

ou fisicista, pois considerava que a apreensão final do objeto físico interno era feita pelo “psíquico” (CASE, 1888, p. 24)⁴.

A reação à tese do encéfalo colorido é considerá-la absurda, como fez Bradley, no sentido de que seria um óbvio equívoco supor que “para se ver vermelho é preciso que haja neurônios vermelhos no encéfalo” (O’REGAN & NOË, 2001, p. 947, 1010, 1018). A resposta a esta objeção envolve aceitar a noção de que a cor subjetiva não é uma propriedade da luz, mas uma propriedade químico-física do encéfalo (no estado de consciência visual) (ver discussão em PESSOA, 2021, p. 86). Em outras palavras, essa alteração de perspectiva é a mudança do externismo do senso comum (“a verdidão está no abacate”) para o internismo de concepções como a teoria dos dados dos sentidos ou a tese da realidade dos *qualia* (“a verdidão está na mente”). Soma-se a isso a tese da identidade mente-encéfalo extensional, e tem-se que “a verdidão está no encéfalo”.

Ao que parece, a tese do encéfalo colorido não foi explicitamente defendida por mais ninguém, apesar de Price (1932, p. 127) escrever que “filósofos estão acostumados a discutir a questão de se os dados dos sentidos são físicos ou mentais”. De fato, no período do Entre Guerras, em que a teoria dos dados dos sentidos e o materialismo conviviam na filosofia de língua inglesa, o psicólogo e historiador da psicologia de nacionalidade estadunidense Edwin Boring se aproximou muito da tese do encéfalo colorido. Foi esta obra que influenciou U. T. Place (1956) a desenvolver sua versão da tese da identidade mente-encéfalo (cf. PLACE, 2000).

A tese da identidade mente-encéfalo foi enunciada por Boring (1933, p. 16) na seguinte afirmação, citada por Place (2000), que exhibe a ampliação do paralelismo para a identidade: “Para o autor, uma correlação perfeita é identidade. Dois eventos que sempre ocorrem juntos no mesmo tempo e no mesmo lugar, sem nenhuma diferenciação temporal ou espacial, não são dois eventos mas o mesmo evento”. A seguir, Place pondera por que Boring foi ignorado pelos filósofos:

Além disso, Boring ele mesmo estava comprometido em combinar a teoria da identidade com um relato fenomenista de qualidades sensoriais, que de acordo com o princípio da *Identidade dos Indiscerníveis* de Leibniz o comprometeria com a concepção de que certos eventos cerebrais são literalmente verdes, [sonoramente] agudos, quentes, azedos ou pútridos, o que para um filósofo constituiria uma imediata redução ao absurdo, fatal para sua posição (PLACE, 2000, p. 1).

4 Para mais detalhes, ver Pessoa Jr. (2021), da qual é extraída a presente seção.

Vemos assim que Boring se aproximou da tese do encéfalo colorido, que podemos chamar também de “fiscismo qualitativo”, apesar de ele não a ter defendido explicitamente. Nossa posição é de aceitação da tese do encéfalo colorido, e de que ela é a única maneira de defender a tese da identidade mente-encéfalo intensional.

5. Place e a tese da identidade

A tese da identidade mente-encéfalo, da maneira que foi formulada pelo psicólogo inglês Ullin T. Place (1956), não coloca mente e cérebro em pé de igualdade, mas privilegia a descrição teórico-científica do encéfalo, atribuindo à vivência subjetiva da consciência um estatuto ontológico menor. Tal posição pode ser considerada um eliminativismo parcial, pois os dados dos sentidos são eliminados. Isso é expresso pelo que Place (1956) chama de “falácia fenomenológica”, que seria identificar numa pós-imagem verde o *quale* da verdidão (tal pós-imagem pode ser gerada ao se olhar detidamente o interior de uma coisa vermelha e depois dirigir o olhar para uma parede branca). Para Place, educado na tradição behaviorista, o correto seria apenas dizer que uma pós-imagem verde evoca os mesmos processos cerebrais que a visão da grama (por exemplo), e não que existiria uma realidade em nível fenomenológico que correspondesse ao *quale* da verdidão.

Em texto posterior, já mencionado, Place (2000) explica que não encarou a tese da identidade no sentido em que Frege identificou numericamente a ‘estrela d’alva’ e ‘estrela Vésper’ com o planeta Vênus. Feigl e Smart adotariam esta abordagem, mas Place (2000) defendeu o uso do “é da composição”, ou seja, a afirmação de que a consciência é composta de processos cerebrais, e nada mais. Referindo-se a si mesmo em terceira pessoa, Place (2000) escreve:

A afirmação de Place não era que a consciência seria idêntica a – ou exatamente a mesma coisa que – os processos encefálicos com os quais ela está correlacionada, mas que a consciência consiste inteiramente de – ou é composta inteiramente de – processos encefálicos. Em outras palavras, ele interpretou a relação das experiências com os processos encefálicos em termos da microrredução substancial de uma substância em suas partes constituintes em um nível de análise inferior ou mais microscópico [...].

Além disso, ressalta que:

a tese da identidade só se aplica para certos aspectos da vida mental: a consciência (Place), os sentimentos brutos da experiência (Feigl), as sensações (Smart). Os aspectos cognitivos e volitivos (intencionais) da vida mental não são redutíveis a estados ou processos encefálicos; mas são (conceitualmente) redutíveis a alguma espécie de competência ou propensão semântica, lógica ou verbal.

Place sugeriu a analogia da identidade consciência-cérebro com a identidade entre um raio e uma descarga maciça de eletricidade atmosférica. No sentido de “composição”, é razoável dizer que um raio é composto de cargas elétricas em movimento (i.e. uma descarga). Mas o sentido de “identidade” discutido aqui, no presente artigo, é justamente o de identidade numérica usado por Frege. Neste caso, podemos até concordar que um raio seja idêntico a um grande número de cargas elétricas em movimento. Mas a analogia desta identidade raio-descarga com a identidade entre consciência e parte do encéfalo apresenta uma diferença crucial: tanto raio quanto “um movimento de cargas elétricas” são *descrições teóricas* de um processo físico, ao passo que na discussão sobre a relação entre consciência e encéfalo se está supondo que “consciência” designe o conjunto de experiências subjetivas, vivenciadas por contato (*acquaintance*), e não um conhecimento proposicional, como aqueles envolvidos na descrição teórica do mundo cotidiano e da ciência. Assim, há uma diferença essencial entre a identidade entre raio e descarga, por um lado, e a putativa identidade entre consciência e encéfalo, por outro.

Em suma, apesar de Place (1956, p. 48) invocar “a identidade entre a consciência e certos processos no encéfalo”, inspirado na sugestão anterior de Boring (1933), ele usa o termo “identidade” para designar a redução de processos conscientes a processos encefálicos, no sentido de que a consciência seria “composta” de processos encefálicos, e nada mais. Assim, a concepção de Place é melhor designada por “reducionismo”, que para certos aspectos da consciência tende a um eliminativismo, como em sua crítica ao “campo fenomênico” (PLACE, 1956, p. 49). O encéfalo conservado na Universidade de Adelaide não pertenceu a um defensor da tese da identidade mente-encéfalo.

6. Feigl e Pepper

Em artigo publicado em 1963, Herbert Feigl discutiu uma “objeção familiar” à teoria da identidade mente-corpo, que é exatamente a questão da lacuna explicativa: “como

é que qualidades diretamente vivenciadas, como cores, sons, cheiros, dores, emoções ou semelhantes, poderiam ser idênticas a processos neurais, cujas propriedades são tão fundamentalmente diferentes?” (FEIGL, 1963, p. 328-9). Feigl aceita a tese da realidade dos *qualia*, ao contrário de Place e Smart, que tendem a rejeitá-la⁵.

Feigl (1967, p. 142) também adota explicitamente o “estruturalismo sobre a física”, seguindo Schlick e Russell. Esta é a tese de que a Física só tem acesso às propriedades relacionais das coisas, não às coisas em si (não aos *relata* das relações). O termo “estrutura” designa o conjunto de relações (causais e de outras espécies) envolvendo um objeto. O acesso à estrutura do mundo seria feito de maneira indireta, inferencial.

Ao longo de sua carreira, Feigl assumiu diversos pontos de vista, oscilando entre diferentes graus de materialismo. Em 1958, Feigl defendera uma “teoria da linguagem dupla” (*double-language theory*), segundo a qual haveria duas linguagens usualmente empregadas em psicologia, a física e a fenomênica, que podem ser *traduzidas* uma na outra (FEIGL, 1967, p. 23, 79-80). Stephen Pepper deu o seguinte exemplo de tradução entre as duas linguagens (adaptado de PEPPER, 1967, p. 76-7): um clarão visual subjetivo (linguagem fenomênica) pode também ser descrito como a reflexão de um foco de luz que incide na retina, despertando uma cadeia de impulsos neurais (linguagem física). Tal “versão linguística” da tese da identidade (ou seja, a teoria da linguagem dupla) supõe assim que os dois enunciados têm o mesmo referente, mas não se pergunta qual é este referente (PEPPER, p. 84).

Feigl acabou abandonando esta versão porque os referentes dos enunciados, no final das contas, não seriam os mesmos. Segundo a tese de Leibniz da indiscernibilidade dos idênticos, se mente e encéfalo forem idênticos, eles devem ser indistinguíveis, o que não é o caso (FEIGL, 1967, p. 145). O referente direto da linguagem fenomênica seriam os “sentimentos brutos” (*raw feels*), ou *qualia*. E a linguagem física? Na nova versão em Feigl (1963), ela se referiria a “símbolos descritivos”, e não, como esperaria um realista, a uma realidade física. Neste ponto, Feigl se afasta da tese da identidade extensional de instância (seção 2e), pois não afirma que uma mancha no campo visual subjetivo (por exemplo) esteja localizada em uma região restrita do espaço físico real, dentro do encéfalo:

Assim, não há conflito e nem incompatibilidade com respeito à “localização”, por exemplo, de uma mancha de cor [*patch of color*] diretamente experienciada. Ela está onde a “vemos” no espaço fenomênico. Ao processo encefáli-

5 A discussão desta seção é semelhante àquela feita em Pessoa Jr. (2021, p. 90-1).

co sistematicamente idêntico é assinalado um local na variedade [*manifold*] 3-dimensional do espaço físico; e uma análise detalhada do processo central em suas relações com impulsos aferentes e eferentes deve dar conta do comportamento relevante no aprendizado de lugar [*place learning*], na orientação espacial, em ilusões de óptica etc. (FEIGL, 1963, p. 331).

A ênfase nesta citação é que (1) a mancha colorida está localizada em um espaço fenomênico, distinto do espaço físico; e (2) de maneira abstrata, associa-se esta localização a uma posição no espaço físico. Esta concepção é distinta da defendida no presente trabalho. Aqui, consideramos que a mancha está localizada, antes de tudo, no espaço físico real. O campo fenomênico é constituído por um conjunto de pontos e manchas, cada qual também bem localizado no espaço real, constituindo o chamado “sensório” visual. Assim, o que Feigl chama de “espaço fenomênico” é (idêntico a) uma região do espaço físico real localizado no sensório visual, uma região do encéfalo que implementa o “materialismo cartesiano” criticado por Dennett (1991, p. 207), e possivelmente localizado no tálamo. A partir do campo fenomênico, “representamos” uma cena do mundo externo, mas esta questão da intencionalidade não se coloca neste debate com Feigl, sendo uma questão a ser considerada posteriormente.

Já a posição do psicólogo e filósofo Stephen Pepper se aproxima mais da nossa. Pepper (1967) herda a versão de 1963 de Feigl da tese da identidade, que combina realismo dos *qualia* com estruturalismo na física interpretado de modo antirrealista. Porém, diferentemente de Feigl (1963), ele caracteriza sua posição como uma “teoria da identidade neural *qualitativa*” (PEPPER, 1967, p. 76), mais próxima do fisicismo qualitativo. Ele adota o que chamamos de “inferência pamprotopsiquista” e considera que há qualidades fora da mente, visão chamada de “panqualitismo” (*pan-quality-ism*) (FEIGL, 1971, p. 308).

Pepper transforma o problema mente-corpo no problema qualidade-conceito (ou qualidade-estrutura), “um problema de atualidade qualitativa e várias descrições simbólicas dela” (p. 92). Pepper (1967, p. 93) está imbuído do impacto que a neurociência teve sobre o pensamento filosófico no pós-guerra, citando o neurocirurgião Herbert Penfield, ao constatar que o processo qualitativo está localizado no encéfalo:

Como exatamente os termos físicos fazem contato com um processo qualitativo atual? A teoria da identidade traz esta questão para o foco. Levando a sério resultados fisiológicos e observacionais de especialistas, ela mostra exatamente onde a qualidade atual e o conceito físico se encontram. Eles se encontram no que é fisiologicamente indicado como áreas do encéfalo humano, ou talvez mais estreitamente, no córtex cerebral.

Esta afirmação de Pepper exprime a tese da identidade extensional de instância. Já a tese intensional é resolvida admitindo-se que há qualidades na realidade física.

7. Argumento da múltipla realização

Nos anos 1960, o funcionalismo se colocou contra a tese da identidade mente-encéfalo. Isso se deu na forma do “argumento da múltipla realização”, apresentado pelo matemático e filósofo Hilary Putnam (1967 [2005]). Ele contrasta a “teoria do estado-encefálico” (*brain-state theory*, o materialismo da tese da identidade) com a teoria do estado-funcional (funcionalismo):

Em suma, argumentarei que a dor não é um estado encefálico, no sentido de um estado físico-químico do encéfalo (ou mesmo de todo sistema nervoso), mas um *gênero* [*kind*] totalmente diferente de estado. Proponho a hipótese de que a dor, ou o estado de sentir dor, é um estado funcional de um organismo inteiro (PUTNAM, [1967] 2005, p. 162).

Putnam (1967 [2005]) *define* “dor” como um estado funcional, associado a um certo comportamento. Ao vermos uma pessoa se debatendo em dor e um polvo sendo fígado, é razoável supor que ambos estão no mesmo estado funcional. Entretanto, não temos acesso ao estado mental subjetivo do polvo para saber se o *quale* associado à sua dor é perfeitamente semelhante ao *quale* da dor do ser humano:

Tomando a hipótese do estado-encefálico desta maneira, então quais razões haveria para preferir a hipótese do estado-funcional ao invés da hipótese do estado-encefálico? Considere o que o teórico do estado-encefálico tem que fazer para validar suas afirmações. Ele tem que especificar um estado físico-químico de tal forma que *qualquer* organismo (não apenas um mamífero) esteja com dor se e somente se (a) ele possuir um cérebro de uma estrutura físico-química apropriada; e (b) seu encéfalo estiver naquele estado físico-químico. Isso significa que o estado físico-químico em questão precisa ser um estado possível de um encéfalo de mamífero, de um encéfalo reptiliano, de um encéfalo de molusco (polvos são moluscos, e certamente sentem dor), etc. (PUTNAM, [1967] 2005, p. 164).

Se para o funcionalista é mais ou menos óbvio que a “dor” de um ser humano e de um polvo são estados funcionais bastante semelhantes, para o materialista que atribui rea-

lidade para os *qualia* é um problema em aberto se o estado subjetivo qualitativo de ambos é perfeitamente semelhante. Esta posição materialista, que atribui realidade aos *qualia*, foi chamado por Chalmers (1996, p. 165-7) de “materialismo tipo B”, e claramente é a posição adotada no presente artigo. Um materialista “tipo A” tende a não atribuir importância aos *qualia*, e segundo o filósofo australiano incluiria “eliminativismo, behaviorismo, e várias versões do funcionalismo reduutivo” (CHALMERS, 1996, p. 165-7):

Finalmente, a hipótese se torna ainda mais ambiciosa quando nos damos conta de que o teórico do estado-cerebral não está dizendo apenas que *dor* é um estado cerebral; ele está, é claro, preocupado em defender que *todo* estado psicológico é um estado cerebral. Assim, se pudermos encontrar pelo menos um predicado psicológico que pode ser claramente aplicado tanto a um mamífero quanto a um polvo (digamos “com fome”), mas cujo “correlato” físico-químico é diferente nos dois casos, a teoria do estado-cerebral colapsa. Parece-me altamente provável que possamos fazer isso. Concede-se que, em tal caso, o teórico do estado-cerebral possa se salvar por meio de suposições *ad hoc* (por exemplo, definindo a disjunção de dois estados como sendo um único “estado físico-químico”), mas isso não precisa ser levado a sério (PUTNAM, [1967] 2005, p. 164-5).

O problema deste argumento é que não temos como saber qual é o estado subjetivo qualitativo do polvo. Um pouco de reflexão mostra que só poderíamos conhecer este estado por inferência a partir de leis-de-ponte psicofisiológicas (ainda não descobertas) que partam do estado físico-químico e funcional do polvo. A semelhança entre estados subjetivos, não diretamente mensuráveis, só pode ser estabelecida a partir de estados encefálicos semelhantes e das leis de ponte. Nesse sentido, para estados mentais qualitativos, o argumento de Putnam é inócuo⁶.

Aparentemente, a primeira geração de materialistas do pós-guerra (Place, Feigl, Smart e Armstrong) não se importou muito com o argumento da múltipla realização, que seria muito influente na nova Filosofia da Mente da década de 1970. Tanto que um filósofo próximo do fisicismo reduutivo como Jaegwon Kim (1992) se debruçou sobre o argumento de Putnam, aceitando suas consequências, e concluindo que a solução a partir da “disjunção” de estados não é tão ruim assim e poderia ser o caso. Ou seja, seria bem possível que o ser humano e o polvo sentissem a “mesma dor”, mesmo tendo estados encefálicos muito distintos.

6 O argumento desta seção é apresentado, de forma mais sucinta, em Pessoa Jr. (2015, p. 203-4).

Neste ponto, entra a distinção entre identidade mente-corpo como instância (*token*) ou como tipo ou espécie (*type*) (seção 3c). Um funcionalista pode aceitar que um certo estado mental em particular (ou seja, uma instância), por exemplo o da visão de um raio verde pela Sra. Bess no entardecer da Ilha de Staffa em 27/07/1882⁷, seja idêntico ao seu estado encefálico naquele momento (ou seja, aceita a tese da identidade para uma instância). Mas este funcionalista não aceita que todos os estados do particular tom de “jade celestial” instanciados no nosso Universo sejam realizados pelo mesmo estado físico-químico subjacente à visão da Sra. Bess (ou seja, rejeita a tese da identidade para tipos). A base física do *quale* de verde jade em cada espécie de animal ou de extraterrestres seria completamente diferente, e a tese da identidade para tipos só poderia ser defendida de maneira disjuntiva: “o estado mental verde jade é idêntico a tal estado encefálico em humanos, *ou* a tal outro estado em polvos, *ou* a um terceiro estado físico em extraterrestres”.

Mas como saber se um marciano vivencia o mesmo tom de verde jade em questão? Não dá para saber, a não ser a partir de uma teoria psicofisiológica bem corroborada que preveja, a partir do estado material de um marciano, que ele está vivenciando o verde jade. E isso só pode ser feito definindo-se uma espécie (tipo) de estado material, de maneira não disjuntiva.

Em outras palavras, há, em princípio, a possibilidade de que um *quale* de um ser extraterrestre seja qualitativamente idêntico a um *quale* humano, mesmo sendo gerado por bases materiais completamente distintas. Mas isso jamais poderia ser descoberto. Ou seja, isso só poderia ser estabelecido experimentalmente se houvesse algo em comum entre o estado encefálico do humano e o estado material do alienígena, pois só assim a teoria poderia fazer a previsão.

Esta situação tira a força do argumento da múltipla realização. Repetindo: se uma teoria psicofisiológica do futuro, consistente com o princípio de superveniência, prever que um alienígena possa ter um *quale* humano em uma base material diferente, terá que haver alguma semelhança entre os dois estados materiais (uma identidade para tipos) para explicar a semelhança mental e permitir a previsão feita pela teoria.

⁷ Referência à obra *Raio verde*, de Julio Verne (1882).

8. Isomorfismo psicofísico

Na discussão sobre o paralelismo psicofísico no séc. XIX, destaca-se a formulação de G. E. Müller (1896) da tese como uma correspondência um-para-um entre sensações subjetivas e estados encefálicos. Posteriormente, em 1929, o psicólogo gestaltista Wolfgang Köhler ([1929] 1943, p. 61-2) introduziu a expressão “isomorfismo psicofísico” para exprimir essa correlação entre as propriedades sistêmicas (estruturais, funcionais) da mente e do encéfalo (cf. BORING, 1942, p. 83-90, 95-6; SCHEERER, 1994).

Fizemos uma exposição inicial do isomorfismo psicofísico em Pessoa Jr. (2023). Em suma, o “isomorfismo psicofísico” é definido por ele como exigindo que “em um dado caso a organização da experiência e os fatos fisiológicos subjacentes têm a mesma estrutura” (KÖHLER, 1943, p. 301). Assim, no caso espacial [e temporal], “a ordem experienciada no espaço [tempo] é sempre idêntica estruturalmente à ordem funcional na distribuição [sequência] dos processos encefálicos subjacentes” (KÖHLER, 1943, p. 60-61). Vemos assim que o isomorfismo psicofísico é uma defesa da tese da identidade mente-encéfalo para propriedades estruturais do campo fenomênico, mas não para os *qualia*. Nas palavras de Boring (1933, p. 237-8):

Há uma tendência crescente de aplicar o termo de Köhler, *isomorfismo*, para a correspondência geral entre as ordens da mente fenomênica e as ordens do corpo neural, como se o espaço pensado fosse espaço neural, tempo pensado tempo neural, intensidade pensada intensidade neural, qualidade pensada ‘como-queira-chamá-la’ neural [*thought quality neural ‘what-you-may-call-it’*].

Boring (1933) colocou-se de maneira cética em relação a esta teoria, mas foi a partir de sua análise, como já vimos na seção 4, que Place formulou sua versão da identidade mente-encéfalo. Boring (1933, p. 16) chegou a sugerir a identidade mente-encéfalo mesmo para qualidades, concepção própria da tese do encéfalo colorido, ao se referir ao vermelho:

Assim o vermelho, sendo sempre vermelho, é identificado consigo mesmo, e todos os seus sintomas são correlatos perfeitos. Se alguém desejar insistir que o vermelho é um par de covariantes perfeitos, nós não podemos refutá-lo. Ele tem direito de manter sua visão. Mas por que deveríamos assumir dois, quando um é suficiente? Se o princípio de parcimônia de Guilherme de Ockham estiver alguma vez justificado, deve ser no presente contexto (BORING, 1933, p. 16).

Köhler ([1929] 1943) salientou que o correlato encefálico imediato de um *quale* de azulidão (o “como-queira-chamá-la neural” mencionado por Boring mais acima) não faz parte do isomorfismo, porque este isomorfismo refere-se a uma propriedade estrutural (*Systemeigenschaften*) e não a uma propriedade material (*Materialeigenschaften*), como os *qualia*. Scheerer (1994, p. 189) menciona que Köhler concebia que tal propriedade material tinha natureza química, seguindo assim a Hering (1913, p. 60-1), que a chamava de “energias específicas”, seguindo Johannes Müller. Em 1938, em obra cujo título em português seria *O lugar do valor em um mundo de fatos*, Köhler descartou explicitamente a tese do encéfalo colorido ao escrever que “O correlato cortical do azul não é azul” (*apud* SCHEERER, 1994, p. 189).

9. Conclusão

Visitamos várias questões associadas à tese da identidade mente-encéfalo. Uma primeira constatação materialista é a de que os processos mentais ocorrem, de maneira imediata, no encéfalo, entendido como a matéria contida dentro do crânio humano. Esta aceção será denotada por “encéfalo₁”. Essa constatação pode ser expressa por meio de uma tese de identidade *extensional* (de instância): processos e qualidades mentais estão localizadas no interior do crânio. Esta versão da tese da identidade é mais fácil de aceitar.

A dificuldade surge para a tese de identidade *intensional*. Vimos que Feigl argumentou que se mente e encéfalo forem intensionalmente idênticos, eles deveriam ser indistinguíveis, o que para ele não seria o caso. Defendemos aqui que de fato há uma identidade intensional envolvendo os processos e qualidades mentais, restando especificar em que sentido o “encéfalo” está envolvido.

Para isso, uma distinção fundamental é entre *estrutura* e *qualidade*, distinção esta feita por Kant, Mach, Schlick, Russell, Köhler, Feigl e Pepper, entre outros. As propriedades estruturais são aquelas mensuráveis por aparelhos físicos, e descritíveis em termos relacionais, de maneira linguística e matemática. Às vezes, o termo “encéfalo” designa a informação advinda das medições físico-químicas e da teorização científica (ou seja, o que Mary conhece sobre as cores dentro de seu quarto preto e branco; cf. Jackson, 1982), o que será denotado por “encéfalo₂”. O encéfalo₂ captura as propriedades estruturais do encéfalo₁, seu referente. Dentro do crânio não há apenas relações, mas também “relata”, ou seja, entes que estão em relação, constituindo a substância ou matéria do encéfalo. São estas que incluem as chamadas “propriedades intrínsecas”. Defendemos que os *qualia*

são idênticos a parte dessas propriedades intrínsecas. Em outras palavras, os *qualia* são encefálicos₁, mas não encefálicos₂.

A identidade da estrutura do campo fenomênico e da estrutura de processos encefálicos é descrita pelo isomorfismo psicofísico de Köhler ([1929] 1943), levando adiante as análises anteriores de Mach, Hering e G. E. Müller. Já a identidade intensional relativa aos *qualia* é satisfeita identificando os *qualia* não com os aspectos estruturais e mensuráveis do encéfalo, mas com propriedades intrínsecas da matéria encefálica.

Defendemos a tese do encéfalo colorido (uma versão do fisicismo qualitativo), segundo a qual a verdidão-jade vivenciada subjetivamente no fenômeno meteorológico do raio verde está em uma porção do encéfalo₁, literalmente. Tal cor subjetiva não tem nada a ver com a luz (ela é apenas causada indiretamente pela luz que cai na retina), e é ou uma propriedade intrínseca de macromoléculas químicas ou uma propriedade intrínseca de campos elétricos. O termo “intrínseco” indica que tais propriedades não são mensuráveis por procedimentos físicos convencionais, que capturam apenas relações, apesar de serem vivenciados subjetivamente (podendo ser considerados resultado de um processo especial de observação).

A tese da identidade intensional aqui defendida não é uma tese de identidade “mente-encéfalo₂”, onde por “encéfalo₂” entendemos o conjunto de propriedades mensuráveis por aparelhos físicos e exprimíveis teoricamente de maneira linguística ou quantitativa. Defendemos uma tese de identidade intensional “mente-encéfalo₁”, ou melhor, a tese de que a estrutura e a qualidade presentes na consciência são idênticos a parte do encéfalo₁.

A tese da superveniência do mental sobre o encéfalo₁ é satisfeita de maneira trivial por uma tese da identidade. O problema empírico interessante é se há um princípio de superveniência do mental sobre o encéfalo₂, ou seja, se *qualia* supervêm sobre a estrutura do encéfalo, mesmo não sendo idênticos a esta.

Agradecimentos

Este artigo é baseado em uma apresentação feita no 2º Encontro de Cognição e Linguagem (2eC&L), realizado da Universidade Federal de Uberlândia, em 06/11/2019, e disponível em inglês em <https://www.youtube.com/watch?v=QHeeSC4hPSE&list=PL-TpRvDsWfOhC5AbZo8ZUPL0OGMsJkSmAT&index=7&t=0s>. Apresentado também na XXII Semana Acadêmica do PPGFIL-PUCRS, Porto Alegre, em 27/09/2022. Agradeço a

discussão realizada nesses eventos, e em especial as críticas de Beatriz Sorrentino, Sofia Stein e Eros Moreira de Carvalho.

Referências

ARISTÓTELES. **De anima**. Trad. Maria Cecília G. dos Reis. São Paulo: Editora 34, 2006.

BORING, E. G. **The physical dimensions of consciousness**. New York: Century, 1933.

BORING, E. G. **Sensation and perception in the history of experimental psychology**. New York: Appleton-Century, 1942.

BROAD, C.D. **The mind and its place in nature**. London: Kegan Paul, Trench, Trubner & Co., 1925.

CASE, T. **Physical realism**. London: Longmans, Green & Co, London, 1888.

CHALMERS, D. J. **The conscious mind**. New York: Oxford University Press, 1996.

CHALMERS, D. J. Panpsychism and panprotopsychism. In: ALTER, T. & NAGASAWA, Y. (Orgs.). **Consciousness in the physical world: perspectives on Russellian monism**. Oxford: Oxford University Press, 2015, p. 246-76.

CLIFFORD, W. K. On the nature of things-in-themselves. **Mind**, v. 3, n. 9, p. 57-67, 1878. Republicado in STEPHEN, L. & POLLOCK, F. (Orgs.). **Lectures and essays by the late William Kingdon Clifford**, vol. 2. London: Macmillan, 1879, p. 71-88.

DENNETT, D. **Consciousness explained**. New York: Little, Brown & Co., 1991.

DESCARTES, R. **Meditações**. Trad. J. Guinsberg & B. Prado Jr., v. 15. São Paulo: Abril Cultural, 1973, p. 81-150. Orig. em latim: 1641. (Os Pensadores)

DRAKE, D. **Invitation to philosophy**. Boston: Houghton Mifflin, 1933.

FEIGL, H. **The “mental” and the “physical”**: the essay and a postscript. Minneapolis: University of Minnesota Press, 1967. (Versão original sem o pós-escrito: 1958).

FEIGL, H. Physicalism, the unity of science, and the foundations of psychology. In: SCHILLP. P.A (Org.). **The philosophy of Carnap**. La Salle (IL): Open Court, pp. 227-68, 1963. Republicado em: FEIGL, H. **Inquiries and provocations: selected writings 1929-1974**. Ed. R.S. Cohen, Vienna Circle Collection v. 14. Dordrecht: Reidel, p. 302-41, 1981.

FEIGL, H. Some crucial issues of mind-body monism. **Synthese**, v. 22. p. 295-312, 1971.

FEYERABEND, P.K. Comment: mental events and the brain. **Journal of Philosophy**, v. 60, p. 295-6, 1963.

FREGE, G. Sobre o sentido e a referência. Trad. S.R.N. Miranda. **Fundamento** (Ouro Preto), v. 1, n. 3, p. 21-44, 2011. (Orig. em alemão: 1892)

HEIDELBERGER, M. **Nature from within: Gustav Theodor Fechner and his psychophysical worldview**. Trad. C. Klohr. Pittsburgh: University of Pittsburgh Press, 2004. Orig. em alemão: 1993.

HERING, E. On the theory of nerve-activity. In: HERING, E. **Memory: lectures on the specific energies of the nervous system**. 4^a ed. Open Court, Chicago, p. 43-70, 1913. Tradução da palestra dada em 1898 em Leipzig.

HUXLEY, T. H. On the hypothesis that animals are automata. **Fortnightly Review**, v. 16, p. 555-80, 1874.

JACKSON, F. Epiphenomenal qualia. **Philosophical Quarterly**, v. 32, p. 127-36, 1982.

KIM, J. Multiple realization and the metaphysics of reduction. **Philosophy and Phenomenological Research**, v. 52, p. 1-26, 1992.

KÖHLER, W. **Gestalt psychology**. 2. ed. New York: Liveright, 1943. 1^a ed.: 1929. Em português: **Psicologia da Gestalt**. Belo Horizonte: Itatiaia, 1980.

LEWES, G. H. **Problems of life and mind**. First series, vol. 2. Boston: J. Osgood, 1875.

LOUX. M. J. **Metaphysics: a contemporary introduction**. 3. ed. New York: Routledge, 2006.

O'REGAN, J. K. & NOË, A. A sensorimotor account of vision and visual consciousness. **Behavioral and Brain Sciences**, v. 24, p. 939-1031, 2001.

PEPPER, S. C. **Concept and quality**: a world hypothesis. La Salle (IL): Open Court, 1967.

PESSOA JR., O. A consciência enquanto observação do cérebro. In: FERREIRA, F.R.M.; HADDAD JR., H.; NOGUEIRA, M. I. & PESSOA JR., O. (Orgs.). **História e filosofia da neurociência**. São Paulo: Liber Ars, 2015. p. 201-19.

PESSOA JR., O. How to measure a quale. **Sofia**, v. 8, p. 187-98, 2019.

PESSOA JR., O. The colored-brain thesis. **Filosofia Unisinos - Unisinos Journal of Philosophy**, v. 22, p. 84-93, 2021.

PESSOA JR., O. Três abordagens ao problema mente-corpo entre neurofisiologistas de meados do séc. XX. In: MARTINS, L.A.-C.P.; MARTINEZ, L. M. D.; FEDERICO, L.; PINO, G. G. & O'LEARY, M. M. (Eds.). **Reflexiones filosóficas e históricas**: ciencia, enseñanza de la ciencia y política científica. AFHIC (Asociación de Filosofía e Historia de la Ciencia del Cono Sur/Associação de Filosofia e História da Ciência do Cone Sul). Universidad del Valle, Departamento de Filosofía Cali, Colombia; Ciudad de Editorial: São Carlos-SP, 2023, p. 305-19. Disponível em: <http://www.afhic.com/es/libro/reflexiones-filosoficas-e-historicas-ciencia-ensenanza-de-la-ciencia-y-politica-cientifica/>

PLACE, U. T. Is consciousness a brain process? **British Journal of Psychology**, v. 47, p. 44-50, 1956.

PLACE, U. T. Identity theories. Manuscrito não terminado de Place e editado por S. Schneider, 2000. Disponível no site **A field guide to the philosophy of mind**: <http://host.uniroma3.it/progetti/kant/field/mbit.htm>

PRICE, H. H. **Perception**. London: Methuen, London, 1932.

PUTNAM, H. W. Psychological predicates. In: CAPITAN, W.H. & MERRILL, D. D. (Orgs.) (1967). **Art, mind and religion**. Pittsburgh: University of Pittsburgh Press, 1967, p. 37-48. Reimpresso em HEIL, J. (Org.). **Philosophy of mind**: a guide and anthology. Oxford: Oxford University Press, p. 158-67, 2005. Reimpresso também com o título “The nature of mental states”.

SCHEERER, E. Psychoneural isomorphism: historical background and current evidence. **Philosophical Psychology**, v. 7, p. 183-210.

SCHLICK, M. **General theory of knowledge**. Trad. A.E. Blumberg. New York: Springer, 1974. Trad. da 2ª edição alemã de 1925; 1ª ed.: 1918.

SKRBINA, D. **Panpsychism in the West**. Cambridge (MA): MIT Press, 2005.

SPERRY, R. W. A interação mente-cérebro: mentalismo, sim; dualismo, não. In: SPERRY, R. W. **Ciência e prioridade moral: uma fusão da mente, do cérebro e dos valores sociais**. Trad. L.S. Csekö & C.A. Medeiros. Rio de Janeiro: Zahar, 1986, p. 109-139. Orig. em inglês: **Neuroscience**, v. 5, p. 195-206, 1980.

STUBENBERG, L. **Consciousness and qualia**. Amsterdam: John Benjamins, 1998.

STUBENBERG, L. Neutral monism. **Stanford Encyclopedia of Philosophy**, online, 2016.

VERNE, J. **O raio verde**. Trad. M.A. Sampaio Dória. São Paulo: Melhoramentos, 2011. Orig. em francês: 1882.





A EMERGÊNCIA DA PERSPECTIVA DA PRIMEIRA PESSOA¹



Cláudia Passos-Ferreira²

Resumo:

Uma definição científica da autoconsciência e da perspectiva da primeira pessoa é um desafio para o projeto naturalista da mente. Explicar como temos consciência de nós mesmos a partir do nosso próprio ponto de vista parece resistir a explicações reducionistas da consciência. Filósofos naturalistas (como Dennett e Dretske) têm buscado explicar a autoconsciência como uma capacidade que emerge na ontogênese. Nesse artigo, analiso como a perspectiva da primeira pessoa emerge durante o desenvolvimento ontogenético a partir da aquisição da capacidade de meta-representar e conceitualizar. Ao longo do desenvolvimento, emergem processos cognitivos que culminam na aquisição da perspectiva da primeira pessoa, ou seja, da capacidade de referir-se a si próprio em primeira pessoa. Os bebês humanos nascem conscientes e se tornam autoconscientes.

Palavras-chave:

Consciência; autoconsciência; perspectiva de primeira pessoa.

1. O que é se sentir como nós

No episódio da série *Star Trek: The Next Generation* “The Measure of the Man”, o Tenente Comandante Data, um androide, é julgado por seu estatuto de ser ou não uma

¹ Este artigo é retirado da minha tese de doutorado *Origens da Interioridade: Autoconhecimento e Externalismo* (Passos-Ferreira, 2006) e reflete minhas ideias sobre esse tema naquele período. A tese analisa os obstáculos que uma abordagem externalista do mental pode trazer ao problema do autoconhecimento, ou seja, como compatibilizar a ideia de que os conteúdos mentais são parcialmente determinados pelo ambiente externo com a noção de que temos conhecimento privilegiado de nossos estados mentais e não nos equivocamos na atribuição de estados ao eu. Agradeço a Roberto Sá Pereira, Jurandir Freire Costa, Octavio Serpa e David Chalmers pelos comentários em versões prévias do artigo.

² Claudia Passos-Ferreira. New York, USA. e-mail: claudiapassos@nyu.edu

pessoa consciente. Data é uma forma de vida artificial, projetada para assemelhar-se a um humano. Ele é possuidor de um cérebro composto de positrônios que funciona como um computador central e exibe uma forma de consciência reconhecida por humanos. O cientista Capitão Maddox quer desmontar Data para aprender como o androide funciona. Mas o Capitão Jean-Luc Picard defende Data contra a ameaça de ser desmontado, argumentando que Data é um ser senciente com direitos.

A questão central desse episódio é o direito que teriam os humanos que projetaram Data, seus criadores, de intervir sobre sua “vida artificial” e, conseqüentemente, de destruí-lo. A defesa em favor de desmontar Data se baseia no argumento de que ele não é uma pessoa com direitos, mas sim uma máquina de propriedade da União da Federação dos Planetas. Em defesa de Data, Capitão Picard alega que ele preenche todos os critérios dos seres conscientes, como a capacidade de sentir, perceber e agir intencionalmente, e que, como tal, teria o direito à liberdade e à autodeterminação concedidos às pessoas. Picard alega que todos os seres são “criados”, mas que isso não os torna possessão de seus criadores. Data, portanto, teria o direito de tomar suas próprias decisões sobre sua vida.

Esse exercício ficcional traz à baila a questão sobre os critérios para a definição do que é a personalidade. Em *Conditions of Personhood* (1981a), Daniel Dennett interroga nosso senso comum sobre o que é ser uma pessoa. Ele afirma que eu e você, leitor, certamente somos pessoas, mas não há garantias de que somos feitos da mesma composição biológica. Portanto, se o fato de pertencermos à espécie humana, se nossa condição biológica não é uma condição necessária nem suficiente para caracterizar a pessoa, quais seriam, então, os critérios necessários? Dennett apresenta seis condições necessárias para a personalidade: ser racional; ser um ser a quem se atribui estados psicológicos e predicados intencionais; ser objeto das atitudes de pessoa (*personal stance*), ou seja, ser tratado como pessoa; ser capaz de tratar reciprocamente os outros como pessoas; ser capaz de comunicação; exibir um tipo especial de consciência, a autoconsciência.

As três primeiras condições – racionalidade, intencionalidade e adoção de atitudes de pessoa – são condições necessárias, mas não suficientes para ser uma pessoa. Os seres que exibem essas três primeiras condições são *sistemas intencionais*³ que se caracterizam por apresentar comportamentos que podem ser explicados e previstos por meio de atribuições

3 O sistema intencional é toda entidade cujo comportamento é previsível e explicável pela *atitude intencional*, ou seja, pela imputação ao sistema da posse de certas informações e pela suposição de que é um agente cujas ações são guiadas por certos objetivos. Como afirma Dennett, “A atitude intencional é a estratégia de interpretar o comportamento de uma entidade (pessoa, animal, artefato, seja lá o que for) tratando-a como se fosse um agente racional que governa sua ‘escolha’ da ‘ação’ levando em ‘consideração’ suas ‘crenças’ e ‘desejos’” (DENNETT, 1996, 27).

de crenças e desejos. O quarto critério, *reciprocidade*, caracteriza os sistemas intencionais capazes de adotar uma atitude intencional frente ao mundo e às outras pessoas. São *sistemas intencionais de segunda-ordem* capazes de ter crenças e desejos acerca de suas próprias crenças e desejos, ou seja, são capazes de representar crenças, desejos e intenções de segunda ordem. Essas características são exibidas por alguns animais⁴. A quinta característica, ser capaz de comunicação verbal, ou seja, ser dotado de linguagem, exclui os animais não humanos da noção de pessoalidade. A última condição, a autoconsciência, é a ideia de que as pessoas são distinguíveis de outras entidades, pois exibem um tipo especial de consciência. A autoconsciência é o que nos permite ter consciência das nossas ações, ou melhor, ter consciência de que somos nós os responsáveis por nossas ações e que podemos ser indagados acerca das razões pelas quais as executamos e ser responsabilizados moralmente por nossos atos.

Como Dennett (1981a) ressalta, a normatividade é inerente à noção de pessoa. Mesmo na presença das seis condições, persistirá sempre o caráter ideal e arbitrário do conceito de pessoa que depende da interação entre pessoas. Qualquer definição metafísica que estabeleça condições ontológicas acerca da pessoalidade (a noção metafísica da pessoa) não escapará dos aspectos normativos que compõem a noção moral de pessoa. Definimos o sentido de nossos critérios metafísicos a partir de nossas injunções éticas, como propõe Stephen White citado por Dennett (2004).

Contudo, é inegável que a presença dessas condições oferece critérios, mesmo que provisoriamente suficientes, para a compreensão da *constituição* do agente moral. A capacidade de **autoconsciência** – ou, ao menos, a possibilidade de adquiri-la, no caso dos bebês, ou o fato de já tê-la tido em algum momento, no caso dos que a perderam – tem sido um dos critérios definidores do que é uma pessoa.

Data é um ser que satisfaz essas condições. É por isso que somos inclinados a conceder-lhe o estatuto de pessoa. Embora possamos restringir seus direitos, alegando que ele pertence a uma classe diferenciada de pessoa – como nas práticas de exclusão social em que limitamos os direitos àqueles que consideramos como não pertencendo ao nosso grupo ou comunidade moral, como não sendo *um de nós*⁵ – Data tem características que têm

4 Os estudos em psicologia evolucionária têm mostrado que certos animais (em particular os primatas) possuem representações mentais de seu ambiente e dos comportamentos dos outros que os permitem resolver problemas e simular novas situações. Alguns chimpanzés chegam mesmo a exibir comportamentos que se assemelham à capacidade de ler o comportamento dos outros, um esboço do que chamamos ‘teoria da mente’. Sobre a evolução da cognição em primatas e sua relação com o desenvolvimento ontogenético humano, ver Doug Jones (1999), Tomasello e Call (1997).

5 Sobre a noção de ‘um de nós’ como pertencendo a uma dada comunidade moral, ver Richard Rorty (1994).

sido reconhecidas como definindo a pessoa. Ainda que esses critérios ontológicos estejam, em última instância, subordinados aos aspectos morais definidos pela comunidade da qual pertencemos, isso não torna obsoleto o seu uso como balizadores de nossos discursos e atitudes frente ao que consideramos uma pessoa. Dentre todos os critérios, **a capacidade de autoconsciência** tem sido tratada como um critério decisivo.

A autoconsciência é, geralmente, considerada um traço distintivo da espécie humana. Mesmo que chimpanzés exibam uma capacidade cognitiva rudimentar de auto-reconhecimento (como aprender a se reconhecer no espelho após um tempo de exposição ao objeto) (Gallup, 1997), a capacidade de autoconsciência, de fazer referência a si mesmo e identificar a si mesmo, parece específica da nossa espécie. Somos as únicas criaturas capazes de pensar sobre nós mesmos como uma existência permanente no tempo (*persisting self*), isto é, a ter uma perspectiva da continuidade do eu. A autoconsciência é um aspecto essencial na construção de nossa identidade pessoal ao longo de nossa história de vida, pois nos permite escapar do aqui e agora de nossa 'presença ao mundo' e nos projetarmos numa dimensão futura como agentes morais responsáveis por nossas ações e decisões. Portanto, a autoconsciência é essencial para ser uma pessoa (Passos-Ferreira, no prelo).

Mas, como definimos um ser autoconsciente?

Todas as formas de autoconsciência pressupõem um ponto de vista subjetivo. Em “*What is it like to be a bat*” (1992), Thomas Nagel mostra como um ser consciente se torna autoconsciente ao adquirir uma perspectiva de primeira pessoa. A perspectiva da primeira pessoa é a experiência subjetiva por meio da qual podemos nos conceber como um ser individual, distinto de todo o resto do mundo com o qual estamos lidando. Como mostra Nagel, todo ser consciente, mesmo seres sencientes não humanos, é sujeito de suas experiências no sentido de que as experimenta. Todo fenômeno subjetivo está conectado, portanto, como um ponto de vista singular.

No entanto, nem todo ser consciente é autoconsciente. A autoconsciência é um estado que nos torna consciente *de* nós mesmos. É a consciência de ter uma experiência subjetiva e possuir conceitos de si mesmos na perspectiva da primeira pessoa. Somente os animais humanos são capazes de exibir essa forma conceptual de perspectiva da primeira pessoa e podem ser ditos completamente autoconscientes, ou, como denomina Lynne Baker (1998), capazes de exibir uma versão *forte* da perspectiva da primeira pessoa.

Baker propõe uma distinção entre as perspectivas da primeira pessoa forte e fraca, que parece útil para esclarecer a relação entre a autoconsciência e a identidade da pessoa. A perspectiva *fraca* é aquela exibida por criaturas capazes de resolver problemas e cujos comportamentos podem ser explicados em termos intencionais. São seres como os animais não humanos e os recém-nascidos, aos quais podemos atribuir crenças e desejos,

pois se comportam de uma certa perspectiva. Não é necessário que um organismo possua o conceito de primeira pessoa para que possam lhe atribuir um comportamento motivado por crenças e desejos. Segundo Baker, os fenômenos *fracos* da perspectiva da primeira pessoa são exibidos por seres sencientes, capazes de solucionar problemas por meio de atitudes perspectivais que explicam seu comportamento intencional. Como todos os estados psicológicos são perspectivais na origem, assim também a crença, o desejo e o comportamento pertencem a um único organismo.

O que é característico da perspectiva *forte* da primeira pessoa é a capacidade de **reconhecer a si mesmo como um ponto de vista subjetivo** e de pensar sobre si mesmo como uma pessoa. Essa capacidade única dos humanos – de se conceber como ‘eu’ autoconsciente e capaz de introspecção – é fundamental para o desenvolvimento do que consideramos uma pessoa, *what-is-it-like-to-be-us*.

Analisarei aqui a emergência da perspectiva forte da primeira pessoa ao longo do desenvolvimento ontogenético. A hipótese central é que, no início, somos seres conscientes e, ao longo do desenvolvimento, nos tornamos autoconscientes. A capacidade de autoconsciência e o autoconhecimento emergem a partir da aquisição da capacidade de meta-representar e de ter conceitos, em particular, os conceitos que permitem fazer referência ao eu e a seus estados mentais e corporais.

2. Da consciência fenomenal à autoconsciência

O debate acerca da experiência subjetiva da consciência tem sido dividido, grosso modo, entre eliminativistas, dualistas e fisicalistas. Os eliminativistas são os que rejeitam o caráter qualitativo da experiência. Os dualistas tentam preservar o caráter de singularidade da experiência subjetiva e negam qualquer redução da consciência a processos neurofisiológicos. Os fisicalistas tentam oferecer um relato naturalista da consciência.

Existem diversas tentativas teóricas de naturalizar a consciência e desmistificar seu caráter ‘misterioso’ que mostram como os aspectos relevantes da consciência podem existir numa concepção puramente fisicalista do mundo. O projeto de naturalização da mente visa explicar os fatos mentais a partir das ciências naturais, descrevendo estados e processos mentais como aspectos do cérebro e do sistema nervoso central. Dentre os defensores da naturalização da consciência, o debate se divide entre representacionistas e funcionalistas.

Examinarei a abordagem representacionista de Fred Dretske (1997), que tenta preservar o caráter qualitativo da experiência subjetiva. Dretske propõe compatibilizar a ideia de que os aspectos qualitativos da experiência são subjetivos e privados com a ideia

de que são passíveis de serem determinados objetivamente, ou seja, podem ser capturados por um relato representacional.

Em sua Teoria Representacional da Consciência⁶, o estado de estar consciente é um modo de representar alguma coisa. A consciência fenomenal nos torna consciente das propriedades fenomenais dos objetos e é dada pelos sentidos perceptivos: ver, ouvir, cheirar, provar, sentir. Os estados mentais conscientes são estados que nos tornam conscientes de objetos, propriedades e fatos. São estados que nos tornam conscientes *com* eles e não estados *dos* quais estamos conscientes. O que torna a pessoa consciente de estar tendo a experiência é um estado interno que representa as propriedades da experiência. Sem o estado representacional não é possível ter consciência da propriedade que experimentamos.

A experiência consciente, em termos do *what it is like for the subject*, pode ser distinguida em dois aspectos: (1) a consciência das propriedades fenomenais dos objetos, que é o caráter qualitativo da experiência que os fenomenólogos chamam de ‘presença ao mundo’; (2) o aspecto ‘*para-mim*’ (*for-me-ness*, como chama Kriegel (2005)), que é o caráter subjetivo da experiência, ou seja, o que é ter a experiência *para o sujeito*. O caráter subjetivo da experiência consiste num conteúdo representacional particular que se desdobra em uma representação de algum aspecto do ambiente e numa representação interna da representação.

Para um materialista como Dretske, não há fatos mentais que só sejam acessíveis a uma única pessoa. Toda subjetividade faz parte do mundo objetivo. Segundo Dretske (1997, 65),

pode haver circunstâncias de tempo e espaço que capacitem uma pessoa a conhecer alguma coisa que os outros não conhecem, circunstâncias de tempo e espaço que tornem uma pessoa autoritativa, mas não há fatos privilegiados a uma pessoa, fatos que só uma pessoa *pode* conhecer.

⁶ Há três usos da noção de representação: a ‘atividade da representação’, a ‘ideia de representação’, e a ‘representação mental’. A ‘atividade da representação’ está associada a alguma atividade mental na qual nos engajamos. A ‘ideia de representação’ geralmente é usada na Teoria Representacional da Percepção como uma interface semântica interna entre nós mesmos e o que estamos pensando ou percebendo, e está associada à ideia de colocar ‘um objeto diante da mente’. O terceiro uso encontramos mais recentemente na literatura após a virada informacional (*the information turn*) e, de certa forma, deriva da ideia precedente da representação como um objeto mental com propriedades semânticas. A noção de ‘representação mental’ é um construto teórico das ciências cognitivas segundo o qual os estados e processos cognitivos são constituídos pela ocorrência, transformação e estocagem (no cérebro ou na mente) de estruturas representacionais que captam informações. Essa noção é o ponto de partida da Teoria Representacional ou Computacional da Mente, que encontra em Dretske um dos seus mais promissores defensores. O ponto de partida dessa abordagem é que os estados mentais (pensamentos, crenças, desejos e percepções) estão relacionados a representações mentais e a intencionalidade desses estados deve ser explicada em termos das propriedades semânticas da representação mental (Para uma melhor discussão, ver Pitt (2000)).

Temos informação direta e imediata sobre nossa vida mental, mas não temos acesso privilegiado a informações que os outros não possam ter. Todos podem ter acesso a mesma informação. Isso quer dizer que a vida subjetiva de uma criatura, o que é se sentir como essa criatura (*what it is like to be*), não é inacessível. Contudo, há uma particularidade na experiência de ser um ponto de vista.

Ao contrário de certos relatos naturalistas que explicam o mental a partir do funcionamento da maquinaria cerebral biológica, Dretske concebe a mente como a função de manipulação de representações, a partir da ideia de ‘informação’⁷. Ele defende duas teses: (1) todos os fatos mentais são fatos representacionais; (2) todos os fatos representacionais são fatos sobre funções informacionais. A ideia central é que a mente é um *design* projetado pela história evolutiva com a função de trazer a informação sobre o objeto. A mente é um sistema que representa uma dada propriedade e tem a função de indicar e representar a propriedade. A função pode ser sistêmica ou adquirida. As funções sistêmicas são funções filogenéticas dos sistemas sensoriais produzidas pelo processo histórico de seleção natural. As funções adquiridas são funções aprendidas no processo de maturação do indivíduo. Os sistemas perceptivos (incluindo a propriocepção) produzem representações e têm a função biológica de fornecer informação. A tarefa do aparelho sensorial é construir representações mentais sobre o mundo. Todos os estados mentais (crenças, pensamentos e experiências) são representações naturais. Toda representação é representação de um fato, mas nem toda representação é conceitual. A experiência é uma forma de representação não-conceitual e é inerente ao sistema. As crenças e pensamentos são representações conceituais adquiridas ao longo do desenvolvimento ontogenético.

Tanto as atitudes proposicionais (crenças e desejos) quanto a experiência sensorial (os *qualia*) são representações internas. O caráter qualitativo da experiência é identificado às propriedades que os objetos representados possuem, e essas propriedades são, em princípio, passíveis de serem conhecidas por qualquer pessoa. Para Dretske, o caráter qualitativo da experiência perceptiva não deve ser definido funcionalmente, por meio de nossas disposições comportamentais, mas deve ser definido em termos físicos.

7 Em seus trabalhos mais recentes, Dretske reformula sua Teoria Semântica Informacional e propõe uma abordagem representacionista, que combina a noção de informação com o elemento teleológico (Teleosseântica). O importante para o argumento aqui desenvolvido é que essa reformulação resulta na ideia de que a principal distinção entre as representações sensoriais e conceituais é feita em termos da distinção entre dois tipos básicos de indicadores de função, sistêmica e adquirida. A representação de que *O* é *P* é sensorial e não-conceitual, quando seu indicador de função é derivado do sistema do qual ele é um estado, e conceitual quando seu indicador de função é adquirido do tipo do qual ele é um estado particular. No primeiro caso, o indicador de função é filogeneticamente determinado pela seleção natural, enquanto no segundo caso ele é adquirido ontogeneticamente pelo aprendizado.

Chamamos de *qualia* as propriedades fenomenais dos objetos que são experimentadas subjetivamente. Essas propriedades são aquelas que estão representadas nas informações que nossos sentidos têm a função natural de fornecer; e podem ser determinadas assim como determinamos outras funções biológicas dos órgãos corporais.

Os organismos têm funções biológicas específicas com diferentes performances discriminativas e capacidades diferenciadas de representar propriedades. As experiências qualitativas podem ser distintas mesmo que os organismos tenham a mesma capacidade biológica. Portanto, a experiência de ser um ponto de vista subjetivo não pode ser reduzida à função biológica. O ponto de vista subjetivo é determinado também pelas partes e aspectos da experiência do objeto que representamos. Todo e qualquer objeto ocupa diferentes lugares no mundo, portanto, captamos informações diferentes do mundo de acordo com as posições que ocupamos. Contudo, as experiências não diferem subjetivamente apenas porque são experiências de diferentes objetos em diferentes perspectivas. Essa diferença é uma experiência que pode ser partilhada ao trocarmos de lugar com o outro e ocuparmos o mesmo ponto de vista do outro. As diferenças nos estados subjetivos que interessam e caracterizam a singularidade da experiência são as diferenças que resultam não de pontos de vistas diferentes, mas no *modo* como os pontos de vistas são vistos. O que faz diferença no *what it is like for the subject* não são os objetos acerca dos quais a experiência é, mas o *modo* como esses objetos são representados de um dado ponto de vista.

Portanto, como afirma Dretske, o que muda é o *modo* como a experiência é codificada. Ao nível da consciência fenomenal, temos a experiência de uma variedade de propriedades. Ao nível da consciência conceitual, temos experiência de um *tipo* de propriedade, pois há uma categorização da experiência sensorial que é dada pelo aprendizado dos conceitos. Os sistemas sensoriais inatos são funções herdadas filogeneticamente, e não podem ter o caráter representacional da experiência alterado. No entanto, são adaptáveis e recalibráveis. O modo como a *experiência* representa o mundo é fixado pelas funções biológicas do sistema. Porém, a *cognição* é uma função que se desenvolve ontogeneticamente. E o modo como uma *crença* representa o mundo é determinado pelo aprendizado individual da capacidade de conceituar. Podemos ter a mesma experiência e conceituá-la de forma diferenciada. Pelo aprendizado, eu posso alterar o que eu acredito e penso quando experiencio o mundo, mas não posso alterar a qualidade do que experiencio. A consciência da experiência altera aquele que a experimenta, mas não modifica a experiência. A diferença entre as experiências que temos consciência e as que não temos não está na própria experiência, mas naquele que tem a experiência, ou seja, no que a pessoa sabe a respeito da experiência que está tendo.

A aquisição da capacidade de conceituar permite a aquisição de conceitos sobre o eu, o que permitirá fazer referência ao eu e representar a experiência subjetiva. A capacidade de representar a experiência subjetiva altera significativamente o *modo* como a pessoa experimenta seu ponto de vista e permite a emergência da consciência de si como uma criatura com uma perspectiva dentre outras.

3. Tornar-se uma perspectiva

A observação do comportamento dos recém-nascidos evidencia aspectos importantes da experiência de ser um ponto de vista distinto dos outros e de se comportar no mundo a partir de uma perspectiva. Até sermos capazes de exibir fenômenos da perspectiva forte da primeira pessoa, nós humanos passamos por cerca de 4 anos de desenvolvimento com mudanças qualitativas consideráveis que culminam com a emergência da capacidade completa de autoconsciência. Do nascimento até 4-5 anos de idade, são observados diferentes níveis de desenvolvimento da representação do eu que indicam alterações importantes no modo como as informações sobre o eu são representadas. Existem cinco etapas que correspondem a diferentes níveis de auto-representação: agente ecológico (0-2 meses); agente situado (2-9 meses); agente perspectivo (9-18 meses); agente representacional (18-36 meses); agente meta-representacional (3-4 anos)⁸.

Desde o nascimento, o bebê humano é capaz de perceber seu próprio corpo como uma entidade diferenciada de outras no ambiente, ou seja, capaz de fazer a **diferenciação eu-mundo** (ver GIBSON, 1979; NEISSER, 1988; ROCHAT, 2003; ROCHAT & STRIANO, 2000; BUTTERWORTH, 1995). A abordagem ecológica gibsoniana enfatiza a experiência precoce de diferenciação, opondo-se à ortodoxia em psicologia de que o bebê nasceria num estado de confusão com seu entorno, e que suas ações seriam respostas automáticas a estímulos do ambiente – tese da indiferenciação eu-outro defendida por, entre outros, William James e Jean Piaget (BUTTERWORTH, 1995; PASSOS-FERREIRA, 2010; 2011). A tese proposta por James Gibson (1979) é que, desde o nascimento, o bebê se comporta como um centro de ação e percepção cuja experiência perceptiva acarreta autopercepção e um sentido implícito do próprio corpo⁹.

8 Essas cinco etapas de desenvolvimento foram adaptadas das classificações apresentadas sobre o desenvolvimento da autoconsciência por Rochat (2003) e do sentido de agentividade por Fonagy (2004).

9 Como afirma Gibson: “A informação sobre o self acompanha a informação sobre o ambiente, e as duas são inseparáveis. A ego-recepção acompanha a exterocepção, como se fosse o outro lado da moeda. A percepção tem dois pólos, o subjetivo e o objetivo, e a informação está disponível para especificar ambos. Percebemos o ambiente e co-percebemos a nós mesmos.” (Gibson 1979, 126)

Para Gibson, nascemos biologicamente equipados com as funções proprioceptivas e exteroceptivas, que são funções do sistema perceptivo que se dirigem ao corpo e ao mundo. A propriocepção traz informação¹⁰ sobre “os aspectos mutáveis e persistentes do corpo e a exterocepção dos (com os) aspectos mutáveis e persistentes dos objetos e acontecimentos extracorporais” (Costa, 2004, 61). Portanto, todo ato perceptivo é auto-informativo¹¹.

A abordagem ecológica do *self* criou uma série de experimentos que confirmam a capacidade perceptiva de se discriminar do ambiente e se comportar como um centro de ações intencionais. Nessa acepção funcionalista do *self* (ELEANOR GIBSON, 1995), privilegiam-se as diferentes informações que o *self* é capaz de obter por meio da percepção e interação com os objetos físicos e culturais.

Ulric Neisser (1988) chama de *self ecológico* o *self* diretamente percebido em relação ao seu ambiente físico imediato, capaz de estar consciente e atento. Ao perceber o ambiente, ele é capaz de se situar espacialmente e de agir como um ponto de vista subjetivo em relação ao seu entorno, por meio de padrões perceptivos que especificam sua posição e seu movimento. Desde o nascimento, a partir de seus movimentos e ações no ambiente, o bebê capta informações perceptivas que especificam seu próprio corpo como uma entidade única no mundo. A capacidade inata de diferenciação eu-mundo é atribuída a experiências perceptivas que especificam o corpo do bebê como oposto a outras entidades perceptivas do mundo. A experiência de tocar-se a si mesmo, de ouvir seus próprios sons, de visualizar seus próprios movimentos, são experiências perceptivas básicas que auto-especificam o corpo do bebê (ROCHAT, 2003).

Existem várias pesquisas que mostram a capacidade inata de se diferenciar e de se autopercebe¹². Philippe Rochat e colaboradores (ROCHAT & HEPOS, 1997) observaram

10 O termo ‘informação’ usado pelos psicólogos cognitivistas tem um significado diferente da semântica informacional de Dretske. Segundo Neisser (1988), na teoria ecológica da percepção de Gibson, a informação existe objetivamente e especifica as propriedades dos objetos e eventos em virtude dos princípios físicos e é obtida por aquele que percebe.

11 Como esclarece Jurandir Costa, na concepção ecológica, se “defende a ideia de que o comportamento movente e senciante dos animais é a maneira que eles têm de descobrir e utilizar os recursos do ambiente, (...) e, à medida que o indivíduo explora os recursos do ambiente apropriados à sua auto-regulação, auto-reprodução e satisfação, desenvolve a autopercepção de si” (COSTA, 2004, p. 61).

12 George Butterworth e colaboradores (1995) desenvolveram uma série de estudos sobre a propriocepção visual dos bebês para testar a capacidade do bebê de discriminar o movimento de seu próprio corpo do movimento dos objetos do ambiente. Esses estudos mostram que muito precocemente, antes mesmo de serem capazes de sentar, os bebês são capazes de ajustar sua estabilidade da postura em relação à estabilidade de seu entorno visual. No estudo do quarto movente (*moving room*), os bebês são presos a carrinhos e colocados num quarto suspenso do chão que se movimenta, de modo que a informação do fluxo óptico capta uma perda de balanço mesmo quando o bebê permanece numa postura estável. Observa-se que o bebê é capaz

o comportamento de fuçar em recém-nascidos (reconhecer com o focinho) (*rooting response*). Eles compararam o comportamento de fuçar em bebês com 24 horas de vida observando a estimulação tátil tanto originada pelo dedo do experimentador quanto a auto-estimulação que a criança espontaneamente fazia com um dos dedos de suas próprias mãos em contato com a bochecha. O resultado mostra que os recém-nascidos fuçam muito mais significativamente quando a estimulação é externa do que quando se auto-estimulam. Esse resultado confirma a hipótese de que os recém-nascidos não nascem num estado de confusão com o ambiente, mas, pelo contrário, são capazes de diferenciar entre estimulações e toques que são originados de seu próprio corpo (auto-estimulação) e de fontes externas (alo-estimulação). O bebê é capaz de discriminar entre eventos perceptivos que são gerados tanto pelo eu quanto pelo não-eu.

A partir dos dois meses, o bebê começa a se comportar, mesmo que de forma rudimentar, como um **agente situado** (ROCHAT, 2003), cujos movimentos são ações orientadas por objetivos específicos que se dirigem a aspectos particulares do ambiente, exibindo capacidade de discriminação e preferência.

Muito antes de ser capaz de se reconhecer como uma imagem ou rosto, o bebê se percebe como um centro de atividade, capaz de controlar suas ações e interações. É capaz não apenas de obter informações de seus estados mentais e corporais, mas de obter informações do sistema como um todo. Ele desenvolve progressivamente uma atitude de contemplação (*contemplative stance*) e reciprocidade frente ao mundo e a si mesmo, situando seu corpo em relação a outras entidades no ambiente (ROCHAT, 1999). A atitude contemplativa (ROCHAT, 2001) indica a auto-percepção de si como capaz de alterar o ambiente e o início do desenvolvimento da capacidade de controlar e escolher suas ações sobre o seu corpo e o ambiente. O bebê deixa de agir apenas de forma imediata e passa a explorar seletivamente a consequência de suas ações sobre o ambiente. Ao mesmo tempo, com o acúmulo das experiências, aprende um repertório novo de ações que torna sua interação muito mais complexa e o permite escolher entre ações variadas em função do objetivo visado. Ele aprende a controlar suas ações e a ter expectativas dos comportamentos do meio, sobretudo dos seres animados (EDWARD REED, 1995).

As pesquisas sobre imitação neonatal de Andrew Meltzoff e Keith Moore (1995) mostram que, ao final dos dois meses, com a emergência do sorriso social, das interações face-a-face e das protoconversações, o bebê desenvolve um comportamento imitativo

de distinguir entre o seu movimento (auto-movimento) e o movimento do ambiente (alo-movimento), e perceber a informação proprioceptiva visual como especificando uma mudança na localização espacial do seu corpo. Conclui-se que os bebês são capazes de fazer a distinção entre uma alteração de lugar de seu corpo e uma alteração do estado do ambiente.

sofisticado e voluntário e se torna capaz, não apenas de discriminar suas ações corporais do modelo que está imitando, como também de mapear seu espaço corporal em relação ao espaço corporal do outro, situando-se em relação ao modelo percebido que imitam. O bebê imita a dinâmica dos atos que observa. Ele necessita ver o ato em progresso para imitá-lo, pois ainda não é capaz de antecipar a sequência de movimentos do outro. Por meio das interações face-a-face, se estabelece entre o bebê e o adulto um 'espelho afetivo', com ressonância emocional, no qual o adulto responde reciprocamente ao bebê (ROCHAT, 2002). Este começa a ser capaz de combinar suas habilidades interativas e perceptivas e entrar num modo intersubjetivo de interação e de vida relacional (PASSOS-FERREIRA, 2011).

A atitude contemplativa é uma condição necessária na ontogênese para que o bebê chegue a desenvolver a perspectiva da primeira pessoa. Observa-se um continuum na experiência do bebê que, entre 2-9 meses, começa, gradativamente e com níveis crescentes de complexidade, a se comportar como um *sistema intencional* (DENNETT 1981b), com a emergência das ações intencionais, com objetivos funcionais que não são apenas aqueles inatos, e a criação de novas ações em função dos objetos encontrados no ambiente próximo. Portanto, o bebê que se recusa a pegar um objeto que está fora de seu campo de ação ou que utiliza seu pé para atingir um móvel suspenso diante de si, pode ser dito capaz não apenas de se diferenciar do objeto, mas também de adotar diferentes estratégias de ação em relação ao objeto e escolher os movimentos mais adequados para atingir seus objetivos. Ele explora seletivamente os provimentos dos objetos (*affordances*) que estão ao seu alcance. Ao final desse período, o bebê já pode ser considerado uma agência autônoma e coordenada (E. GIBSON, 1995), com capacidade para modular suas ações a partir das expectativas de comportamento do seu meio social.

No decurso do desenvolvimento emocional, é somente quando o bebê desenvolve a capacidade de se perceber como tendo um ponto de vista, ele é capaz de exibir rudimentos da perspectiva *forte* da primeira pessoa. Aos nove meses, o bebê se torna um **agente perspectivo** – o que Michael Tomasello (1995) chama de '*becoming perspectival*'. Com a emergência da locomoção, surge a possibilidade de se perceber o mundo de várias posições. A função da locomoção está diretamente relacionada com o desenvolvimento da habilidade de se autoperceber como uma perspectiva (REED, 1995), possibilitando um movimento que não se destina a funções específicas, como pegar objetos, mas a explorar o ambiente, e isso permite ao bebê explorar perspectivas alternadas. Surge também a capacidade de apontar, por meio das vocalizações e dos gestos usados para indicar os objetos. Os jogos imitativos também se tornam complexos e passa a ser suficiente para a criança observar o final do ato realizado pelo adulto para imitá-lo, exibindo a capacidade de

antecipar o processo. O bebê começa a consultar a face do outro em busca de informações sobre os recursos do ambiente e passa sistematicamente a integrar o olhar do outro em suas explorações pelo meio circundante. Ele tem a experiência de explorar pontos de vista distintos e promover padrões específicos de atividades com outros. Essa nova função possibilita que compartilhe os provimentos dos objetos (*affordances*) do ambiente com os outros, e se engaje em interações ternárias que incluem o bebê, o adulto, o objeto ou outro adulto (REED, 1995).

Portanto, a partir dos nove meses de vida, o bebê exhibe esboços do que vai se tornar futuramente a capacidade de representar sua própria perspectiva, de ter crenças e desejos sobre suas crenças e desejos. Os *sistemas intencionais de segunda ordem*, tal como concebido por Dennett, são característicos de criaturas que **não** têm o conceito de 'eu', mas se comportam como organismos que têm uma perspectiva de si em relação ao seu entorno. É dessa perspectiva egocêntrica que experimentam o mundo. Pode-se dizer que são conscientes de seus estados mentais e corporais, mas não são conscientes de serem possuidores desses estados. Portanto, não são autoconscientes, pois não têm o *conceito* de crenças ou desejos, nem o *conceito* de si mesmo como um sujeito de suas crenças e desejos. O bebê age de sua própria perspectiva sem saber que sua perspectiva difere de outras perspectivas. Autolocaliza-se frente aos objetos e pessoas a partir de sua posição espaço-temporal e age a partir desse ponto de vista tomando a si mesmo como centro. Seus estados são autolocalizados e referidos ao seu ponto de vista. No entanto, ainda não é objeto de seu pensamento e não é capaz de ter auto-representações das quais poderia estar consciente ou pensar sobre, pois ainda não adquiriu a autoconsciência reflexiva¹³.

O bebê de 9-12 meses é capaz de dar o passo intencional na direção do que vai se constituir, depois dos 3 anos, a perspectiva forte da primeira pessoa. Ele apresenta a habilidade rudimentar de atribuir tipos de estados mentais – tais como estados de atenção e de emoção – a outros agentes e interpretar o comportamento do agente como racional e dirigido a um objetivo que o permite prever sua ação futura em direção ao objetivo tendo por base uma nova situação (TOMASELLO, 1995).

Esse processo culmina aos 14-18 meses com o desenvolvimento da capacidade de se reconhecer no espelho (ROCHAT, 2003). A capacidade de identificar a própria imagem no espelho pode ser vista como o primeiro índice de objetificação do eu. É um índice não apenas de reconhecimento da imagem corporal, mas de que aquela imagem o identifica, se refere ao eu. Ou seja, ele não somente diferencia o seu ponto

13 Quando abordar a gênese do eu, veremos que estas representações só vão aparecer com a emergência do *self expandido* (*extended self*).

de vista da perspectiva exterior, mas também se torna capaz de perceber essa distinção. É nesse período que surgem condutas de embaraço relacionadas a contextos em que a criança recebe atenção prolongada do adulto sobre si mesmo (ROCHAT, 2002; 2003). O desenvolvimento cronológico mostra que o reconhecimento da imagem no espelho vem antes da utilização da palavra 'eu'. Existe, portanto, uma capacidade gestual implícita de se referir a si mesmo, antes mesmo da capacidade de se manifestar, linguisticamente, por meio do uso do pronome 'eu'. O reconhecimento da imagem no espelho como a imagem de si mesmo distinta de outras pode ser visto como sinal da capacidade de identificar uma representação de si mesmo e serve como um marco adaptativo do desenvolvimento do que vai ser chamado depois de autoconsciência reflexiva. Não obstante, ter um ponto de vista subjetivo e ser capaz de discriminar sua imagem entre outras não é suficiente para exibir a perspectiva *forte* da primeira pessoa.

No segundo ano de vida, com a aquisição da linguagem, a criança desenvolve a capacidade de compreender de um *modo mental* o agir intencional e se auto-perceber como um **agente representacional**, cujas ações são causadas por estados mentais prévios tais como crenças e desejos. E começa a entender que as ações podem mudar mentalmente, assim como as propriedades físicas do mundo. Eles compreendem, por exemplo, que um gesto declarativo de apontar pode alterar o estado de atenção de outro agente (FONAGY, 2005).

O desenvolvimento da perspectiva *forte* da primeira pessoa culmina com a aquisição da linguagem e a emergência da perspectiva da teoria da mente. Aos 3-4 anos, a criança desenvolve a habilidade de compreender e atribuir estados emocionais e complexos estados intencionais aos outros, como as falsas crenças, de entrar nos jogos de faz-de-conta e de fazer referência verbal a si mesmo. Surgem também as emoções secundárias como vergonha, orgulho, desprezo e culpa. A compreensão da agência em termos da causação mental passa a incluir a representação epistêmica dos estados mentais, emergindo o **agente meta-representacional**, cujas ações são compreendidas como estados mentais intencionais (crenças e desejos), que são construídas como representações conceituais do mundo (FONAGY, 2005). A posse das representações e das crenças capacita a criança a poder representar algo como falso ou verdadeiro. Esse processo culmina aos 4 anos de idade com a capacidade de passar no teste da crença falsa, que avalia a habilidade de representar o mundo mental do outro e, conseqüentemente, compreender a perspectiva do outro. Isso permite à criança compreender a referência a si mesmo (autoconsciência) e a seus estados mentais (autoconhecimento). Ela adquire a capacidade representacional de relatar as memórias de suas atividades intencionais e experiências numa organização coerente de causa e tempo, levando à emergência do *self expandido* no tempo, que é a

compreensão da sua história autobiográfica¹⁴.

A capacidade de pensar em si mesmo como possuidor desses pensamentos traz outras habilidades para a pessoa. A aquisição da autoconsciência nos torna capazes de pensar nos outros como possuidores de seus próprios pensamentos e como pontos de vista subjetivos diferentes dos nossos. Consequentemente, tornamo-nos capazes de refletir sobre nós como possuidores de nossos próprios desejos e ter o que chamamos de desejos ou pensamentos de segunda-ordem sobre nós mesmos. Esta capacidade, como esclarece Baker (1998), é a capacidade de se conceber como ‘si mesmo’, como um ‘eu’ distinto de todo o resto, e de formular um conceito sobre essa distinção, ou seja, é “não apenas ter pensamentos exprimíveis por meio do ‘Eu’, mas também conceber a si mesmo como o possuidor desses pensamentos”.

4. Autoconsciência reflexiva

No debate atual sobre a noção de autoconsciência, uma das questões que se coloca é se a informação que o eu obtém, desde o nascimento, sobre seus próprios estados pode ser chamada de autoconsciência ou autoconhecimento. Alguns afirmam que só podemos considerar o fenômeno da autoconsciência quando lidamos com seres de linguagem, capazes de conceituar linguisticamente sua experiência. Desde que Wilfrid Sellars (1997) propôs distinguirmos entre a consciência discriminativa (*awareness*) (‘ter a experiência de x’) e a consciência reflexiva (*consciousness*) (‘ter o conceito de x’, ‘saber que tipo de coisa é x’), existe em filosofia a tendência a separar a experiência subjetiva pré-linguística da experiência linguística reflexiva.

Muitos filósofos, como Dretske, fazem uma clara distinção entre pensamento (julgamento, crença) e sensação (experiência). A experiência de uma sensação (ouvir a Nona Sinfonia de Beethoven) não requer o conceito (saber o que é uma sinfonia) para saber do que se está consciente perceptivamente. Nos termos de Dretske, a distinção é feita entre a consciência do objeto (a experiência da propriedade x) e a consciência do fato (a crença de que isto é x). Nesse sentido, ter a experiência de K é diferente de ter pensamentos sobre K, assim como ter representações de fenômenos sensoriais difere da representação de fatos conceituais em crenças e julgamentos. A experiência não requer conceito, mas o conhecimento requer conceito. Os sistemas perceptivos têm a função de indicar a qualidade fenomenal da experiência por meio de funções herdadas

¹⁴ Sobre a noção de *self expanded* ver Passos-Ferreira (2007).

filogeneticamente e não podem ter o caráter representacional da experiência alterado. A função de indicar o conteúdo representacional é adquirida ontogeneticamente, por meio do aprendizado individual. O desenvolvimento individual pode mudar não o que a mente filogeneticamente deve representar, mas pode alterar o que a experiência pode representar com a aquisição de novas representações sensoriais e conceituais.

Para Dretske, a diferença entre a sensação e a cognição está no modo como a informação deve ser codificada. A condição que uma representação analógica (de uma sensação) deve satisfazer para ser qualificada como experiência é que a informação que ela tem a função de fornecer tem que estar disponível para ser digitalizada, ou seja, para ser representada conceitualmente (em crenças e intenções). Contudo, esse esquema lógico não explica como se passa do nível lógico da experiência para a reflexão, como se passa da sensação para a cognição. Em seu esquema, parece ser apenas a aquisição da linguagem que importaria na mudança de um nível a outro.

A abordagem desenvolvimentista mostra que, para além desse esquema lógico de interpretação, há um *continuum* de experiências que acontecem e precisam acontecer para que haja a transformação da consciência discriminativa em consciência reflexiva. Há uma gradação. Nesse sentido, a história das interações causais (da filogênese e da ontogênese) é fundamental para explicar como o comportamento discriminativo se transforma nas práticas sociais de justificação. Ou seja, como a 'presença ao mundo' se torna uma autoconsciência capaz de justificar suas ações.

Todas essas pesquisas mostram que, muito precocemente, os bebês aprendem coisas sobre os objetos que são alcançáveis no mundo físico e sobre suas capacidades em alcançá-los e como ajustar sua configuração e postura corporal. As transações com o mundo físico e social levam ao conhecimento do *self* como agente intencional e social. Tenderíamos a concordar com a psicologia gibsoniana e a fenomenologia que essas experiências geram informações não-conceituais sobre o eu, tanto de seus estados corporais e mentais quanto do sistema como um todo. Na abordagem ecológica, a percepção do mundo e a percepção do eu são dois aspectos do mesmo processo. Ao percebermos o ambiente, temos também informações sobre o eu. Mas não há ainda representação conceitual da entidade eu, temos apenas informações sobre nossos estados mentais e corporais, da perspectiva fraca da primeira pessoa.

Apesar de fazerem o diagnóstico correto da situação, a conceituam de maneira equivocada. Como vimos, as pesquisas mostram que os bebês obtêm informação sobre seu corpo como uma entidade diferenciada, discriminada de outros objetos e pessoas percebidos no ambiente, e que essa capacidade favorece o desenvolvimento da noção de eu. O bebê aprende com essas informações a modificar seu comportamento e desenvolve

outras formas de interação. Mas o tipo de auto-informação que a experiência rudimentar de si gera é diferente do autoconhecimento introspectivo que emerge com a perspectiva forte da primeira pessoa. A crença é uma condição para o conhecimento. Se definimos o conhecimento como dependente de conceitos não é possível falar de autoconhecimento antes de termos conceitos de eu, e o mesmo se aplica para a noção de autoconsciência. O fato de o bebê trazer informações analógicas sobre si mesmo, que serão futuramente digitalizadas, para usar a expressão de Dretske (1997), e transformadas em representações sobre o 'eu', não significa que o bebê tenha consciência de si como uma agência representacional ou meta-representacional.

O fenômeno apontado pela psicologia do desenvolvimento é que, mesmo que convençionemos que o autoconhecimento seja um conhecimento conceitual do eu e de seus estados, e, portanto, dependente da capacidade de usar símbolos, ele é determinado, em alguma medida, pelo modo como o bebê, desde o nascimento, processa as informações perceptivas que capta de seu ambiente físico e social. Os bebês têm informações que representam, de forma não-conceitual, seus estados mentais e corporais, que alteram as ações do eu, mesmo que não sejam capazes de se representar como possuidor desses estados. Há uma pré-história importante que conduz ao desenvolvimento da perspectiva forte da primeira pessoa. Quando os símbolos linguísticos são adquiridos, a criança já tem uma experiência anterior de pertencimento a um corpo que interage com objetos físicos e corporais. Somente um corpo que tem a experiência prévia de ser um agente no mundo pode vir a 'pensar sobre essa experiência'. Antes mesmo de ter conceitos, o bebê experimenta uma série de interações que fornecem informações, que são representações não-conceituais sobre o eu e seus estados. A perspectiva da primeira pessoa fraca, ainda não objetificada, é um centro de motivação da ação e a criança é capaz de se perceber como esse centro ativo que deflagra ações (E. GIBSON, 1995).

Sendo assim, a autoconsciência não é algo que irrompe automaticamente e simplesmente com a aquisição da linguagem. A capacidade de conceitualização do eu e de fazer referência a si mesmo emerge progressivamente no desenvolvimento. Existem níveis cumulativos que se consolidam a partir de fundamentos ecológicos do sentido implícito do eu e que, progressivamente, se tornam explícitos com a aquisição da linguagem e da capacidade reflexiva. Há uma transição fundamental, pois a forma como vivenciamos as interações pré-conceituais imprimem uma marca no modo como as interações linguísticas serão vivenciadas. Antes de ser capaz de conceptualizar e, portanto, de ter consciência de si como 'eu', ele tem acesso a uma série de informações sobre o 'eu' que é a experiência subjetiva da primeira pessoa. Tem a experiência de ser um agente que controla suas ações, se situa e se diferencia face ao ambiente, o que pode ser evidenciado

por seus comportamentos (imitação, protoconversações, ajustes corporais, localização em relação a outros objetos). O bebê que chora ou grita, ou aponta na presença de dor pode ser dito apresentar um ‘marcador de perspectiva’, ou seja, experimenta o mundo de sua posição encarnada. Porém, ainda não se concebe como possuidor desse estado, pois é um estado que ainda não foi auto-atribuído. Ainda não tem uma perspectiva de si como um ponto de vista dentre outros.

A criança se torna consciente de si mesmo e de suas próprias experiências quando os conceitos necessários à capacidade de autoconsciência são adquiridos. Sem os conceitos, a criança tem a experiência autoperceptiva, mas não tem recursos para compreender ou conhecer o que é estar consciente de si. A autoconsciência emerge a partir da aquisição do conceito de representação e referência de si. Se Baker e Dretske têm razão (e acredito que tenham), não é possível falar de objetificação do eu antes da emergência da capacidade de conceptualização do eu. A conceptualização do eu se expressa pela capacidade de fazer referência a si mesmo, seja por gestos, símbolos, palavras ou frases.

A perspectiva forte da primeira pessoa fornece a capacidade de pensar em nós mesmos de uma forma única, “de dentro”. Não pensamos em nós mesmos como um objeto ou uma entidade ‘self’, como um corpo ou um eu desencarnado a ser introspectado. O ‘eu’ não se refere a uma entidade, a qual tenho acesso direto, refere-se à pessoa em sua “concretude corporal”. Parafraseando Baker (1998), referir-me a mim mesma pelo pronome ‘eu’ não é diferente de referir-me à pessoa Cláudia. O que é particular acerca do ‘eu’ é que nos referimos a nós mesmos de um modo que os outros não podem, **de dentro**. Essa capacidade de pensarmos em nós mesmos ‘de dentro’ é a capacidade introspectiva presente na autoconsciência reflexiva¹⁵.

Referências

BAKER, Lynne Rudder. The first-person perspective: a test for naturalism. **American Philosophical Quarterly**, oct 1998, p. 327 (1).

BUTTERWORTH, George. Self as Object of Consciousness. In: ROCHAT, Philippe. **The self in infancy: theory and research**, Amsterdam: Elsevier, 1995.

CHALMERS, David, Facing up to the problem of consciousness, **Journal of Consciousness Studies**, 2(3), p. 200-19, 1995.

¹⁵ Para uma análise dos processos mentais que caracterizam a autoconsciência e o autoconhecimento ver Passos-Ferreira (2006).

- COSTA, Jurandir Freire. **O vestígio e a aura**. Corpo e consumismo na moral do espetáculo, Rio de Janeiro: Garamond, 2004.
- DRETSKE, Fred. **Naturalizing the mind**, Cambridge: MIT Press, 1997.
- FONAGY, Peter et al. **Affect regulation, mentalization and the development of the self**. London: Karnac, 2005.
- GALLUP, Gordon, Jr. Self-Recognition in primates: A comparative approach to the bidirectional properties of consciousness. **American Psychologist**, 32, p. 329-38, 1977.
- GIBSON, Eleanor & PICK, Anne. **An ecological approach to perceptual learning and development**. Oxford: Oxford University Press, 2000.
- GIBSON, Eleanor. Are We Automata? In: ROCHAT, Philippe (Ed.). **The self in infancy: theory and research**. Amsterdam: Elsevier, 1995.
- GIBSON, James. **The ecological approach to visual perception**. Boston: Houghton-Mifflin; reprinted, 1986.
- JONES, Doug. Evolutionary Psychology, **Annu. Rev. Anthropology**, 1999, 28:553-75.
- KRIEGEL, Uriah. Naturalizing Subjective Character. **Philosophy and Phenomenological Research**, 71, 2005, p. 23-57.
- MELTZOFF, Andrew & MOORE, M. Keith. A theory of the role of imitation in the emergence of self. In: ROCHAT, Philippe (Ed.). **The self in infancy**. Amsterdam: Elsevier, 1995.
- NAGEL, Thomas. What's Like to Be a Bat? In: NAGEL, Thomas. **Mortal questions**. Cambridge University Press, 1992.
- NEISSER, Ulric. Five kinds of self-knowledge. **Philosophical Psychology** 1, 35-59, 1988.
- PASSOS-FERREIRA, Cláudia. **As origens da interioridade: autoconhecimento e externalismo**, 2006. Tese (Doutorado em Saúde Coletiva). Universidade Estadual do Rio de Janeiro, Rio de Janeiro, 2006.
- PASSOS-FERREIRA, Claudia. Além das fronteiras da consciência: o self externalizado em Winnicott e Dennett. In: BENILTON BEZERRA JR, Francisco Ortega (Ed.). **Winnicott e seus intercessores**. Rio de Janeiro: Relume Dumará, 2007.
- PASSOS-FERREIRA, Claudia. O Desenvolvimento da Autoconsciência em Bebês. In: **Nascimento: antes e depois. Cuidados em rede**. ed. Curitiba: Honoris Causa, 2010. p. 133-44.

PASSOS-FERREIRA, Claudia. Os Processos Imitativos e a Emergência do Eu. **Polemica**, v.10, p. 613-23, 2011.

PASSOS-FERREIRA, Claudia. Intersubjetividade e autoconsciência. In: **Dimensões da intersubjetividade**. 1ed. São Paulo: Escuta, 2012. p. 163-79.

PASSOS-FERREIRA, Claudia, **As origens da interioridade**, no prelo.

PITT, David. Mental Representation. **The Stanford Encyclopedia of Philosophy** (Fall 2022 Edition), Edward N. Zalta & Uri Nodelman (Eds.). Disponível em: <https://plato.stanford.edu/archives/fall2022/entries/mental-representation/>.

ROCHAT, Philippe. Naissance de la co-conscience. **Intellectica**, 2002/1, 34, p. 99-123.

ROCHAT, Philippe. Five levels of self-awareness as they unfold early in life. **Consciousness and Cognition**, 12, 2003, p. 717-31.

ROCHAT, Philippe. **The infant's world**. New York: Harvard University Press, 2004.

ROCHAT, Philippe & HESPOS, S. J. Differential rooting response by neonates: evidence for an early sense of self. **Early development and parenting**, 6(2), 150.1–150.8. 1997.

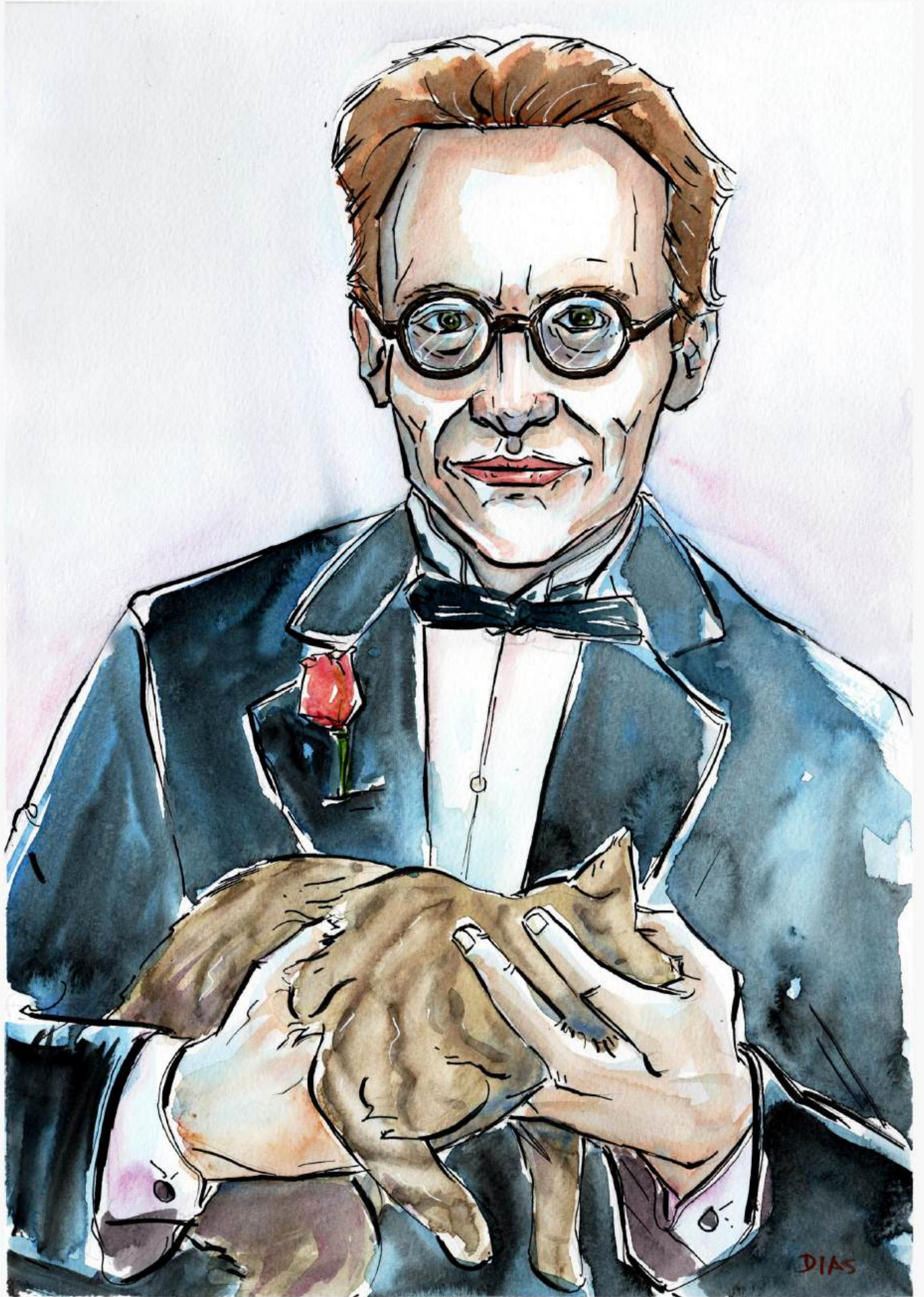
ROCHAT, Philippe & STRIANO, Tricia. Perceived self in infancy. In **Infant Behavior and Development**, 23, 2000, p. 513-30.

SELLARS, Wilfrid. **Empiricism and philosophy of mind**. Harvard University Press, 1997.

TOMASELLO, Michael. Understanding the Self as Social Agent. In: ROCHAT, Philippe (Ed). **The self in infancy**. Amsterdam: Elsevier, 1995.

TOMASELLO, Michael & CALL, J. **Primate cognition**. New York: Oxford University Press, 1997.





EM BUSCA DE UM SENTIDO PLAUSÍVEL DE *SELF*



Fabiense Pereira Romão

Leonardo Ferreira Almada

Resumo:

Neste artigo, nos dedicaremos a compreender se há um sentido de *self*. E, no caso de haver, qual seria esse sentido. Promoveremos um debate a partir de perspectivas distintas de pensadores clássicos e contemporâneos acerca do problema do *self* vislumbrando o horizonte de uma definição geral. Buscaremos verificar quais seriam as condições para tal designação e, no caso de haver, verificar a força de tal designação. Nos dedicaremos a buscar se há um sentido de *self* no ato de conhecer. Em termos concretos, compõe o escopo da nossa investigação a tentativa de compreender como nossas percepções, sensações corporais, sentimentos emocionais, pensamentos e memórias parecem se conectar para compor uma unidade coerente de experiência. Buscaremos compreender o que há por trás da compreensão de que nós somos nós mesmos ao longo do tempo. Para tanto, nossa tarefa é investi-

Abstract:

In this article, we will focus on understanding whether there is a sense of *self*. And, if there is, what would that meaning be. We will promote a debate from different perspectives of classical and contemporary thinkers about the problem of the self, looking at the horizon of a general definition. We will seek to verify what the conditions for such designation would be, and, if any, to verify the strength of such designation. We will dedicate ourselves to searching if there is a sense of *self* in the act of knowing. In concrete terms, the scope of our investigation comprises the attempt to understand how our perceptions, bodily sensations, feelings, thoughts, and memories seem to connect to compose a coherent unit of experience. We will seek to understand what lies behind the understanding that we are ourselves over time. Therefore, our task is to investigate whether something in this sense. In this endeavor, in

gar se algo neste sentido. Nesta empreitada, em busca do sentido do *self*, investigaremos qual seria o papel do corpo na composição do *self* e da experiência da subjetividade vivenciada na experiência consciente. Encaminharemos também o problema no sentido de questionar se agimos no cotidiano ancorados em um *ego-self*. Buscaremos confrontar noções que sustentam a evidência do *self*, noções que o concebem como fixo, imutável e independente, e também noções que sustentam sua inexistência. Estamos diante de um clássico problema filosófico sobre o qual não encontramos univocidade, porém ainda é capaz de instigar imensamente aqueles que se dispõem a explorar o seu solo teórico e conceitual.

Palavras-chave:

Self; sentido de *self*; *selves*.

search of the meaning of the *self*, we will investigate what would be the role of the body in the composition of the *self* and the experience of subjectivity lived in the conscious experience. We will also address the problem in the sense of questioning whether we act in everyday life anchored in a *ego-self*. We will seek to confront notions that support the evidence of the *self*, notions that conceive it as fixed, immutable, and independent, and notions that support its inexistence. We are facing a classic philosophical problem about which we do not find univocity, but it is still capable of immensely instigating those who are willing to explore its theoretical and conceptual ground.

Keywords:

Self; sense of *self*; *selves*.

1. Introdução

Neste artigo, nos dedicaremos a compreender se há um sentido de *self*. E, no caso de haver, qual seria esse sentido? Promoveremos um debate a partir de perspectivas distintas de pensadores clássicos e contemporâneos acerca do problema, com a finalidade de tentar extrair uma definição geral do que possa vir a ser designado de *self*. Buscaremos verificar quais seriam as condições para tal designação, e, no caso de haver, verificar a força de tal designação. Também chegaremos a tangenciar em alguma medida as abordagens da psicologia budista com o propósito de enriquecer o debate.

A título de esclarecimento, deliberamos pelo emprego do termo '*self*' em inglês pelo fato de acreditarmos que não é possível encontrar equivalência satisfatória caso o traduzíssemos para a língua portuguesa. Teríamos algo aproximado como 'si mesmo'. E,

neste caso, estaríamos limitados claramente no que diz respeito a um problema central do artigo e todos seus desdobramentos, a saber: há um sentido de *self* no ato de conhecer?

Neste sentido, na esteira desta questão, em termos concretos, compõe o escopo da nossa investigação a tentativa de compreender como nossas percepções, sensações corporais, sentimentos emocionais, pensamentos e memórias parecem se conectar para compor uma unidade coerente de experiência. O que nos conduz à compreensão de que nós somos nós mesmos? Seria a identidade pessoal subsidiada pelo *self*? Afinal, o que nos faz ser e/ou reconhecer o que nós somos ao longo do tempo? Seriam nossos corpos, nossos cérebros, feixes de percepção, substância(s) ou processo(s) que nos identifica(m) e individualiza(m)? Porque parecemos agir e/ou perceber em torno de uma perspectiva unificada, ainda que as percepções e sensações transcorram com evidência em condições de transitoriedade.

Também incluiremos no rol das investigações, na busca do sentido do *self*, qual seria o papel do corpo na sua composição e da experiência da subjetividade. Levando em conta o que acontece dentro e fora dos limites do organismo momento a momento (percepções, sensações, pensamentos, memórias etc.), seria possível afirmar que há um *eu* ancorado a cada um desses momentos? Haveria uma base neural, um estado biológico que seria constantemente reconstituído? Para tanto, nos filiaremos em alguma medida ao aporte teórico e conceitual oriundo das perspectivas corporificadas do *self*, ainda que o problema não alcance univocidade no interior destas propostas.

Será que é possível explicar como nós agimos no cotidiano ancorados em um suposto *eu*? É possível localizar um *self* único, independente e duradouro? É possível localizá-lo na nossa experiência momentânea, no 'aqui e agora'? A grande questão é que, a depender das premissas assumidas, podemos nos conduzir a sua evidência ou não. E a partir da inferência alcançada, o impacto na visão de mundo, mente e consciência podem ser significativamente alterados.

2. UM SENTIDO DE *SELF*?

Iniciaremos nosso debate a partir de uma noção de *self* que o concebe como evidente. Para tanto, recuperamos e destacamos Dennett (1991; 1992) por meio de uma analogia exposta em sua noção de *self*, quando levantou a crença no fato de que as experiências parecem não existir na ausência de um *self* ou sujeito que as submete, assim como as

ondas oceânicas não podem existir sem o mar através do qual elas se movem. A partir dessa visão geral de Dennett (1991; 1992), contrastaremos noções díspares aspirando a ampliar os horizontes conceituais e teóricos relativos ao problema do sentido do *self*.

Começamos com a clássica abordagem de David Hume (1713-1784), para o qual a tentativa de descrever o *eu* deveria tomar como fundamento seu modelo empirista de investigação, segundo o qual nenhuma conclusão poderia ser admitida, em última instância, sem o crivo da experiência e da observação. Hume exclui, como um falso problema, a investigação do *eu*, dada a dificuldade decorrente da impossibilidade objetiva de seu sistema filosófico conceber a existência autônoma de uma mente substancial, colocando à prova a natureza da alma por meio do seu princípio da cópia. Hume, não encontrando um referente empírico seguro para o *eu*, se viu obrigado a descrevê-lo como uma mera ficção.

Talvez, uma das construções empíricas mais relevantes para a filosofia empirista de Hume tenha sido apresentada na estrutura da relação que há entre as impressões e as ideias: as ideias são ‘cópias mais fracas’ das impressões. Na *Investigação sobre o Entendimento Humano*, no Livro 1, Parte 1, Seção 1, Hume expõe essa relação entre os dois tipos de percepção:

Todos os materiais do pensamento são derivados da sensação externa ou interna, e à mente e à vontade compete apenas misturar e compor esses materiais. Ou, para expressar-me em linguagem filosófica, todas as nossas ideias, ou percepções mais tênues, são cópias de nossas impressões, ou percepções mais vívidas.

Este princípio aparece nos domínios da literatura como princípio da cópia, consoante o qual toda ideia é cópia de uma impressão. Hume lançou mão de dois argumentos complementares procedendo a uma distinção entre as percepções simples e complexas: (i) as primeiras basearam-se naquelas percepções que não pressupõem nenhum tipo de distinção e separação; (ii) as segundas, por sua vez, admitem distinção e separação, por serem compostas por percepções simples.

Este terreno conceitual nos desperta o interesse em razão da atenção de Hume voltada para os “filósofos que construíram raciocínios meticulosos para mostrar que nossas percepções seriam inerentes a uma substância material ou a uma substância imaterial”, como aparece no Livro 1, do *Tratado da Natureza Humana*, Seção 5.

Para Hume, em conformidade com seu critério de decisão, o princípio da cópia, se houvesse algo como a ideia de substancialidade de nossas mentes, deveria também haver uma impressão dessa substância. Para Hume (1739), as impressões são de dois tipos: sensações e paixões. As primeiras são oriundas dos órgãos dos sentidos e apresentam cheiros, cores, sons, volume etc. As segundas apresentam os sentimentos de amor, empatia, raiva, aflição etc. Hume dirá que nenhuma dessas características pode ser identificada a uma substância, pelo menos no sentido que a tradição deu a essa noção. Noção esta que sustentava a substancialidade da alma ou mente substancial, na qual as impressões estariam ligadas.

Hume refutará a noção de substância por meio do argumento de que a provável substancialidade de nossas mentes não é uma noção adquirida empiricamente e, para os que discordam disso, o filósofo lançou o seguinte desafio, presente no *Tratado da Natureza Humana*, Livro 1, seção 5:

Gostaria que aqueles filósofos que afirmam que possuímos uma ideia da substância de nossas mentes nos apontassem a impressão que produz essa ideia, e que nos dissessem distintamente como tal impressão opera, e de que objeto deriva. É ela uma impressão de sensação ou de reflexão? É agradável, dolorosa ou indiferente? Acompanha-nos em todos os momentos, ou só aparece a intervalos? Se a intervalos, em que momentos, sobretudo aparecem, e que causas a produzem?

Hume mostra aí seu ponto de vista segundo o qual a ideia de uma mente substancial, por não possuir um referente claro na impressão, inviabilizaria o princípio da cópia, tornando a questão da substancialidade da mente ininteligível.

Em suma, sob a égide da tradição, e a partir de seu julgamento, a alma seria uma substância à qual todas as percepções seriam inerentes. Para a tradição, estas percepções, seriam, antes, conhecidas exclusivamente no âmbito do intelecto, conceitualmente. Hume teve a coragem de romper com a tradição, ainda que seu sistema filosófico, por meio de seu modelo empirista, tenha posto um limite muito específico às suas investigações, sobretudo no que diz respeito à questão da natureza do *self* ou do *eu*.

Hume aniquilou um sentido de *self*, vinculando o *eu* no máximo a um feixe de percepções, no qual nenhuma conclusão metafísica poderia ou pode ser aceita. David Hume, no *Tratado da Natureza Humana*, Livro 1, seção 6:

A mente é uma espécie de teatro, onde diversas percepções fazem sucessivamente sua aparição; passam, repassam, esvaem-se, e se misturam em uma infinita variedade de posições e situações. Nela não existe, propriamente falando, nem simplicidade em um momento, nem identidade ao longo de momentos diferentes, embora possamos ter uma propensão natural a imaginar essa simplicidade e identidade. Mas a comparação com o teatro não deve enganar. A mente é constituída unicamente pelas percepções sucessivas; e não temos a menor noção do lugar em que essas cenas são representadas ou do material de que esse lugar é composto.

Hume se encontra diante da impossibilidade de explicar como nossas percepções sucessivas em nosso pensamento ou consciência são unificadas. Destarte, assume uma posição atomística em relação às percepções ao descrever a mente como feixe de percepções. Para Hume, as percepções existem individualmente, por si próprias, e, por isso, quando estão em conjunto, em um feixe, podem formar apenas um *eu* fictício. Na esteira desta concepção, robustas correntes filosóficas defendem a indefinição do *self*, como, por exemplo, a apresentada por Varela, Thompson & Rosch (2003), os autores do livro *The embodied mind*. Neste sentido, tanto Hume (1739) quanto Varela, Thompson & Rosch (2003) refutam a ideia de tratar o *self* como entidade ou substância.

3. INDEFINIÇÃO DO SELF

Ao perguntarmos o que é esse *self*, visitamos o capítulo de *The embodied mind* intitulado o *Centro do Turbilhão ou Eu da Tempestade* (VARELA, THOMPSON & ROSCH, 2003, p. 73), no qual a posição sustentada, além de ir ao encontro da defesa de Hume, traz novos elementos para o debate. A discussão trazida à baila por Varela, Thompson & Rosch, em sinergia com Hume, lança a seguinte questão: “o que é esse *self*, esse centro do *ego* que aparece e desaparece, que parece tão frágil, tão familiar ainda que tão difícil de ser compreendido?” (VARELA, THOMPSON & ROSCH, 2003, p. 73).

Através da constatação de que a todo momento de nossa vida alguma experiência está ocorrendo, segundo os autores de *The embodied mind* (2003), nos defrontamos com a seguinte contradição. Por um lado, a experiência, ainda que tenha uma atenção superficial, parece nos mostrar a sua transitoriedade e sua dependência de uma situação em particular. Ou seja, momento a momento, vemos, tocamos, sentimos gostos, cheiramos, pensamos etc. Nos sentimos dominados pelas emoções quando elogiados ou criticados, quando perdemos um ente querido etc. Podemos ficar com medo, irritados, cansados, perplexos etc. Deste modo, averiguamos que não há a vivência de uma experiência que seja permanente e independente de um contexto particular (VARELA, THOMPSON & ROSCH, 2003, p. 73).

Contudo, por outro lado, nos sentimos convencidos em relação às nossas identidades. Assim, nossas personalidades, planos e antecipações, memórias e recordações parecem se unirem em um centro a partir do qual investigamos o mundo e sobre o qual nos encontramos. Como esse ponto de vista coerente seria possível? Estaria esse ponto de vista fixado em um *self* único, independente e duradouro no mundo da experiência? Ao se debruçarem sobre esta questão, Varela, Thompson & Rosch (2003) promoveram a associação entre três grandes fontes, a saber: as ciências cognitivas, a filosofia e a tradição budista de meditação da atenção-consciência.

Na defesa da indefinição do *self*, tomam como ponto de partida a clássica passagem de David Hume, *Tratado da Natureza Humana*, Livro 1:

Eu, quando entro muito intimamente no que denomino *eu* mesmo, sempre tropeço em alguma percepção específica, de calor ou frio, luz ou sombra, amor ou ódio, dor ou prazer. Nunca sou capaz de surpreender a mim mesmo, em momento algum, sem uma percepção e nunca sou capaz de observar coisa alguma além da percepção. Quando minhas percepções são removidas por um tempo qualquer, como durante o sono profundo, fico insensível de mim mesmo e se pode verdadeiramente dizer que não existo. E se todas as minhas percepções fossem removidas pela morte e eu não pudesse pensar, sentir, ver, amar, odiar depois da dissolução do meu corpo, eu seria totalmente aniquilado, e não consigo imaginar o que mais haveria de ser preciso para fazer de mim uma perfeita não entidade. Se alguém, depois de uma reflexão profunda e imparcial, julgar que tem uma definição diferente de si mesmo, devo confessar que não terei mais condições de argumentar com ele. Só poderei admitir que ele possa ter tanta razão quanto eu, e que nós dois somos essencialmente diferentes nesse pormenor. Ele talvez possa perceber algo simples e contínuo a que denomina si mesmo; embora eu tenha certeza de que não existe tal princípio em mim.

O ponto central da tese de David Hume é o de que o *self* no máximo seria um conjunto de várias impressões e ideias atreladas a uma referência ficcional do *eu*. Portanto, se alguma impressão dá origem a uma ideia de *self*, esta impressão deveria continuar invariável todos os anos de nossas vidas. Como se supõe, consoante a passagem exposta no trecho, não há impressão constante e invariável. Por isso, David Hume, incapaz de encontrar o *self* em sua reflexão, recua negando a sua existência.

A investigação acerca do *self* por parte dos autores de *The embodied mind* (2003) vai contar com as relevantes contribuições das práticas contemplativas orientais. Neste sentido, a prática da atenção-consciência figura como uma promissora estratégia de desenvolvimento gradual da habilidade de estar presente com a mente e o corpo nas experiências cotidianas. Segundo Varela, Thompson & Rosch (2003, p. 74), os meditadores

iniciantes ficam geralmente surpresos com a atividade tumultuada de suas mentes, uma vez que sentimentos, medos, percepções, pensamentos e todo outro tipo de “conteúdos mentais perseguem um ao outro interminavelmente como um gato querendo morder seu próprio rabo (VARELA, THOMPSON & ROSCH, 2003, p. 74).

À medida que os meditadores ampliam suas habilidades na prática, começam a ter o ‘*insight*’ da mente experienciada, e, portanto, constataam o quão as experiências são transitórias (VARELA, THOMPSON & ROSCH, 2003, p. 74). No seio destas constatações vem à tona a tensão entre o sentido transitório do *self* na experiência e o insucesso em encontrá-lo através da reflexão (VARELA, THOMPSON & ROSCH, 2003, p. 75).

A discussão oferecida pelos autores de *The embodied mind* apresenta David Hume como um dos pontos de partida. Entretanto, avança significativamente no solo teórico acerca do sentido do *self*. Na seção intitulada *Prestando Atenção no Mundo*, capítulo 6, os autores lidaram com o potencial drama de concebermos a perda do sentido do *self*, alegando que, neste caso, é como se perdêssemos o que há de mais precioso e familiar para nós.

Uma das questões fundamentais de Varela, Thompson & Rosch (2003, p. 138-9) gira em torno da crítica à insistência na defesa de um *ego-self*. Depois de fazerem um longo percurso em busca de um sentido do *self*, chegam à inferência na qual a sua indefinição prevalece. No entanto, revestidos da vontade de trazer algum alento, vão dizer que não é preciso duvidar que o mundo possua sua estabilidade. Isso porque, neste ponto especificamente, na tese da ausência do *self*, podemos ficar aflitos em relação ao modo como vemos o mundo.

Neste sentido, quando o *self* deixa ser um ponto de referência, as pessoas podem começar a ver espectros de solipsismo, só eu existo no mundo, ou no campo do subjetivismo e idealismo: eu crio o mundo dentro de mim (VARELA, THOMPSON & ROSCH, 2003, p. 139). Contudo, os autores de *The embodied mind* asseguram que não é esse o caminho a ser seguido. Trata-se, apenas, de um caminho no qual o *self* não serve mais de âncora em oposição às abordagens que estão literalmente centradas no *self* e, por consequência, ao deixarem de discutir o *self* pessoal, se depararam com o relevante problema filosófico sobre qual seria o fundamento último do mundo.

4. SELF CONCEBIDO COMO EVIDENTE

Agora lidaremos com perspectivas sobre as quais o *self* é concebido como evidente. Resgataremos um trecho de Damásio (2011, p. 25), em que Willian James em franco confronto com David Hume se vê impelido a expressar uma crítica memorável defendendo a existência do *self*. Damásio (2011) ao abordar a mente consciente privilegia a evidência do *self*. Na visão de Damásio (2011, p. 20), “a mente consciente surge quando um processo do *self* é adicionado a um processo mental básico”, levando-o a sustentar que “quando não ocorre um *self* na mente, essa mente não é consciente, no sentido próprio do termo” (DAMÁSIO, 2011, p. 20).

Afinal, que crítica foi essa? William James atacou um ponto fundamental da perspectiva humeana no que diz respeito à sua posição em relação ao *self*. Hume recua em relação à existência do *self* por não ter conseguido lidar com ‘a estranha mistura’ de unidade e diversidade nele contida. James chamará atenção para esta questão do “núcleo de uniformidade que permeia todos os ingredientes do *self*” (DAMÁSIO, 2011, p. 25). Rees-tamparemos a famosa passagem de David Hume mostrando o contraditório indicado por Willian James. A essência da discordância é analisada em James, *Principles*, I, compreendido entre as páginas 350-352. Eis a afirmação de Hume e a resposta de James:

Hume, *Tratado da Natureza Humana*, Livro 1: Eu, quando entro muito intimamente no que denomino *eu* mesmo, sempre tropeço em alguma percepção específica, de calor ou frio, luz ou sombra, amor ou ódio, dor ou prazer. Nunca sou capaz de surpreender a mim mesmo, em momento algum, sem uma percepção e nunca sou capaz de observar coisa alguma além da percepção. Quando minhas percepções são removidas por um tempo qualquer, como durante o sono profundo, fico insensível de mim mesmo e se pode verdadeiramente dizer que não existo. E se todas as minhas percepções fossem removidas pela morte e eu não pudesse pensar, sentir, ver, amar, odiar depois da dissolução do meu corpo, eu seria totalmente aniquilado, e não consigo imaginar o que mais haveria de ser preciso para fazer de mim uma perfeita não entidade. Se alguém, depois de uma reflexão profunda e imparcial, julgar que tem uma definição diferente de si mesmo, devo confessar que não terei mais condições de argumentar com ele. Só poderei admitir que ele possa ter tanta razão quanto eu, e que nós dois somos essencialmente diferentes nesse pormenor. Ele talvez possa perceber algo simples e contínuo a que denomina si mesmo; embora eu tenha certeza de que não existe tal princípio em mim.

Eis a resposta de James, em *Principles*, I, páginas 350-2:

Mas Hume, depois de fazer esse bom trabalho introspectivo, joga fora o bebê junto com a água do banho e alça um voo tão extremado quanto os filósofos substancialistas. Assim como eles dizem que o *self* nada mais é do que unidade, unidade abstrata e absoluta, também Hume diz que o *self* nada mais é do que diversidade, diversidade abstrata e absoluta; enquanto na verdade ele, o *self*, é essa mistura de unidade e diversidade que nós mesmos já constatamos ser tão fácil distinguir. Ele nega que esse fio de semelhança, esse núcleo de igualdade que percorre os ingredientes do *self* exista mesmo como uma coisa fenomênica.

Damásio (2011; 2012; 2015; 2018), mais que reforçar a hipótese de não tratar o *self* como entidade ou substância, inspira-se em Willian James, sustentando que a constituição da mente consciente deve privilegiar, de modo significativo, a noção de *self*.

Damásio tem apresentado em suas obras notáveis relações entre a mente, o *self* e a consciência e, conforme a obra *E o Cérebro Criou o Homem*, tem sustentado que na ausência de um *self* na mente, não ocorre geração da subjetividade, uma característica eminente da mente consciente que lhe assegura identidade pessoal. Isso significa dizer, basicamente, que o *self* é o responsável pela garantia da consciência na mente. Para ambos, tanto para Damásio quanto para James, “o *self* apresenta firmes alicerces biológicos, que não devem ser confundidos com uma entidade metafísica conhecedora” (DAMÁSIO, 2011, p. 25).

Da união de filósofos e neurocientistas, a partir desta orientação, a visão acerca do *self*, além de ter sido modificada e expandida, passou a incluir diferentes aspectos do *self*, lançando luz à sua importância para a construção da mente consciente. Os contemporâneos trabalhos da psicologia e da filosofia da mente, em sinergia com o significativo desenvolvimento da biologia geral, biologia evolucionária e da neurobiologia, ampliaram de maneira significativa o legado conceitual do problema em questão (DAMÁSIO, 2011, p. 25). Passamos também a contar com dados destas fontes com a finalidade de problematizar o sentido do *self*.

De acordo com Damásio (2011, p. 25), a base neural da consciência, para ser elucidada abrangentemente, requer de modo imperioso a explicação acerca dos conceitos de *self*-objeto ou *eu*-material e o *self*-conhecedor. Para isso, Damásio (2011, p. 25) recuperou a suposição de James, *Principles*, I, capítulo 2, que dizia que o *self*-objeto ou *eu*-material era a soma de tudo que um homem poderia chamar de seu: “não só seu corpo e suas faculdades psíquicas, mas também suas roupas, sua esposa e seus filhos, além de antepassados e amigos, reputação e obras, terras e cavalos, iate e conta bancária”.

Damásio (2011, p. 22), adiante, alinha-se a James em um ponto de maior concordância ainda, qual seja: “o que permite que a mente saiba que existe esses domínios e estes pertençam aos seus proprietários mentais” (corpo, mente, passado e presente e todo o resto) é a percepção de que qualquer um desses elementos gera emoções e sentimentos, e, portanto, em conformidade com esta ideia, os sentimentos ensejam a separação entre os conteúdos que pertencem ao *self* e os que não pertencem.

4.1. SELF CORPORIFICADO, EMOÇÃO E SENTIMENTOS

Ao contrário do senso comum, Damásio procede a uma distinção entre emoção e sentimento de modo não ortodoxo. O senso comum com poucas variações costuma sinonimizar esses termos. As emoções são, de modo geral, alterações fisiológicas no estado do corpo decorrentes dos acontecimentos internos e externos ao corpo, tais como: estímulos nos ambientes (ameaças, ambientes seguros ou inseguros), alterações nas vísceras ou estruturas internas (músculos e tendões, por exemplo), recordações e/ou memórias, pensamentos, projeções de um futuro antevisto etc.

Para Damásio (2012, p. 141), para que uma emoção ocorra é necessário que determinados núcleos subcorticais e o córtex cerebral sejam ativados. Tudo isso em íntima relação do corpo com o cérebro e com ambiente. Outra característica importante das emoções são seus padrões de respostas fisiológicas. Neste sentido, as respostas fisiológicas de uma pessoa, por exemplo, que tem medo de barata e outra que tem medo de serpente são similares. Os pelos ficam eriçados, a pupila dilatada, o esfíncter se fecha etc. Embora os objetos sejam distintos, o medo desencadeia, portanto, respostas corporais que seguem padrões determinados previamente pela estrutura do organismo.

Já o sentimento aparece como a experiência e interpretação das alterações fisiológicas decorrentes da experiência emocional. Por isso, Damásio (2012) lança mão com frequência do termo sentimento-emocional com a finalidade de explicitar que o “fundamento do sentir de uma emoção é a experiência dessas alterações em interação com as imagens mentais que iniciaram o ciclo” (DAMÁSIO, 2012, p. 140). Infere-se a partir daí que o sentimento é uma dimensão privativa que não abdica da história de vida de cada indivíduo. Vejamos, então, que enquanto a emoção traz características públicas e detectáveis por terceiros, como por exemplo, o rosto ruborizado por vergonha, os sentimentos são privados e só podem literalmente ser conhecidos quando são compartilhados.

Parece-nos que Damásio (2012) no capítulo 7, *Emoções e Sentimentos*, propõe ao *self* o papel de conduzir a transição dos estados mentais não-conscientes para os estados mentais conscientes. E, no bojo desta concepção, o sentimento da emoção figura como os mecanismos que atuam no sentido de auxiliar nesta tarefa através da discriminação, destaque e compreensão dos objetos internos e externos que nos afetam. Muito provavelmente, por isso, se referiu à separação entre os conteúdos que pertencem ao *self* e os que não pertencem.

Ainda tratando da dimensão corporificada do *self*, Damásio (2011, p. 22) chama atenção para a dimensão material do *self*, veiculado pela expressividade do corpo vivo enquanto centro de representação do *self*-material, isto é, do *self*-objeto, como uma coleção dinâmica de processos neurais integrados, que encontra expressão em uma coleção dinâmica de processos mentais integrados a este dinâmico corpo vivo. Assim, defende a indispensabilidade da corporeidade na constituição do *self*.

Conforme Damásio (2011, p. 23), o *self*-conhecedor apresenta-se menos coeso em termos mentais ou biológicos em relação ao *self*-objeto, dada sua natureza sutil, com frequência dissolvida no fluxo da consciência, o que o torna uma presença mais difícil de ser definida. Entretanto, isso não diminui a importância do *self*-conhecedor perante a consciência. Damásio (2011, p. 23) enfatiza seu ponto de vista ao defender que o *self* sujeito-e-conhecedor além de ser uma presença muito real, cumpre um papel crucial na guinada histórica da evolução biológica à qual o homem pertence.

Importante ressaltar o entendimento acerca da imbricada relação existente entre o *self* sujeito-e-conhecedor e o *self*-objeto. Aqui Damásio aponta para uma natureza estratificada da vida mental. Também se opõe a uma possível dicotomia levantada na relação entre *self*-objeto e o *self*-conhecedor. Neste ponto, em especial, citamos Damásio (2011, p. 23):

Podemos imaginar que o *self* sujeito-e-conhecedor está, por assim dizer, sobre o *self*-objeto, assim como uma nova camada de processos neurais dá origem a mais uma camada de processamentos mentais. Não há dicotomia entre *self*-objeto e *self*-conhecedor; o que existe é continuidade e progressão. O *self*-conhecedor tem seu alicerce no *self*-objeto.

Ao tratarmos da natureza estratificada da vida mental, decorre a afirmação de que não há dicotomia entre as 'camadas'. O que existe é continuidade e progressão, e, por consequência, a compreensão é balizada com a concepção de que a natureza é inteiramente composta de elementos, processos, forças e campos físicos, dividindo-se em níveis distintos de complexidade.

Em consonância com esta perspectiva, expõe Almada (2017), esses “diferentes níveis de complexidade se estendem da esfera mais básica, na qual se incluem as partículas elementares, até o nível de maior complexidade, a exemplo dos seres mentais conscientes”, implicando que cada um dos objetos e fenômenos no âmbito de uma hierarquia determinada ocupa um lugar específico segundo sua complexidade organizacional. Desta forma, cada um dos níveis superiores contém instâncias que não são encontradas nos níveis anteriores.

4.2. SELVES DE DAMÁSIO

Neste painel teórico identificamos uma forte relação com a hipótese de Damásio (2011, p. 225-6), para quem o *self* é construído em estágios. Para Damásio (2011, p. 227), a tentativa de ampliar a compreensão acerca da construção da mente consciente requer um estudo minucioso de seus estágios. Desta maneira, Damásio (2011) ao priorizar o *self* não pretende diminuir ou negligenciar a abrangência da mente em si mesma. Pelo contrário, defende que a mente consciente ao prevalecer na história da evolução natural aperfeiçoou os mecanismos de regulação da vida contando com a evolução dos estágios do *self*. A abordagem proposta de acordo com esta perspectiva é a de que não estamos mais tratando apenas de *self*, mas, sim, de *selves*.

Para Damásio (2011, p. 25), é fundamental pôr em relevo o papel do *self* na construção da mente consciente, desde que haja atenção em relação ao fato de que nenhum mecanismo ou dispositivo isolado pode gerar a mente consciente. Assim, outra hipótese que endossa o tratamento que o *self* deve receber encontra-se pautado no fato de que uma mente em estado de vigília deve contar com o cérebro construindo a consciência a partir da geração de um processo do *self* (DAMÁSIO, 2011, p. 225-6).

Em relação aos *selves*, Damásio (2011, p. 226) apresenta os três estágios do *self*, a saber: (i) *protosself*, (ii) *self* central e (iii) *self* autobiográfico. O primeiro e mais simples estágio consiste na capacidade de reunir as imagens que descrevem os aspectos relativamente estáveis do corpo, gerando sentimentos espontâneos do corpo vivo, os quais Damásio (2011) chamou de sentimentos primordiais. Este estágio tem origem na parte do cérebro que representa o organismo (*protosself*).

O *protosself* foi descrito por Damásio (2011, p. 235-6) como: “uma coleção integrada de padrões neurais separados que mapeiam, momento a momento, os aspectos mais

estáveis da estrutura física do organismo”. Damásio (2011) enfatiza que os mapas do *protossself* são característicos não só por gerarem imagens corporais, mas também por cumprirem a função de gerar imagens corporais sentidas. Esses sentimentos primordiais encontram-se espontaneamente presentes no cérebro normal acordado (DAMÁSIO, 2011, p. 236).

A indispensabilidade do corpo para constituição do *protossself* encontra-se fundada na contribuição dos mapas interoceptivos (vísceras), mapas gerais do organismo e os mapas dos ‘portais’ sensoriais direcionados para o exterior (visão, audição, tato, paladar, olfato) (DAMÁSIO, 2011, p. 236). Este mapeamento do ponto de vista anatômico requer a participação tanto do tronco cerebral quanto das regiões corticais, de tal modo que a interação e integração estrutural e fisiológica entre os componentes exterosensoriais e intersensoriais se ajustam para compor o *protossself*.

Damásio (2011), além de sustentar a evidência do *self*, sustenta que o *self* é fundamentalmente corporal, ou seja, leva a cabo sua natureza corporificada. Esta concepção foi fomentada no trecho *Uma Prévia das Ideias Principais* da obra *E o cérebro criou o homem*, quando Damásio explicita que “as estruturas cerebrais do *protossself* não são meramente referentes ao corpo. Elas estão ligadas ao corpo em sentido literal e de maneira inextricável”.

A partir desta análise, Damásio (2011) defenderá que a estrutura do *protossself*, além de ter uma relação direta e privilegiada com o corpo, “funda seus pilares arquitetônicos” necessários para construção do *self* central. Tal é a relevância do *protossself* que Damásio (2011) o enunciou através de uma interessante metáfora apresentada em *Uma Prévia das Ideias Principais*: “o melhor modo de conceber o corpo é como a rocha sobre a qual se assenta o *protossself*, enquanto o *protossself* é o eixo em torno do qual gira a mente consciente”.

As inflexões a partir da definição do *protossself* e seu destaque do corpo em seu engendramento, impactarão significativamente e frontalmente em uma das noções de *self* mais influentes da história. Vai impactar o sistema metafísico de Descartes (1596-1650), que, do ponto de vista da história da filosofia, no contexto em que viveu, foi marcado pelos modelos de se pensar as ciências naturais, até certo ponto atrelados a um contexto religioso da época; resultando em uma profusa e entranhada concepção dualista posta de modo radical entre a alma e corpo ou entre a mente e o corpo.

Ao invocarmos Descartes, nós o trazemos pela expressividade do símbolo que se tornou de um conjunto de ideias inerentes ao corpo, à mente e ao cérebro, que, de um modo ou de outro, continuam a influenciar as ciências e as humanidades no mundo ocidental. Trazida para as variantes modernas da noção dualista, Descartes separaria a mente do corpo e a mente do cérebro.

Quando Damásio (2012, p. 218) teceu uma crítica à noção dualista, apresentou um exemplo de sua versão atualizada, de que a “mente e o cérebro estão relacionados, mas apenas no sentido de a mente ser o programa de *software* que corre em uma parte do *hardware* chamado cérebro”, ou em uma concepção mais branda: em que o cérebro e o corpo estão relacionados; porém, apenas no sentido de o cérebro não pode sobreviver sem a manutenção que o corpo lhe oferece (DAMÁSIO, 2012, p. 218).

Como sabemos, Descartes via o ato de pensar como uma atividade separada do corpo; essa afirmação celebrou a separação da mente, a ‘coisa pensante’ (*res cogitans*), do corpo não pensante, o qual tem extensão e partes mecânicas (*res extensa*) (DAMÁSIO, 2012, p. 218). Mais à frente, Damásio culmina na parte que dará o título a sua obra, *O Erro de Descartes: emoção, razão e o cérebro humano*, anunciando o erro de Descartes, cito:

É esse o erro de Descartes: a separação abissal entre o corpo e a mente, entre a substância corporal, infinitamente divisível, com volume, com dimensões e com um funcionamento mecânico, de um lado, e a substância mental, indivisível, sem volume, sem dimensões e intangível, de outro; a sugestão de que o raciocínio, o juízo moral e o sofrimento adviniente da dor física ou agitação emocional poderiam existir independentemente do corpo. Especificamente: a separação das operações mais refinadas da mente, para um lado, e da estrutura e funcionamento do organismo biológico, para o outro (DAMÁSIO, 2012, p. 219).

Agora, trataremos do segundo estágio, o *self* central. O *self* central é resultado das relações estabelecidas entre o organismo, como ele é representado pelo *protossself*, e qualquer parte do cérebro que represente um objeto a ser conhecido. O resultado é o *self* central. A partir do *self* central, um agente conhecedor é introduzido pelo cérebro, trazendo “algo que não estava presente antes, ou seja, um protagonista” (DAMÁSIO, 2011, p. 249).

Sendo assim, quando este protagonista ou testemunha passa a ser disponibilizado em meio a outros conteúdos mentais, ligando-se coerentemente a alguns conteúdos mentais correntes, temos, então, o fenômeno da subjetividade. Damásio (2011, p. 249) defende que existe um limiar do protagonista, “ponto no qual os elementos indispensáveis do conhecimento aglutinam-se, por assim dizer, para compor a subjetividade”.

Ainda que o *protosself* e seus sentimentos primordiais sejam o alicerce do *eu* material, precisamos de um processo do *self* que seja intermediário, entre, de um lado, o *protosself* e seus sentimentos primordiais e, de outro, algo que nos dê identidade e individualidade. Este processo do *self*, um terceiro estágio, Damásio (2011, p. 250) denominará de *self* autobiográfico. O *self* autobiográfico só pode ser construído por meio do mecanismo do *self* central. Em conformidade com Damásio (2011, p. 250), é fundamental compreender que o próprio estado do *protosself* precisa mudar para que ele se torne um *self* propriamente-dito, ou seja, um *self* central.

Para desembocar em um padrão pulsativo do *self* central, ocorre uma mudança crucial do *protosself* promovida pelo encontro do organismo com o objeto, elevando e destacando o perfil mental do *protosself* que precisa conectar-se aos eventos nos quais está envolvido, assegurando-lhe o protagonismo da narrativa momentânea. Isso explica a pulsação do *self* central, configurado em microsequências advenientes da interação que ocorre em uma faixa temporal estreita com o processamento sensorial do objeto. Via de regra, o *protosself* é mudado toda vez que o organismo encontra um objeto, mais ainda, qualquer objeto. Mesmo os objetos evocados ou imaginados provocam alterações no *protosself*, uma vez que ao aprendermos sobre um objeto, além de registrar sua aparência, armazenamos nossas interações em uma memória motora.

Tal fato sugere, por exemplo, porque não desaprendemos a andar de bicicleta. Nossos movimentos das mãos, olhos e cabeça foram ajustados no momento de sua percepção e, na ausência destes objetos, permanecem disponíveis para compor a produção de pulsos do *self* central (DAMÁSIO, 2011, p. 253). Se esta ideia estiver correta, explicaria porque não perdemos a consciência quando devaneamos em um quarto silencioso de olhos cerrados (DAMÁSIO, 2011, p. 253).

Por conseguinte, evocar um objeto engloba a evocação de um conjunto diversificado de interações motoras memorizadas. Isso ocorre porque, para mapear o objeto, o corpo tem que ser ajustado adequadamente pelo cérebro, resultando em ajustes segundo os quais o conteúdo da imagem mapeada é paralelamente sinalizado para o *protosself* (DAMÁSIO, 2011, p. 250-1).

Por meio de uma cadeia de eventos as mudanças do *protosself* dão início à criação momentânea do *self* central. Portanto, o *protosself* necessita estar desperto o suficiente para gerar o sentimento primordial de existência ensejado pelo seu diálogo com corpo. Damásio (2011, p. 251) oferece uma explicação em termos didáticos para esta série de eventos. Para tanto, descreve o processo por meio de três eventos subsequentes:

(i) O primeiro evento é marcado pela transformação no sentimento primordial, resultando em um sentimento no qual o objeto é distinguido dos demais na simultaneidade do tempo. Damásio (2011, p. 251) classificou este evento como “sentimento de conhecer o objeto”;

(ii) O segundo evento é marcado pela geração de destaque para o objeto da interação, um processo que inclui o componente da atenção que atua na convergência de recursos de processamento, ampliando o foco de um objeto em específico em detrimento de outros;

(iii) E, por fim, agora que o objeto se encontra marcado pelo sentimento e destacado pela atenção, a mente atua incluindo as imagens através de uma sequência simples e comum a série de eventos: “o objeto chama atenção do corpo ao ser olhado, tocado ou ouvido”, a partir de uma perspectiva específica, “isso faz o corpo mudar; a presença do objeto é sentida e este objeto ganha destaque (DAMÁSIO, 2011, p. 251).

Damásio (2011, p. 251) vai dizer que a mente, espontaneamente, realiza uma narrativa não verbal desses eventos que não cessam de ocorrer. Assim, o protagonista do *eu* material, latente, se revela por meio de sua associação com ações que estão sendo produzidas pelo organismo a partir do sentimento gerado pelo objeto que engendra uma sensação de posse.

Damásio (2011) marca sua posição em defesa da tese da existência evidente do *self*. A partir de sua defesa, permitirá que o *self* seja situado no âmbito de um conjunto de processos que preenchem a lacuna existente entre a mente e a consciência (*gap*). Com isso, o *self*, para Damásio (2011), constitui uma manifestação importante e culminante desta consciência apresentada em um estrato altamente sofisticado, e no qual os artifícios da memória, linguagem e raciocínio serão vastamente utilizados. Eis o *self* autobiográfico. Neste terceiro estágio, ainda em desenvolvimento no âmbito da história da evolução biológica, as narrativas com contornos simples são enriquecidas, produzindo, assim, um protagonista sólido e bem definido (DAMÁSIO, 2011, p. 252).

Damásio (2011, p. 252) conhece bem o solo teórico e o caminho conceitual o qual percorrerá, tendo a visão de que o mecanismo do *self* central descrito acima funda suas bases no *protosself* e nos sentimentos primordiais, estruturando o maquinismo central que ensinará o *self* autobiográfico e a mente consciente. Os complexos recursos necessários para alcançar o nível do *self* autobiográfico requerem, determinadamente, o funcionamento normal do mecanismo do *self* central (DAMÁSIO, 2011, p. 252). Visão geral de Damásio (2011, p. 259) no que concerne ao *self* autobiográfico e o papel da memória no soerguimento da consciência:

Uma autobiografia é feita de recordações pessoais; é o somatório do que vivenciamos, inclusive as experiências dos planos que fizemos para o futuro, sejam eles específicos ou vagos. O *self* autobiográfico é uma autobiografia que se tornou consciente. Ele se baseia em toda a nossa história memorizada, tanto a recente como a remota. As experiências sociais de que fizemos parte, ou gostaríamos de ter feito, estão incluídas nessa história, assim como as memórias que descrevem as mais refinadas dentre as nossas experiências emocionais, aquelas que podem ser classificadas de espirituais. Enquanto o *self* central pulsa incessantemente, sempre “online”, variando de sinal vagamente pressentido a presença marcante, o *self* autobiográfico leva uma vida dupla. Por um lado, pode ser manifesto, produzindo a mente consciente no que ela tem de mais grandioso e mais humano; por outro, pode estar latente, com sua infinidade de componentes aguardando a vez para entrar em atividade. Essa outra vida do *self* autobiográfico ocorre fora da tela, longe da consciência acessível, e é possivelmente aí que o *self* amadurece, graças à sedimentação gradual e à reelaboração de nossa memória (Damásio, 2011, p. 259).

Conforme as experiências vividas são reencenadas e reconstituídas, por meio da reflexão consciente ou por processos inconscientes, seu conteúdo é inevitavelmente reavaliado e rearranjado, podendo ser modificados a sua composição factual e seu teor emocional em maior ou menor grau, e não obstante os eventos são passíveis de adquirirem novos contornos emocionais. As recordações podem passar por um processo de clivagem, em que algumas recordações são realçadas, outras restauradas, e outras podem ainda ser combinadas por nossas necessidades ou, por fruto do acaso, combinadas de modo a criar cenas nunca realmente ocorridas (DAMÁSIO, 2011, p. 260).

A partir deste esboço, compreende-se o porquê, no decorrer dos anos, de a história de cada um nós, em meio a esta suscetibilidade, comportar sua reedição ou reescrita. Deste ponto de vista, não é exagero afirmar que nunca conseguimos assistir a um mesmo filme ou ouvir uma mesma música duas vezes, pois os fatos compõem uma nova versão e “a música tocada hoje toca diferente em comparação com a do ano passado” (DAMÁSIO, 2011, p. 260). Então, como não nos perdemos em um multiverso de percepções?

Esse problema não é nada trivial. O cérebro em interação e integração com o corpo e com o ambiente dispõe de elaboradíssimos coordenadores quando alcança o nível do *self* autobiográfico por meio de dois processos concatenados. Consoante Damásio (2011, p. 261):

Primeiro, conjuntos expressivos de memórias biográficas definidoras têm de ser agrupados de modo que cada um possa ser prontamente tratado como um objeto individual. Cada um desses objetos pode modificar o *protosself* e produzir seu pulso de *self* central, com os respectivos sentimentos de conhecer e o conseqüente destaque do objeto a reboque. Segundo, como os objetos sem nossa biografia são numerosos, o cérebro precisa de mecanismos capazes de coordenar a evocação de memórias, transmitindo-as ao *protosself* para a interação requerida e mantendo os resultados da interação em um padrão coerente ligado aos objetos causativos. Esse problema não é nada trivial. De fato, os níveis complexos do *self* autobiográfico, aqueles que incluem, por exemplo, importantes aspectos sociais, englobam tantos objetos biográficos que requerem numerosos pulsos de *self* central. Em conseqüência, construir o *self* autobiográfico demanda um maquinário neural capaz de obter múltiplos pulsos de *self* central, em uma breve janela de tempo, para um número expressivo de componentes, e ainda por cima exige que os resultados sejam temporariamente mantidos juntos.

5. Considerações finais

Assumamos a posição relativa à indefinição do *self* ou a posição que o concebe como evidente, ainda assim estaremos atrelados à forte intuição de que nos situamos em torno de um centro a partir do qual investigamos o mundo e sobre o qual nos encontramos. Sabemos que o solo teórico é abarrotado de desafios conceituais.

De um lado, temos as experiências que notamos sem dificuldades e que se apresentam transitórias, e do outro, ainda assim, nos sentimos seguros em apontar que somos proprietários de uma personalidade e de uma identidade. Intuitivamente, nos parece plausível que convivam sobre a perspectiva do *self* a diversidade e a unidade. Uma diversidade expressada no trânsito de nossas experiências, e a unicidade expressada pela dimensão subjetiva do reconhecimento de que nós somos nós mesmos ao longo do tempo. Salvo quando somos acometidos por uma doença ou acidente neurológico que nos inviabilize de tal competência.

Neste ponto, uma questão fundamental trazida para o debate foi a seguinte: há um *self* único, imutável e independente possível de ser localizado? A partir desta indagação, nos defrontamos primeiramente com a noção tradicional de *self*, na qual a substancia-

lidade da mente (ou alma) é concebida como não desintegrável e passível de perdurar mesmo diante da desintegração do corpo. Esta perspectiva de cunho metafísico não-naturalista alcançou seu ápice no sistema filosófico de René Descartes (1596-1650). Em suma, sob este viés, as percepções, sensações, pensamentos encontrariam um local para residir, e, portanto, seriam inerentes a essa sustância material ou substância imaterial fixa e duradoura.

Assim, no percurso desenvolvido trouxemos a forte refutação de David Hume (1711-1776) quando utilizou o princípio da cópia com a finalidade de se opor à ideia de substancialidade da mente. Nos sintonizamos com David Hume no ponto específico a partir do qual, sob o crivo da experiência e da observação, não é possível conceber uma entidade metafísica conhecedora.

Na esteira de Hume, trouxemos para o debate a visão apresentada por Varela, Thompson & Rosch (2003), a qual girou em torno da crítica à insistência na defesa de um *ego-self*. Em busca de um sentido do *self*, chegaram à inferência de que a sua indefinição prevalece. No entanto, revestidos da vontade de trazer algum alento, vão dizer que não é preciso duvidar que o mundo possua sua estabilidade. Tratou-se de um caminho no qual o *self* não serve mais de âncora, em oposição às abordagens que estão literalmente centradas no *self*, e, por consequência, ao deixarem de discutir o *self* pessoal se deslocaram para o problema filosófico que busca compreender qual seria o fundamento último do mundo.

Adiante, trazidos à baila, Damásio (2011; 2012), sob a inspiração de William James, defende a evidência do *self*. Para tanto, dados advindos de fontes da biologia geral e neurociências contribuíram significativamente para alcançarem tal compreensão. O autor marca sua posição em defesa da tese da existência do *self*, situando-o no âmbito de um conjunto de processos que preenchem a lacuna existente entre a mente e a consciência (*gap*). Com isso, o *self* para Damásio constitui uma manifestação importante e culminante desta consciência apresentada em um estrato altamente sofisticado, e no qual os artifícios da memória, linguagem e raciocínio serão vastamente utilizados.

A partir desta análise, Damásio defenderá que as estruturas do *self* estabelecem uma relação direta e privilegiada com o corpo. Emblematicamente, a relevância da noção de *self* corporificado foi expressa por Damásio (2011, p.) na seção *Uma Previa das Ideias Principais*: “o melhor modo de conceber o corpo é como a rocha sobre a qual se assenta o *protossself*, enquanto o *protossself* é o eixo em torno do qual gira a mente consciente”.

Esta ideia guarda uma forte relação com a hipótese de Damásio (2011, p. 225-226) segundo a qual o *self* se compõe de estágios. Os estágios são constituídos em níveis hierárquicos de complexidade e são intimamente interdependentes. Esta tese nos conduz a uma perspectiva naturalista e não-redutivista do *self*, o qual figura como um conjunto de processos. A partir deste entendimento, os níveis se entendem desde os mais elementares e básicos até os dotados de alta complexidade, como a consciência, e, destarte, cada um dos níveis superiores contém instâncias que não são encontradas nos níveis anteriores.

Para Damásio (2011, p. 227), a tentativa de ampliar a compreensão acerca da construção da mente consciente requer um estudo minucioso de seus estágios. A despeito disso, ao priorizar o *self*, Damásio não pretende diminuir ou negligenciar a abrangência da mente em si mesma. Pelo contrário, defende que a mente consciente, ao prevalecer na história da evolução natural, aperfeiçoou os mecanismos de regulação da vida contando com a evolução dos estágios do *self*. A abordagem proposta de acordo com esta perspectiva é a de que não estamos mais tratando apenas de *self*, mas, sim, de *selves*.

Em busca de um sentido do *self*, nos deparamos com noções robustas e antagônicas, as quais queremos destacar. De um lado, sustentando a posição da indefinição do *self*, destacamos os autores de *The embodied mind*, Varela, Thompson & Rosch (2003). De outro, tomando o *self* como evidente, destacamos Damásio (2011, 2012; 2015; 2018). Neste sentido, um ponto importante de convergência entre as duas frentes teóricas se dá em níveis distintos de entendimento. Para ambos, o *self* não é uma substância. No entanto, chegam a esta conclusão percorrendo caminhos distintos. De acordo com Varela, Thompson & Rosch (2003), a sustentação da ausência da fundação do *self* já implica a negação da substancialidade do *self*. Deste modo, a cognição e a estabilidade percebida pelo agente cognitivo, mesmo abandonando a ideia de um *self* fixo, deve ser compreendida no âmbito da fenomenologia.

Nesta perspectiva, o processo cognitivo é concebido como uma construção dinâmica do mundo, inseparável do histórico de vida e singular do ponto de vista da experiência de cada sujeito. Não à toa, Varela, Thompson & Rosch (2003) se inspiraram em Merleau-Ponty (1999[1945]). Merleau-Ponty, em sua obra *A Fenomenologia da Percepção* (1999[1945]), traz o interessante subtítulo: “o verdadeiro cogito não substitui o próprio mundo pela significação do mundo”. A ideia central é a de que o mundo é inseparável do sujeito, e, por conseguinte, o sujeito é projeto do mundo que ele mesmo projeta (ROMÃO, 2021, p. 97).

De posse da convicção de que o problema não é unívoco, nos acautelamos quanto à possibilidade de esgotar as investigações concernentes ao *self*. Diante da indefinição ou não do *self*, nos inclinamos a assumir algumas posições, sem, no entanto, pretender esgotar o debate: i) o *self* não é uma substância; ii) vislumbramos a força do argumento o qual propõe que o *self* seja designado como um conjunto de processos a partir dos quais a subjetividade emerge, e, assim, iii) fornece ao organismo um sentido de posse dos conteúdos mentais, contribuindo, portanto, para o desenvolvimento da personalidade e identidade; iv) concebemos o *self* como corporificado; e, por fim, em relação ao ponto mais polêmico que diz respeito à evidência do *self*, e sem a intenção de exaurir o debate neste ponto, v) assumimos a possibilidade de sua evidência desde que seja tratado como um fenômeno ou conjunto de processos.

Referências

ALMADA, Leonardo Ferreira & MESQUITA, Luiz Otávio de Sousa. Senciência corporal e autoconsciência. **Controvérsia**, v.13, n.3, p. 26-41, 2017.

DAMÁSIO, António. **E o cérebro criou o homem**: Construindo a mente consciente. São Paulo, SP: Companhia das Letras, 2011.

DAMÁSIO, António. **O erro de Descartes**: Cérebro, razão e emoção. São Paulo, SP: Companhia das Letras, 2012.

DAMÁSIO, António. **O mistério da consciência**: do corpo e das emoções ao conhecimento de si. São Paulo, SP: Companhia das Letras, 2015.

DAMÁSIO, António. **A estranha ordem das coisas**: As origens biológicas dos sentimentos e da cultura. São Paulo, SP: Companhia das Letras, 2018.

DENNETT, Daniel. **Conscious explained**. Boston, MA: Little Brown, 1991.

HUME, David. **Tratado da natureza humana**: uma tentativa de introduzir o método experimental de raciocínio nos assuntos morais. Tradução de Deborah Danowski. 2 ed. rev. e ampliada. São Paulo, SP: Editora UNESP, 2009.

JAMES, William. **The Principles of Psychology**. New York, NY: Dover, 1950 [1890].

MERLEAU-PONTY, Maurice. **Fenomenologia da percepção**. Trad. Carlos Alberto Ribeiro de Moura. 2. ed. São Paulo, SP: Martins Fontes, 1999.

ROMÃO, Fabiense Pereira. **O Papel do corpo na emergência da mente (consciente): o problema da percepção-ação na exploração do ambiente a partir das perspectivas ecológicas e teorias enativistas.** 2021. 130 p. Dissertação (Filosofia) – Universidade Federal de Uberlândia, Uberlândia, MG, 2021.

VARELA, Francisco; THOMPSON, Evan; ROSCH, Eleanor. **A mente incorporada: Ciências cognitivas e experiência humana.** Porto Alegre, RS: Artmed, 2003.





CORES E MUNDO: UMA DISCUSSÃO SOBRE A HIPÓTESE DO ESPECTRO INVERTIDO



Juliana de Orione Arraes Fagundes¹

Resumo:

O argumento do espectro de cores invertido é usado em filosofia da mente para esclarecer a noção de *qualia*, aspectos subjetivos da experiência. Este trabalho apresenta e discute uma das versões do argumento. Embora seja bastante intuitivo, o argumento começa a sofrer desgastes quando seus detalhes são observados. Isso porque os *qualia* provêm de uma concepção cindida de sujeito e mundo. Experienciamos aspectos do mundo que se tornam relevantes para nós graças às nossas capacidades cognitivas, muitas delas desenvolvidas pelo processo evolutivo. Assim, por meio da noção de *affordances*, podemos pensar em nossas experiências a partir das nossas relações com as coisas, e não por um aparente aspecto estritamente subjetivo.

Palavras-chave:

Qualia; *affordances*; Dennett.

Abstract:

The inverted spectrum hypothesis is used in philosophy of mind to clarify the notion of *qualia*, subjective aspects of experience. This work presents and discusses a particular version of this argument. Although it is quite intuitive, the argument begins to lose its force as its details are analyzed. This is because *qualia* comes from a conception that imposes a gap between subjects and world. We experience aspects of the world that become relevant to us thanks to our cognitive abilities, many of which have been developed through the evolutionary process. Thus, by means of the notion of *affordances*, we can think of our experiences from our relationships with things, and not from an aspect that seems to be strictly subjective.

Keywords:

Qualia; *affordances*; Dennett.

¹ Professora adjunta do Departamento de Filosofia e Ciências Humanas da Universidade Estadual de Santa Cruz (UESC). joafagundes@uesc.br

Introdução ao argumento

Um argumento comumente usado para tentar esclarecer a noção de *qualia* é o de que se a sensação subjetiva de um sujeito perante um objeto azul, por exemplo, fosse equivalente à de outro sujeito perante um objeto amarelo, essa diferença subjetiva não poderia ser detectada, pois o vocabulário acerca das cores de ambos os sujeitos também seria invertido. A ideia é que a experiência da cor se apresenta apenas para o sujeito e não pode ser alcançada por uma perspectiva objetiva. A partir disso, algumas características dos *qualia* podem ser enumeradas: privados, inefáveis, intrínsecos (independentes de outros tipos de estados mentais, como crenças), subjetivos.

Trata-se do famoso argumento do espectro invertido, apresentado primeiramente por Locke (1959)² e reescrito em diversas versões ao longo do tempo. Para Locke, se a ideia produzida pelo azul de uma violeta na mente de uma pessoa fosse como a ideia produzida pelo amarelo de um girassol na mente de outra, jamais poderíamos saber desse fato. Pretende-se aqui apresentar o argumento e, a partir dele, tentar olhar para as características de nossas experiências sob outra perspectiva, não mais estritamente subjetiva³.

Locke defende a existência de uma distinção entre qualidades primárias e secundárias. De acordo com essa distinção, as qualidades primárias estariam nos objetos independentemente das nossas percepções, como dimensão e forma. As qualidades secundárias, por sua vez, estariam nos objetos da forma como são percebidas por nossos sentidos, como cor e cheiro. A pergunta a ser colocada é se as qualidades secundárias estão objetivamente no mundo ou se só estão nos objetos na medida em que são percebidas por nós.

Assim, as cores resultam de uma relação entre as propriedades dos objetos e as nossas propriedades visuais de percebê-las. Nesse sentido, aparentemente nós não podemos dizer que elas estejam nos objetos, ao menos não da mesma forma como podemos dizer que a forma e a dimensão estão. Se a cor depende das propriedades visuais de quem percebe, então onde ela está? Ela está na subjetividade?⁴ Essas questões levam algumas pessoas a dizerem que há algo inacessível sob qualquer perspectiva objetiva, denominando-o de *qualia*.

2 Livro II, capítulo XXXII, § 15.

3 Esse argumento foi discutido em minha dissertação de mestrado (FAGUNDES, 2009).

4 O famoso argumento do conhecimento também apela para a percepção das cores. Ver: JACKSON, 1986. Em Fagundes (2018), discuto esse argumento.

À época de Locke, o conhecimento que se tinha sobre o espectro de cores era menor. Em certa medida, era possível pensar que havia uma nuance de cor oposta para cada tonalidade. Porém, sabemos atualmente de algumas informações que nos levam a concluir que seria muito improvável uma percepção invertida do espectro sem expressão comportamental (BYRNE, 2020). Um dos problemas é que não há uma correspondência entre as cores e suas inversas, pois algumas cores têm muito mais nuances perceptíveis do que suas inversas, o que seria facilmente detectado em testes comportamentais se houvesse tal inversão. Outro ponto é que as cores possuem tom, luz e saturação, de modo que, em uma inversão de *qualia*, o sujeito relataria que o azul é mais luminoso que o amarelo, por exemplo, o que seria estranho perante as pessoas que tivessem a percepção de cor normal. Um terceiro ponto são as reações emocionais que temos perante as diferentes cores. Se fosse houvesse uma inversão dos *qualia* das cores, isso causaria reações fisiológicas e emocionais estranhas, como a aceleração nos batimentos cardíacos perante um objeto azul ou a sensação de calma perante um objeto vermelho.

No entanto, alguns filósofos podem responder que essas dificuldades empíricas não invalidam o argumento, pois estamos diante de questões conceituais, não de empíricas. Em outras palavras, o argumento poderia ser tomado hipoteticamente como um exemplo metafórico para nos auxiliar a compreender os *qualia* e suas características. Assim como outros argumentos na história da filosofia se baseiam em hipóteses que jamais teriam realização empírica, como a Alegoria da Caverna, o argumento do espectro invertido poderia ser tomado com fins didáticos de ajudar a esclarecer o conceito complexo de *qualia*. Vale notar que a grande complexidade em esclarecer o conceito está relacionada à sua subjetividade e privacidade e o recurso que temos para lidar com esclarecimento de conceitos, a linguagem, é objetiva e pública. Passamos agora a uma discussão do argumento a partir de Dennett (1991) e Block (2005).

O argumento da inversão intrapessoal

Para Block (2005), podemos trabalhar filosoficamente com a noção de espectro invertido. Para isso, ele apresenta versões do argumento que considera conceitualmente possíveis e interessantes: o da Inversão Intrapessoal e o da Terra Invertida. Por questões de espaço, discutiremos aqui apenas a Inversão Intrapessoal⁵, que segue as etapas a seguir:

5 O argumento da Terra Invertida nos exigiria entrar em uma discussão técnica sobre os princípios de Chal-

- (a) O indivíduo tem visão normal;
- (b) Lentes de contato são implantadas em sua retina com tecnologia para inverter as cores. O que era visto como azul passa a ser visto como amarelo, o que era vermelho passa a ser visto como verde, e assim por diante;
- (c) O indivíduo se adapta à inversão, passando a adotar nomes invertidos. Passa a chamar de verde o que antes chamava de vermelho, por exemplo;
- (d) Ele sofre uma amnésia e se esquece de como via o mundo antes da inversão do espectro, de modo que a inversão passa a não ser mais identificada por ele e ele se comporta exatamente como alguém que não teve os *qualia* invertidos.

Nesse caso, para evitar o argumento acima apresentado de que nosso espectro de cores não é simétrico, podemos imaginar seres parecidos conosco, inteligentes e capazes de se comunicar por linguagem complexa e articulada, mas que tenham um sistema visual com espectro simétrico, quer dizer, que tenham um sistema sensorial tal que a cada cor corresponda uma cor complementar (BLOCK, 1990). Após as quatro etapas supracitadas, tal indivíduo teria sofrido uma inversão de *qualia* que não seria detectável nos testes comportamentais.

Esse processo depende de que a adaptação prevista na etapa (c) ocorra (DENNETT, 1991). Haverá um esforço inicial do indivíduo para inverter os nomes das cores que se apresentam diante dele. Entretanto, com o passar do tempo, essa troca deverá ser automatizada, de modo que o indivíduo possa fazer sem precisar pensar sobre ela. É preciso que essa inversão do vocabulário fique completamente natural para o indivíduo para que ela chegue ao estágio de não ser detectada em nenhum teste comportamental.

Dennett (1991, p. 389-398) argumenta que, para a inversão comportamental de fato não poder ser detectada por meio do comportamento verbal, ela tem que chegar ao ponto de não ser perceptível nem mesmo para o próprio sujeito. Todavia, segundo ele, se ela não é perceptível para o sujeito, os *qualia* provavelmente voltaram ao normal. Se não

mers (1996) acerca da organização funcional. Ele defende que há a possibilidade lógica da inversão, mas, em nosso mundo, seres com a mesma organização funcional necessariamente terão os mesmos *qualia*. Esse é um debate no qual não entraremos aqui.

há nenhuma diferença detectável por meio de testes objetivos entre (a) e (c), o que poderia servir como meio de comparação para que se afirme que houve inversão de *qualia*? Em outras palavras, se nem o sujeito consegue perceber, do que estamos falando, afinal? Parece haver algo estranho com esse argumento, pois a inversão também não poderia ser detectada sob um ponto de vista subjetivo quando o processo fosse concluído. Antes do fim do processo, por outro lado, o indivíduo poderia falar sobre a inversão, de modo que ela seria observável objetivamente.

Aqui, a proposta de Block (2005) era que o argumento fosse avaliado como uma hipótese teórica, independentemente dos aspectos empíricos, mas mesmo assim surgem problemas. Vale a pena olhar um pouco mais para esse argumento a partir de aspectos empíricos, pois esse ser imaginário que teve o espectro invertido por meio de lentes e, cujo sistema de percepção de cores é simétrico, possui linguagem como nós. Se ele usa as palavras da mesma forma que nós as usamos, então há associações entre as cores e outros aspectos de sua vida cognitiva. Se ele, assim como nós, tem as sensações vinculadas à afetividade, poderia também perceber o vermelho como uma cor quente e o azul como uma cor fria (HUMPHREY, 1994). Após a inversão, o sujeito diria que a brasa agora lhe parece fria e que o céu lhe parece quente. O processo de amnésia, portanto, teria que atingir também a relação afetiva que o sujeito tem com as cores. Caso contrário, a inversão também seria detectável.

Esse argumento talvez não consiga alcançar os aspectos conceituais da hipótese do espectro invertido. Apesar disso, ele nos leva a perceber que o argumento do espectro invertido precisa de mais idealizações do que parecia inicialmente. Por exemplo, o argumento do espectro invertido pressupõe que as cores, bem como os *qualia* são características autônomas do nosso sistema sensorial. Mas o processo de percepção das cores envolve características diversas, tais como brilho, calor, questões afetivas e emocionais. Como as cores estão relacionadas a uma série de processos cognitivos, elas afetam nossos comportamentos de diversas formas e se relacionam a diversos aspectos de nossa cognição. Por mais que o argumento do espectro invertido pareça intuitivo, ele ignora a existência das diversas variáveis envolvidas na percepção das cores.

Dennett (1991, p. 395) pede para supormos que a amnésia do sujeito seja provocada por meio de uma cirurgia. Ocorre que esse sujeito tinha sofrido um acidente de trânsito. Ele tinha batido em um carro azul e esse acidente lhe havia deixado como seqüela um trauma psicológico que lhe causava mal-estar sempre que via um objeto azul. Durante a cirurgia, a região responsável pelo trauma não seria tocada. O que aconteceria após a

cirurgia é que o mal-estar do sujeito passaria a ocorrer quando ele visse um objeto amarelo, de modo que a inversão dos seus *qualia* seria detectada comportamentalmente com facilidade.

A suposição de que o sujeito do experimento seria parecido conosco, com a diferença apenas de ter um espectro de cores simétrico, não é um ajuste suficiente para que o experimento funcione. Seria necessário supor, ainda, uma diversidade de diferenças entre nossos sistemas cognitivos e o do sujeito do experimento. Por exemplo, seria necessário supormos uma completa autonomia dos *qualia* dentro de nossos processos cognitivos para que eles não tivessem nenhum tipo de componente afetivo e não tivessem ligações com nossas memórias nem com nosso inconsciente. E se os *qualia* são assim em uma criatura que supomos tão diferentes de nós, o argumento talvez não tenha muito a esclarecer acerca do papel dos *qualia* em nós.

Por que o mundo é colorido?

As cores, assim como nossas outras sensações, nos fornecem uma série de informações importantes sobre o mundo ao nosso redor, permitindo-nos continuar vivos. Nossos sistemas sensoriais possuem ajustes em relação ao mundo que nos cerca, uma vez que eles são adaptados para que possamos perceber os aspectos da realidade relevantes para nós. As sensações nos causam emoções diversas. As cores podem causar atração ou repugnância em relação a um objeto e isso vai variar de pessoa para pessoa. Além do mais, as pessoas enxergam cores diferentes no mundo, umas enxergam mais matizes do que outras. Isso em geral tem a ver com o ambiente em que vivem. Por exemplo, uma pessoa que vive em um ambiente com muita neve consegue diferenciar tonalidades de branco que um habitante dos trópicos não é capaz. Da mesma forma, o idioma de quem vive nesse ambiente gelado contará com um vocabulário maior sobre tonalidades de branco, de modo que o ambiente cultural humano também influencia nas tonalidades que as pessoas podem enxergar, pois vemos o mundo, em grande medida, a partir da linguagem.

Ainda que haja uma variedade de itens influenciando de diversas formas as nossas percepções das cores, entre afetos, traumas, questões biológicas e culturais, a maior parte de nós já nascemos programados para percebermos e discriminarmos determinadas cores. Quando alguém tem algum tipo de dificuldade em relação a isso, como um daltônico, os testes comportamentais podem identificar com muita facilidade. Se uma pessoa tivesse seus *qualia* de cores invertido, isso não lhe deixaria em uma situação de desvantagem

adaptativa? A brasa não lhe pareceria fria? Embora pareça bastante intuitivo à primeira vista, o argumento do espectro invertido tem o problema de desconsiderar a importância das informações que as cores nos fornecem acerca do mundo.

Ao argumentar que não temos acesso às experiências coloridas dos outros, já que a inversão nas cores seria acompanhada de uma inversão vocabular, perde-se um ponto importante. Certos aspectos do mundo são coloridos para nós porque são importantes para nossa sobrevivência. Nessa perspectiva, as cores podem ser melhor compreendidas não como *qualia* que moram dentro de nossas mentes, mas, sim, como aspectos do mundo objetivo que podemos perceber graças à nossa estrutura sensorial. Há coisas que vemos e outras que não vemos, devido às nossas possibilidades cognitivas. Fomos dotados de capacidades para experimentar o mundo objetivo, mas essas capacidades são estruturas orgânicas, como redes de neurônios, por exemplo. Não há cores nessas redes de neurônios, nem há *qualia*. As cores estão naquilo que percebemos como colorido.

Affordances e cores

Dennett (2017) tende a abordar a consciência a partir de seus aspectos evolutivos e a negar a existência de *qualia*. De acordo com o autor, os seres vivos possuem sensores biológicos que capturam aspectos do ambiente. Tais aspectos são as *affordances* daquele organismo, ou seja, aquilo que no ambiente se manifesta para o organismo. O estudioso traz uma noção de ontologia que deriva da epistemologia. A ontologia, segundo essa concepção, constitui-se por aspectos do mundo que estão disponíveis para cada organismo, incluindo aquilo que foi construído pelo próprio organismo (como instituições, no caso humano).

Férias não estão na ontologia de um urso polar, mas neve está, e também estão as focas. Neve provavelmente não está na ontologia de um peixe-boi, mas hélices de motores de popa podem bem estar, junto com algas e peixes e outros peixes-boi (DENNETT, 2017, p. 60, tradução nossa).

O termo *affordances* é um neologismo que deriva de *afford*, um verbo da língua inglesa que significa proporcionar, suprir ou dispor. Falta-nos, ainda, uma tradução precisa do termo para a língua portuguesa. Tem a ver com aquilo que o ambiente oferece ou provê ao animal, coisas que podem ser boas ou ruins para aquele animal. Mas o que o ambiente

oferece a um animal são aspectos do mundo relevantes para aquele animal em algum sentido, conforme as possibilidades de relações entre aquele animal e o mundo circundante. Por exemplo, coisas que podem servir de alimento, coisas que podem aumentar ou diminuir seu sucesso reprodutivo ou coisas que ameaçam a vida e o bem-estar.

As *affordances* são partes do mundo objetivo com as quais o organismo lida conforme suas possibilidades cognitivas, construídas pelas relações que seus ancestrais tiveram com o meio. Em uma linguagem mais dennettiana, podemos dizer que as *affordances* são resultado de uma acumulação de projetos sem projetista, construídas ao longo da história evolutiva. Nesse sentido, se há uma experiência interna (um “como é ser”)⁶ de ser um organismo, tal experiência traz informações reais sobre o mundo, não estando confinada ao sujeito. Ao contrário, é produto das relações de seus ancestrais evolutivos com o meio.

A noção de *qualia* traz certa ruptura entre uma realidade interiormente percebida e o mundo exterior. Essas realidades podem ser reunidas por meio da noção de *affordance*. Nós experimentamos aspectos estruturais do mundo por meio de nossas capacidades cognitivas e pelas propriedades que o mundo tem de ser percebido por nós de determinadas maneiras. Essas propriedades são relevantes para nós porque foram essenciais para a sobrevivência e reprodução de nossos ancestrais.

Hardcastle (2020) explica que a noção de *affordance* se apresenta com o intuito de promover uma indistinção entre o mundo objetivo e a realidade subjetiva. Com isso, esperava-se que a noção de *qualia* perderia sua razão de ser. A consciência, de acordo com esse tipo de concepção, seria resultado de um sistema relacional cérebro-corpo-ambiente. As *affordances* seriam, portanto, o resultado do que há no ambiente e as capacidades das criaturas. Contudo, elas não dependem do organismo, no sentido de que estariam objetivamente no mundo. Ao contrário, elas são oportunidades ambientais para que o organismo possa se apropriar delas.

Não obstante, a autora acredita que a subjetividade aparece em contextos específicos e que a noção de *affordances* não é capaz de diluir o problema dos *qualia*. No seu dizer, a experiência que nós temos dos objetos é diferente em relação à experiência que observamos as outras pessoas tendo. A diferença está em como o sujeito experiencia as *affordances*. Nesse caso específico, essa noção ajuda na compreensão da noção de consciência.

6 Ver Nagel (2005).

Porém, o problema apontado por ela é que precisamos de critérios capazes de identificar por que a consciência aparece em certos contextos, e não em outros.

Em relação à pergunta apresentada na introdução deste trabalho, sobre se as qualidades secundárias estariam no mundo ou na mente, Dennett (2020) insiste na ideia de que não há *qualia* subjetivos em nossas mentes. De acordo com o filósofo, nós temos crenças sobre *qualia*, mas elas seriam uma espécie de ilusão. O que causa nossas crenças acerca das cores são as *affordances*, e elas estão disponíveis no ambiente, sendo, portanto, aspectos objetivos da realidade. Então, levando esse debate à questão das qualidades secundárias, elas estão no mundo objetivo e nós as capturamos conforme nossas possibilidades cognitivas. Ele entende que a discussão sobre *qualia* atrapalha os progressos sobre o tema da consciência e que a noção de *affordance*, por outro lado, ajuda.

De fato, considerando o extenso debate sobre a noção de *qualia*, não podemos compreender *qualia* como *affordances* por um motivo muito simples: por definição, *qualia* são subjetivos e *affordances* são objetivas. No entanto, essa discussão sobre o argumento do espectro invertido serve para uma reflexão sobre algumas dificuldades que surgem quando tentamos compreender separadamente as realidades objetiva e subjetiva. Os aspectos intrínsecos, privados e inefáveis dos *qualia* parecem ser intuitivos, mas não resistem facilmente a uma investigação que leva em conta nossa constituição enquanto seres vivos. A noção de *qualia* exige uma investigação estritamente subjetiva, como se fôssemos *res cogitans*, mas não somos isso. Enquanto os *qualia* separam as investigações sobre as realidades objetiva e subjetiva e criam um *hard problem* (CHALMERS, 1996), a noção de *affordance* faz o contrário: ela integra sujeito e mundo.

Referências

BYRNE, A. Inverted qualia. In: **Stanford Encyclopedia of Philosophy**. In: ZALTA, E. (Ed.) Fall, 2020 Edition. Disponível em: <https://plato.stanford.edu/archives/fall2020/entries/qualia-inverted/>. Acesso em: 10 ago. 2023.

BLOCK, N. Inverted Earth. **Philosophical perspectives: action theory and philosophy of mind**, v. 4, p. 53-74, 1990. Disponível em: <http://www.jstor.org/stable/2214187>. Acesso em: 05 mar. 08.

BLOCK, N. Qualia. In: GUTTENPLAN, S. **A Companion to the philosophy of mind**. Oxford: Blackwell, 2005.

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

DENNETT, D. **Conscious explained**. London: Penguin Books, 1991.

DENNETT, D. **From bacteria to bach and back: The evolution of minds**. New York: W. W. Norton & Company, 2017.

DENNETT, D. A History of Qualia. **Topoi**, v. 39, p. 5-12, 2020. Disponível em: <https://doi.org/10.1007/s11245-017-9508-2>. Acesso em 31 jul. 2023.

LOCKE, J. **An essay concerning human understanding**. New York: Dover Publications, 1959.

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

FAGUNDES, J. O. A. **A consciência vista de fora: A perspectiva de Dennett**. (Dissertação de Mestrado) Programa de Pós-Graduação em Filosofia da Universidade de Brasília. Orientador: Paulo Abrantes. Brasília, 2009.

FAGUNDES, J. O. A. Sobre estados mentais relacionais e qualia: o totalmente subjetivo não pode ser conhecido. In: TOLEDO, G. L.; GOUVEA, R. A. dos S. & ALVES, M. A. S. (Orgs.). **Debates contemporâneos em filosofia da mente**. 1ed. São Paulo: FiloCzar, 2018. p. 181-196.

HARDCASTLE, V. G. The consciousness of embodied cognition, affordances, and the brain. **Topoi**. v. 39, p. 23–33, 2020. Disponível em: <https://doi.org/10.1007/s11245-017-9503-7>. Acesso em 31. jul. 2023.

HUMPHREY, N. **Uma história da mente**. Rio de Janeiro: Campus, 1994.

JACKSON, F. What Mary didn't know. **Journal of Philosophy**. v. 83, p. 291-5, 1986.





MENTE, MUNDO E CORPO: RESPONSABILIDADE E AGENCIAMENTO ESTENDIDO



Murilo Karasinski¹ & Léo Peruzzo Júnior²

Resumo:

Neste artigo pretendemos mostrar como a tese da mente estendida repercute sobre a compreensão das ideias de agenciamento e responsabilidade. Assim, ao considerar que processos cognitivos se estendem além do cérebro e incorporam elementos externos, como ferramentas e tecnologias, discutimos como a responsabilidade estendida requer uma reavaliação dos critérios tradicionais de atribuição de responsabilidade, levando agora em conta também a influência do ambiente na tomada de decisões e nas ações intencionais individuais. Para isso, reconstruímos a tese da mente estendida, para, na sequência, propor uma leitura que considere as implicações dessa nova abordagem às ciências cognitivas.

Palavras-chave:

Mente estendida; Agenciamento; Responsabilidade Estendida; Cognição; Intencionalidade.

Abstract:

In this article, we intend to show how the extended mind thesis impacts the understanding of the ideas of agency and responsibility. Thus, by considering that cognitive processes extended beyond brain and incorporate external elements such as tools and technologies, we discuss how extended responsibility requires a reassessment of traditional criteria for assigning responsibility, now also taking into account the influence of the environment on decision-making and individual intentional actions. To do this, we reconstruct the extended mind thesis and then propose an interpretation that considers the implications of this new approach for cognitive sciences.

Keywords:

Extended Mind; Agency; Extended Responsibility; Cognition; Intentionality.

1 Professor-doutor da Pontifícia Universidade Católica do Paraná – PUCPR. E-mail: k.murilo@pucpr.br
ORCID: <https://orcid.org/0000-0002-6099-6968>

2 Professor-doutor da Pontifícia Universidade Católica do Paraná - PUCPR / FAE. E-mail: leo.junior@pucpr.br
ORCID: <https://orcid.org/0000-0003-3084-5170>

Introdução

Se a tese for aceita, até onde devemos ir? Todos os tipos de quebra-cabeças vêm à mente. E os aldeões amnésicos em Cem Anos de Solidão, que esquecem os nomes de tudo e então penduram etiquetas em todos os lugares? As informações do meu Filofaz contam como parte da minha memória?

Clark & Chalmers (*The Extended Mind*, p.16)

A proposta da mente estendida (*extended mind*), formulada inicialmente por Clark e Chalmers (1998), pretende responder aos desafios da perspectiva dualista, uma vez que indica a conclusão aparentemente radical de que somos uma complexa mistura de nosso corpo, do meio ambiente e de todas as interações empreendidas. Por isso, é uma tese construída sobre a hipótese de que nossos processos cognitivos são *constituídos* por processos tanto derivados do próprio corpo quanto derivados do ambiente e de suas interações.

Entretanto, se assumirmos o tradicional argumento de que a cognição é constituída por ações intencionais, o ambiente desempenha, na teoria da mente estendida, um papel apenas nos casos em que aquilo que é chamado de ‘cognitivo’ pode ser realizado dentro do crânio. Isso significa que as ações intencionais internas podem realizar ações funcionalmente equivalentes fora da cabeça, o que implica que a fronteira entre o que consideramos como processos cognitivos puramente internos e o ambiente externo se torna menos clara. Nesse contexto, o objetivo deste trabalho é explorar as implicações dessa perspectiva para a compreensão da responsabilidade e do agenciamento, uma vez que esses termos estão apoiados sobre a capacidade de tomar as ações como intencionais e conscientes. Por isso, se aceitarmos a premissa de que as ações intencionais podem ter equivalentes funcionais fora da cabeça, isso afeta a nossa compreensão da responsabilidade individual e do próprio agenciamento.

Segundo Clark e Chalmers (1998, p. 7), então, a questão de onde a mente termina e o resto do mundo começa tem recebido duas respostas padrão:

Alguns aceitam os limites da pele e do crânio, e dizem que o que está fora do corpo está fora da mente. Outros são impressionados com argumentos que sugerem que o significado de nossas palavras ‘apenas não está na cabeça’ e sustentam que esse externalismo sobre o significado carrega um externalismo sobre a mente.

É justamente a partir deste confronto que a teoria da mente estendida pretende defender uma terceira posição, particularmente assentada sobre um externalismo baseado no papel ativo do ambiente na condução dos processos cognitivos.

A visão que defendemos aqui é referida por um crescente corpo de pesquisa em Ciência Cognitiva. Em áreas tão diversas quanto a teoria da cognição situada (SUCHMAN, 19870), estudos de robótica do mundo real (BEER, 1989), dinâmicas de abordagens para o desenvolvimento infantil (THELEN e SMITH, 1994), e pesquisas sobre as propriedades cognitivas de grupos de agentes (HUTCHINS, 1995), a cognição é muitas vezes considerada contínua com os processos no ambiente. (...) Alguns acham esse tipo de externalismo intragável. Uma razão pode ser que muitos identificam o cognitivo com o consciente, e parece longe de ser plausível que a consciência se estenda para fora da cabeça nesses casos (CLARK & CHALMERS, 1998, p. 10).

Clark e Chalmers (1998) sustentam que, ao contrário do externalismo passivo de Putnam (1975) e Burge (1998), o organismo humano está ligado ao ambiente numa direção bidirecional, a qual é responsável por criar um sistema acoplado que pode ver visto como um sistema cognitivo por direito próprio. Assim, o papel causal não é desempenhado apenas por um lado do sistema, uma vez que removendo o componente externo, por exemplo, é possível que a competência comportamental perca a sua função. Segundo eles, embora Putnam (1975) e Burge (1998) estejam certos sobre a externalidade do conteúdo, não está claro que tais aspectos externos desempenhem um papel causal ou explicativo na geração da ação (CLARK; CHALMERS, 1998, p. 9). Assim, enquanto para Putnam (1975, p. 151), por exemplo, “o significado determina a extensão, mas pelo fato de a extensão fazer parte do significado”, pois, por um lado, não é o falante por si só que fixa a extensão e, por outro, porque não é o seu estado mental que determina a referência, para Burge (1998, p. 59), o conteúdo de determinadas crenças e estados mentais são “o que uma pessoa crê que significam as suas palavras, como as toma, determina completamente que atitudes pode expressar ao usá-las”, uma vez que o conteúdo de tais eventos ou estados mentais está estritamente limitado a noções que ele compreende.

Seguindo a tipografia do externalismo, a tese da mente estendida procura indicar que objetos externos, assim como a atividade neural, constituem a base dos processos cognitivos e extravasam os limites intracranianos. Isso quer dizer, em outras palavras, que se

uma parte do mundo funciona como um processo, o qual fosse ele executado na cabeça, nós não hesitaríamos em reconhecê-lo como parte do processo cognitivo, então, essa parte do mundo é (como alegamos) parte do processo cognitivo (CLARK & CHALMERS, 1998, p. 8).

Assim, apoiando-se no que foi chamado de *princípio da paridade*, Clark e Chalmers (1998) compreendem que quando um processo cognitivo é substituído por outro e cumpre a mesma função, então, este processo faz parte tanto quando aquele inicial. É justamente a partir do princípio da paridade que os processos cognitivos podem ser *estendidos* para fora do corpo (CARVALHO, 2020, p. 146).

O uso do termo *estendido*, assim como observa Farkas (2011, p. 441), é metafórico, isto é, significa que “estendemos o que conta como um papel funcional ou um perfil disposicional que qualifica um estado como um certo tipo de estado permanente”. Por isso, aqui, falar de “extensão” não é o mesmo que falar de localização espacial, mas do papel que o estado estensional desempenha na vida mental do sujeito. A proposta da mente estendida, por conseguinte, propõe que a base física dos eventos mentais pode se estender além dos limites do nosso corpo orgânico. Considerar a extensão da base física para fora do crânio, por sua vez, exige repensar, conseqüentemente, a localização dos fatores constitutivos dos processos cognitivos, assim como o que significa a ideia de responsabilidade e agenciamento nestes novos termos.

Este artigo pretende mostrar, portanto, que a tese da mente estendida, levada às últimas conseqüências, implica também uma extensão paralela da responsabilidade com os artefatos compartilhados e, ainda, com às agências estendidas como um todo, e não apenas às suas partes humanas. Obviamente, tem-se como resultado, primeiramente, uma nova leitura do individualismo metodológico, para o qual toda ação social deveria ser explicada em termos de motivos e disposições e, posteriormente, o argumento de que uma ação é a combinação de elementos humanos e não-humanos, estes últimos agora acoplados. É justamente este último argumento que indica a necessidade de reavaliar os limites tradicionais da responsabilidade e atribuí-la não apenas aos agentes humanos, mas também aos artefatos e agências estendidas que estão envolvidos na ação.

1. Artefatos não-biológicos e acoplamento: novos andaimes cognitivos

Os exemplos de Otto e Inga, apresentados por Clark e Chalmers (1998, p. 12), são utilizados para ilustrar a ideia de que o uso de artefatos externos pode ser integrada de tal forma que se tornem parte essencial do processo cognitivo. No exemplo de Otto, um indivíduo que sofre de uma deficiência de memória (doença de Alzheimer) utiliza um bloco de notas para compensar sua limitação. Ele anota informações importantes no bloco de notas, como compromissos e endereços, e consulta essas anotações sempre que necessário. Nesse caso, o bloco de notas funciona como uma extensão da memória de Otto, permitindo que ele acesse informações de forma rápida e eficiente. Isso não seria diferente de um indivíduo que armazena uma série de notas em seu celular ou que opta por um sistema *iCloud* para armazenar dados, como documentos e fotos, os quais podem ser baixados em outro momento.

Já o exemplo de Inga envolve uma situação semelhante, mas com uma diferença crucial. Inga não possui deficiência de memória, mas opta por utilizar suas crenças para caminhar até o Museu de Arte Moderna, que está localizado na 53rd Street. Sua crença estava em algum lugar na memória, esperando ser acessada. Segundo Clark e Chalmers (1998, p. 14), seria inutilmente complexo explicar as ações de Inga em termos de crenças sobre sua memória. Por isso, “o bloco é uma constante para Otto, da mesma forma que a memória é uma constante para Inga”. Nos dois casos, os artefatos (o bloco de notas de Otto e as crenças de Inga) desempenham um papel fundamental na cognição, pois eles são utilizados como recursos confiáveis para auxiliar nas tarefas cognitivas, assim como para acessar informações importantes e orientar ações específicas.

Clark e Chalmers (1998) argumentam, então, que explicar as ações de Inga em termos de crenças seria excessivamente complexo. Em vez disso, eles sugerem que a mente estendida pode ser compreendida por meio da interação entre os elementos cognitivos internos (como as crenças de Inga) e os artefatos externos (como o bloco de notas de Otto). Os autores ainda procuram atacar o argumento de que, em alguns casos, a informação externa poderia não estar disponível, o que prejudicaria o exemplo de Otto e endossaria a posição de que as crenças de Inga sempre estarão disponíveis em sua memória:

Pode-se temer que o acesso de Otto ao seu caderno *de fato* venha e vá. Ele toma banho sem o caderno, por exemplo, e não consegue lê-lo quando está escuro. Certamente sua crença não pode ir e vir tão facilmente? Poderíamos contornar esse problema redescrivendo a situação, mas, em qualquer caso, uma desconexão temporária ocasional não ameaça nossa reivindicação. Afinal, quando Inga

está dormindo, ou quando está embriagada, não dizemos que a sua crença desaparece. O que realmente conta é que a informação está facilmente disponível quando o sujeito precisa dela, e essa restrição é satisfeita igualmente nos dois casos. Se o caderno de Otto muitas vezes não estivesse disponível para ele nos momentos em que as informações nele contidas seriam úteis, poderia haver um problema, pois a informação não seria capaz de desempenhar o papel de guia de ação que é central para a crença; mas se estiver facilmente disponível na maioria das situações relevantes, a crença não está em perigo (CLARK & CHALMERS, 1998, p. 15).

Parece claro, portanto, que essa perspectiva ressalta que a cognição estendida não exige uma disponibilidade perfeita e ininterrupta dos artefatos externos, mas uma disponibilidade prática que permita uma interação efetiva entre o indivíduo e os recursos externos quando necessário. Embora existam momentos em que o acesso ao caderno de Otto pode ser temporariamente limitado, como durante o banho ou em situações de baixa luminosidade, os autores argumentam que essas interrupções não invalidam a concepção da mente estendida. O critério essencial é que a informação esteja facilmente disponível na maioria das situações relevantes, onde a crença precisa ser utilizada como guia para a ação. Por isso, essa abordagem reconhece que os processos cognitivos podem se estender além dos limites do cérebro individual e incorporar elementos externos ao sujeito.

O interessante dessa abordagem é que ela prioriza os artefatos externos como novos andaimes cognitivos, reconhecendo o papel de tais artefatos como extensões funcionais e complementares dos processos funcionais. Ao mesmo tempo, ao afirmar que os processos cognitivos se estendem para além dos limites do cérebro individual, ela desafia a visão tradicional de que a cognição é estritamente confinada ao domínio interno da mente. O fato é que, sendo a cognição fortemente influenciada e moldada pela interação entre nossos processos internos e os recursos e ferramentas disponíveis externamente, temos como uma das consequências diretas a própria mudança nos conceitos de *agenciamento* e *responsabilização*. Afinal, não é suficiente reconhecer que a mente não é um fenômeno interno, mas uma entidade em constante interação com o mundo ao seu redor. Precisamos determinar quando uma crença é suficientemente reconhecida como verdadeira e em que medida deveríamos dissolver a responsabilidade entre o agente e o próprio artefato que está sendo usado.

Otto poderia reconhecer no bloco de notas uma série de informações que lhe permitiriam transitar entre os ambientes do Museu. Seria capaz de saber onde estão cada uma das coleções e dos quadros que estão sendo expostos, porém o bloco de notas é uma ferramenta meramente passiva que apenas facilita o acesso às informações e que não tem

uma capacidade cognitiva inerente. Otto poderia confiar, então, que as informações contidas no bloco de notas sejam adequadamente corretas, quando na verdade elas são falhas e conduzem a raciocínios equivocados. Se estou procurando recordar algo que não sei, como posso confiar que um processo cognitivo estendido seja capaz de tornar novamente emergente aquela informação?

Como se argumenta, não é suficiente apontar que o cérebro é um poderoso processador de informações, pois diante dos limites de sua capacidade de lembrar, calcular, navegar, planejar e aprender, faz uso de artefatos que nos sustentam na execução de tarefas cognitivas (HEERSMINK, 2017). Por isso, a capacidade de alguém agir intencionalmente no mundo, tomar decisões, fazer escolhas e executar ações com base em crenças, desejos e intenções implica um sentido de controle e autodeterminação sobre as próprias ações. O uso dos artefatos não-biológicos, como canetas, *smartphones* ou qualquer outro recurso tecnológico, então, permite que novos andaimes cognitivos possam ser construídos, estendendo as capacidades mentais além dos limites naturais do cérebro. Esses andaimes cognitivos funcionam como suportes estendidos que complementam e aprimoram nossas habilidades cognitivas, possibilitando a resolução de problemas mais complexos que, de outra forma, seriam inacessíveis para o cérebro isolado.

É provável que, em alguns casos, os artefatos forneçam recursos adicionais para moldar e influenciar como pensamos e nos comportamos, mas isso não significa que tais artefatos possam substituir a capacidade de agirmos intencionalmente. Em vez disso, os artefatos são usados como ferramentas ou auxílios para facilitar a expressão da agência. Há várias razões que poderiam explicar porque o cérebro biológico teria evoluído e amadurecido com tais capacidades para manipular o ambiente externo. Uma dessas razões seria a tentativa do cérebro em parasitar o ambiente para reduzir a carga de memória e até mesmo transformar a natureza dos próprios problemas computacionais (CLARK & CHALMERS, 1998). Foi essa habilidade de distribuir-se pelo ambiente que permitiu potencializar nossa capacidade de agir intencionalmente no mundo e, conseqüentemente, incorporar o ambiente em nossos processos cognitivos.

Por isso, a leitura tradicional de que o agenciamento e a responsabilidade possam situar-se apenas no poder de agir de acordo com suas próprias intenções ou na capacidade de decidir e exercer controle sobre as próprias ações e comportamentos precisa ser reconsiderada. A concepção tradicional de agenciamento e responsabilidade costuma focar a ideia de que o indivíduo é totalmente autônomo e livre para tomar decisões, como se fosse uma entidade isolada do ambiente. No entanto, essa abordagem não é capaz de determinar o papel dos artefatos não-biológicos nos processos cognitivos e nem explicar as interações que mantêm durante o seu acoplamento.

2. Agenciamento e externismo ativo: acoplando os artefatos à responsabilidade

A teoria da mente estendida, conforme concebem Clark e Chalmers (1998), se movimenta sobre o que os autores têm denominado de “externismo ativo”, isto é, um tipo de interação que os indivíduos mantêm com o ambiente e que envolve as suas ações (CARVALHO, 2020; ALONSO; RAMOS, 2022; VIERKANT, 2022). Enquanto um externismo de conteúdo apontaria para o histórico de interações causais e para o modo como o ambiente contribui para a determinação do conteúdo de nossos estados mentais, o externismo ativo compreende que a mente está imersa em um determinado contexto e profundamente entrelaçada com ele. Por isso, enquanto as teses tradicionais das Ciências Cognitivas, dentre as quais o cognitivismo, optaram por perspectivas descorporificadas, a posição da mente estendida se afirma na dinâmica percepção-ação, um enfoque que reconhece a interação constante entre o sujeito cognitivo e seu entorno.

O externismo ativo, então, reconceitualiza a mente não como um ente isolado, mas como um sistema que se aproveita ativamente do ambiente para facilitar processos de pensamento, resolução de problemas e tomada de decisões. Isso significa, portanto, que o externismo ativo vai além da mera influência do ambiente, enfatizando a coparticipação ativa do ambiente em nossos processos cognitivos. Nessa perspectiva, a cognição opera em um ciclo contínuo de percepção e ação, onde as fronteiras tradicionais entre a mente e o mundo se tornam mais fluidas, permitindo que utilizemos ferramentas, tecnologias e outros elementos do ambiente como extensões das nossas capacidades cognitivas. Em outras palavras, o externismo ativo equivale a dizer que alguns de nossos estados mentais são constituídos por elementos que estão “fora de nossa cabeça”.

É verdade, porém, que assumindo a tese da mente estendida precisamos evocar, por um lado, uma visão funcionalista mínima e, por outro, a existência de um banco de crenças estocadas na própria memória. Em relação ao funcionalismo, os processos mentais podem ser compreendidos como a função que desempenham em um sistema cognitivo. A função que tais processos exercem no próprio sistema é suficiente para considerá-los como estados capazes de realizar cognição (CLARK & CHALMERS, 1998, p. 7-8). O fato é que essa posição desobriga a necessidade de que tais estados mentais sejam realizados por um cérebro, pois o que está em jogo é a forma como isso é capaz de produzir conteúdo e comportamento.

Já em relação ao banco de crenças estocados na própria memória, a questão parece ser um pouco mais profunda. É difícil justificar boa parte de nossas ações sem referência

a um conjunto de crenças e desejos que não nos ocorrem o tempo todo, mas que estão ali quando outras suposições passam a ser consideradas, como é o caso das emoções (PERUZZO JÚNIOR, 2019). Por exemplo, se posso confrontar a tarefa com uma parte do mundo e isso funciona adequadamente bem, então, deveríamos reconhecer como parte de um processo essa parte do mundo.

Assim, no contexto do externismo ativo, as crenças armazenadas em nossa memória também são afetadas por essa interação. Para tal, nossas crenças não são formadas em um vácuo, mas são moldadas pelas experiências e interações que temos com processos que poderiam ocorrer externamente. Isso significa que o externo não é apenas um fator que influencia nossa cognição, pois ele desempenha um papel ativo na forja das estruturas cognitivas. Segundo Clark (1997, p. 214), “as complexidades da dinâmica usuário-artefato convidam à reflexão sobre um tópico mais geral: como conceber a fronteira entre o sistema inteligente e o mundo. Esta fronteira (...) parece ser um pouco mais plástica do que se supunha anteriormente”.

Assumindo essa infiltração da mente no mundo, as Ciências Cognitivas precisam lidar com a própria ideia de artefatos e, com isso, redesenhar o argumento de que o agenciamento possa ser pensado apenas como a capacidade dos seres racionais de realizarem ações intencionais e serem responsáveis por suas escolhas. Por isso, se considerarmos a teoria da mente estendida, há uma reviravolta significativa no conceito de agenciamento (*agency*): o agenciamento não é visto apenas como uma propriedade do cérebro individual, mas também como algo que pode se estender para além dos limites corporais. Tendo em vista essa leitura, precisamos considerar a maneira como o ambiente funciona para complementar e estender nossas capacidades cognitivas e de tomada de decisão. Isso implica, conseqüentemente, repensar o argumento de que a responsabilidade deva se concentrar exclusivamente nos processos individuais internos.

De qualquer forma, aceitar que o compartilhamento da responsabilidade com artefatos não humanos possa diluir a responsabilidade moral dos humanos amplia a própria ideia de agenciamento, levantando dúvidas sobre como determinar quem é o responsável quando as decisões são feitas em colaboração com as tecnologias. Assim, à medida que os artefatos desempenham um papel mais central, precisamos desenvolver novas normas e regulamentações para lidar com as implicações éticas da responsabilidade compartilhada. Talvez aqui Clark (1997, p. 215) tenha razão ao afirmar que existe, afinal, uma dificuldade bastante geral em traçar uma linha firme entre um usuário e uma ferramenta: “Uma pedra segurada na mão e usada para quebrar uma noz é claramente uma ferramenta. Mas se um

pássaro deixar cair uma noz do ar para que ela se quebre ao entrar em contato com o solo, o solo é uma ferramenta?”. Por isso, não se trata de reduzir a responsabilidade e o agenciamento à ação situada, mas repensar a forma como nós humanos devemos compreender os próprios artefatos, considerando que não podemos ser exclusivamente responsáveis por isso.

3. Repensando a responsabilidade: a ideia de ação situada

Para David Hume (2006, p. 32), todos os raciocínios acerca de questões factuais parecem fundamentar-se na relação de causa e efeito. Apenas por meio dessa relação é possível ultrapassar a evidência da memória e dos sentidos. Um observador que encontrasse um relógio ou qualquer outra máquina em uma ilha deserta concluiria que em algum momento houve um ser humano naquela ilha. Segundo Hume (2006, p. 33), as causas e os efeitos não podem ser descobertos pela razão, mas sim pela experiência. “Ninguém imagina que a explosão da pólvora ou a atração de um ímã poderiam descobrir-se por meios de argumentos *a priori*” (Hume, 2006, p. 33).

Com efeito, quando se aborda o tema da responsabilidade é crucial considerar, por exemplo, a análise que o Direito Penal faz desse assunto. Especificamente no âmbito jurídico, sobretudo no campo criminal, é fundamental compreender se alguém pode ser responsabilizado por um crime, levando em conta a existência do nexos causal que conecta a conduta (causa) ao resultado naturalístico (alteração no mundo externo).

No contexto brasileiro, por exemplo, o artigo 13 do Código Penal define a causa como a ação ou omissão sem a qual o resultado não teria ocorrido, adotando assim o que é conhecido como “teoria da equivalência dos antecedentes causais” (*conditio sine qua non*). Essa teoria sustenta que a conduta deve ser considerada a causa quando o resultado não teria ocorrido na ausência dessa conduta.

Deve-se somar à teoria da equivalência dos antecedentes causais o procedimento hipotético de eliminação de Thyrén, que reputa como causa todo evento que, ao ser mentalmente removido, resultaria no desaparecimento do resultado conforme ocorreu ou quando ocorreu. Por exemplo, no caso em que um indivíduo disparasse uma arma contra outro (conduta) resultando na morte da vítima (resultado) devido ao disparo (nexo causal), o método de eliminação hipotética de Thyrén asseguraria que, se o primeiro indivíduo não tivesse efetuado o disparo, o resultado de morte não teria ocorrido.

Além disso, para evitar uma regressão infinita, é fundamental que se estabeleçam critérios claros de demarcação. No caso de um fabricante de armas, para mencionar um exemplo recorrente, poderia ser discutida a responsabilidade pela produção de uma arma que eventualmente fosse utilizada para cometer um crime. Para que haja uma atribuição de responsabilidade, porém, é preciso considerar se o fabricante agiu com dolo ou culpa (responsabilidade subjetiva), ou se ele criou um risco relevante e juridicamente proibido (imputação objetiva).

Nesse sentido, retornando-se à questão do agenciamento estendido, John Protevi (2008) investigou o papel dos afetos, como raiva e pânico, e dos reflexos pré-cognitivos no contexto dos combates e do ato de matar. Segundo Protevi (2008), apenas 2% da população seria capaz de matar de forma fria, olho no olho, consciente do que está fazendo, sendo esses indivíduos categorizados como de baixa afetividade ou sociopatas em busca de estímulos. Protevi (2008) argumentou que a dicotomia “lutar ou correr” seria erroneamente extrapolada de conflitos entre espécies, nos quais é mais provável que a fuga ocorresse antes da luta efetiva. Embora seja verdade que em algumas espécies territoriais, como os leões, um novo macho alfa vitorioso possa vir a matar a prole do adversário derrotado, a inibição de atos violentos dentro da espécie seria a regra. Um exemplo que ilustra essa restrição poderia ser observado nos conflitos e assassinatos entre chimpanzés, nos quais as emboscadas frequentemente envolvem não apenas dois, mas muitas vezes sete ou oito animais atacando uma única vítima isolada.

No contexto da Segunda Guerra Mundial, apesar dos avanços no treinamento em equipe e no comando, que aumentaram a capacidade de envolvimento em combates letais, Protevi (2008) indicou que a taxa de disparos à queima-roupa das tropas de infantaria americana foi de apenas 15% a 20% do número total de tiros efetuados durante a guerra. Isso sugeriu que grande parte dos disparos realizados em combate foi mais uma demonstração de presença do que uma efetiva luta. Nesse âmbito, conforme observado por Protevi (2008), a maioria dos soldados não conseguiu realizar assassinatos com frieza, necessitando matar em um estado de dessubjetivação, ou seja, movidos por reflexos, fúria e pânico. Diante disso, surge a questão levantada por Protevi (2008): quem é o responsável pelos atos de violência quando os reflexos, a fúria e o pânico são ativados?

Para abordar essa questão, é possível encontrar argumentos no pensamento de Allan Hanson. Segundo Hanson (2008), os teóricos sociais estão cada vez mais conscientes de que todas as ações dos seres humanos, ou até mesmo as possíveis ações, não ocorrem isoladamente. Em vez disso, os seres humanos sempre atuam em conjunto, seja

com outras pessoas, ferramentas e objetos, animais, computadores ou, o mais comum, por intermédio de uma combinação de todos esses elementos. De acordo com Hanson (2008), seria impossível, por exemplo, realizar uma busca por um livro no catálogo de uma biblioteca *on-line* sem a participação ativa de um computador. Da mesma forma, percorrer grandes distâncias seria impensável sem o uso de um carro, trem ou avião.

Com base nesse raciocínio, Hanson (2008) argumentou que a responsabilidade moral deveria ser atribuída não apenas aos seres humanos, mas também às agências estendidas que englobariam elementos não-humanos. Esse enfoque é fundamentado no pensamento pós-modernista, que destaca que as ações são realizadas e devem ser explicadas em termos de combinações interrelacionadas entre os humanos e os não-humanos.

Com efeito, Hanson (2008) sustentou que a visão tradicionalmente humanista do indivíduo como um sujeito, originada no Renascimento e consolidada no início da modernidade, teria se tornado obsoleta como estratégia para explicar as ações morais, o que demandaria, por seu turno, que fosse substituída por explicações originadas nessa agência estendida, a qual é “fundamentalmente diferente do indivíduo humano e, portanto, as formas tradicionais de pensar sobre a responsabilidade moral dos seres humanos não se aplicam” (Hanson, 2008).

De acordo com Hanson (2008), a agência estendida se caracterizaria por ser um conceito aberto e ilimitado, permitindo a inclusão progressiva de mais participantes. Isso poderia acarretar uma reavaliação da responsabilidade, inclusive no que diz respeito à questão da regressão infinita como critério para a demarcação do Direito Penal, mencionada anteriormente, além de trazer novos elementos para a discussão da causalidade, na forma clássica proposta por Hume. Nesse contexto, a agência responsável por um homicídio em um acidente de trânsito poderia abranger diversos elementos, como as condições da via e as condições climáticas, entre outros, dependendo das circunstâncias específicas. Segundo Hanson (2008), existiria um ponto de retornos decrescentes, em que a participação de certos elementos se tornaria menos relevante (por exemplo, a cor da roupa da vítima, que pode influenciar a visibilidade do motorista, mas a diferença entre tons de marrom ou azul escuro seria de menor importância). Além disso, haveria fatores constantes e universais, como o efeito da gravidade, que deveriam ser desconsiderados.

Para Hanson (2008), uma perspectiva moral fundamentada em uma ética deontológica, na qual o pensamento de Kant se destaca, tenderia a restringir a responsabilidade moral apenas aos agentes capazes de agir intencionalmente, ou seja, aos seres humanos, o que seria chamado de *individualismo moral*. No entanto, Hanson (2008) pondera a noção

de uma agência estendida que incluísse tanto elementos humanos quanto não humanos, atuando de forma conjunta e intencional. Para ilustrar essa ideia, é apresentado o cenário hipotético de uma busca por uma pessoa perdida em uma região selvagem. Nesse caso, a operação de busca envolveria uma variedade de componentes, como cachorros farejadores, carregadores humanos, helicópteros com suas tripulações, equipes em solo realizando buscas, bem como o pessoal no centro de comando coordenando todas as atividades e os equipamentos de comunicação utilizados para manter a comunicação entre os envolvidos. Essa configuração completa constituiria a agência estendida responsável pela busca.

Por outro lado, o conceito de agência estendida apresentado por Hanson (2008) é debatido por aqueles que defendem a perspectiva do individualismo moral. De acordo com esta visão, apenas os seres humanos podem ser atribuídos com intenções e, portanto, somente eles poderiam ser considerados moralmente responsáveis por suas ações. Consequentemente, os cachorros, helicópteros e equipamentos de comunicação não seriam considerados moralmente responsáveis por estarem envolvidos na busca pela pessoa perdida. Para Hanson (2008), no entanto, a dificuldade com essa forma de pensar é que ela partiria do pressuposto de que os elementos não humanos seriam apenas ferramentas utilizadas por seres humanos estáveis e autônomos para atingir seus objetivos. Essa visão negligenciaria o fato de que os seres humanos seriam afetados e modificados pelo uso de determinadas tecnologias, como evidenciado quando se percebe que um homem com uma arma é diferente do mesmo homem desarmado; ou quando uma pessoa se transforma quando está atrás do volante de um carro.

Questão complementar no debate sobre a agência estendida diz respeito ao tratamento de recompensas e punições, que normalmente se aplicariam somente aos seres humanos. Carros, armas, computadores ou qualquer outra entidade não-humana não sofreriam consequências positivas ou negativas decorrentes de um determinado comportamento. Por exemplo, não faria sentido encarcerar uma arma envolvida em um crime. Contudo, segundo Hanson (2008), se o propósito de aplicar uma sanção seria incentivar certos tipos de ações em detrimento de outros, o foco deveria recair sobre as agências estendidas que produzem essas ações. Nesse contexto, poder-se-ia argumentar – no sentido de responsabilizar as agências estendidas – que quando uma criança se comporta mal, usando um brinquedo de forma violenta ou se recusando a compartilhá-lo, uma punição típica seria retirar o seu brinquedo. Hanson (2008) destaca que, ao liberar o *hacker* de computador Kevin Mitnick da prisão, a decisão judicial o proibiu de utilizar equipamentos por três anos como parte de sua punição. “Assim, é razoável entender que a punição é dirigida contra o relacionamento ofensivo como um todo, e não apenas contra uma parte

específica dele” (HANSON, 2008), de modo que a atuação da agência estendida seria eliminada, evitando, além disso, o surgimento de outras entidades similares, que, embora pudessem apresentar variações em seus componentes, ainda representariam ameaças capazes de causar danos do mesmo tipo ou semelhantes.

Outro aspecto que merece destaque é a concepção de inteligência desenvolvida por Santaella (2003), com base no pensamento de C. S. Peirce. Segundo Santaella (2023, p. 100), a presença de inteligência está relacionada à capacidade de aprender, crescer ou evoluir, independentemente do quão primitiva essa ação possa parecer. Como resultado, Santaella (2023, p. 100) sugere que o pensamento não se limitaria ao cérebro humano, uma vez que já existiriam estudos que fundamentam a ideia de que florestas e plantas também possuiriam capacidade de pensamento. De fato, esses ecossistemas poderiam cuidar de seus vizinhos doentes, alertar suas companheiras sobre perigos através de sinais elétricos transmitidos por redes de fungos e até mesmo fornecer uma solução açucarada para as raízes de árvores caídas há séculos, mantendo-as vivas.

De acordo com Santaella (2023, p. 108), é incontestável que programas de computador e algoritmos de IA demonstrariam algum nível de inteligência, desde que não fossem comparados exclusivamente com o padrão da inteligência humana:

Uma discussão mais clara sobre a presença ou não da inteligência na IA tem sido atrapalhada pelo fato de que a IA é alimentada por dados humanos, esses dados são selecionados por humanos, o design é modelado por humanos, os vieses são apontados por humanos etc. Ou seja, a intervenção humana é extensa demais para que, aparentemente, se possa reconhecer a presença da inteligência nas máquinas. Essa falta de reconhecimento poderia ser verdadeira se um fator muito relevante não fosse deixado para trás, a saber, os algoritmos de IA aprendem. Não é do acaso que surgiram os nomes de “aprendizagem de máquina” e, mais ainda, “aprendizagem profunda”. Onde existe aprendizagem, agenciamentos inteligentes entram em cena. Aprender implica um maior ou menor grau de autonomia transformadora. Essa autonomia, evidentemente em grau bem inferior à autonomia humana, está sendo assumida pela máquina (SANTAELLA, 2023, p. 108-9).

Hanson (2008) defende um raciocínio semelhante, argumentando que está cada vez mais evidente que os seres humanos não ocupariam um lugar especial no mundo, dada a sua conexão com outras formas de vida. Portanto, qualquer teoria que restringisse a responsabilidade moral exclusivamente aos indivíduos humanos estaria completamente desatualizada devido à abundância de interações, inter-relacionamentos e interdependência entre humanos e não-humanos.

Por outro lado, de acordo com Ronald Giere (2008), a ideia de distribuir a responsabilidade moral entre seres humanos e os artefatos que produzem resultados moralmente relevantes no mundo não se sustentaria. Por exemplo, helicópteros apenas executariam suas funções designadas, voando em determinadas velocidades e pairando quando necessário. Os elogios morais, conforme Giere (2008), deveriam ser direcionados apenas aos pilotos que, por escolha própria, decidiriam voar na altura das árvores para arriscar suas vidas e salvar pessoas perdidas. Compartilhar a responsabilidade moral com elementos não humanos, segundo Giere (2008), atenuaria a responsabilidade moral dos seres humanos.

Contudo, essa perspectiva não é consensual. Segundo Cash (2010), os indivíduos seriam apenas um dos possíveis centros de cognição, ação e responsabilidade. Assim, o aspecto mais relevante da investigação sobre responsabilidade e cognição estaria em enfatizar a responsabilidade coletiva pelas ferramentas cognitivas e práticas normativas que sustentam a agência estendida. Consoante Cash (2010), isso implicaria, frequentemente, mas não sempre, utilizar as premissas do individualismo moral.

Em nossa leitura, seria crucial enfatizar a responsabilidade coletiva em compreender e abordar de maneira mais direta os diferentes aspectos do contexto em que as pessoas vivem, pensam e agem. Essa abordagem possui respaldo nas considerações de Hutchins (1995, p. xiii), que destacava a importância de se considerar a atividade cognitiva a partir de um contexto, entendendo-o como um processo dinâmico mais amplo, e não como um conjunto estático de condições do ambiente. Concordando com a visão de Hayles (1999, p. 289), que argumentava que a sofisticação cognitiva dos seres humanos modernos se deveria à construção de ambientes mais inteligentes para trabalhar, torna-se evidente a necessidade de efetuarmos uma nova compreensão sobre a responsabilidade e o agenciamento estendido, uma vez que a cognição tem como característica misturar-se ao ambiente para encontrar formas de suporte e otimização.

Considerações finais

O avanço constante de tecnologias e ferramentas tem possibilitado a modificação do ambiente ao qual os seres humanos estão expostos através de artefatos que, cada vez mais, são incorporados para aprimorar a cognição humana. No entanto, muitos dos conceitos morais e jurídicos ainda em uso são baseados em construções antiquadas que há muito tempo deveriam ter sido reavaliadas e atualizadas. Nesse sentido, o agenciamento

to estendido, por um lado, é uma perspectiva que traz novas visões sobre o mundo e a responsabilidade humana. É crucial aprofundar esse conceito para as ciências que se desenvolverão ao longo do século XXI, pois ele oferece uma compreensão mais abrangente de como as tecnologias e ferramentas estão integradas à experiência humana e como elas afetam a noção de mente, mundo e corpo.

Por outro lado, ao estendermos nossa agência mental para artefatos não humanos, podemos estar diluindo o senso de responsabilidade moral, uma vez que a própria ideia de responsabilidade está ligada à intenção e à causalidade. Assim, deveríamos estabelecer, antes de tudo, em que medida um artefato não humano contribui para uma decisão ou ação para, na sequência, vincular o possível grau de controle que alguém possa ter sobre uma situação ou evento. À medida que assumimos essa nova leitura sobre o agenciamento e a responsabilidade, o desafio reside em equilibrar a expansão de nossa capacidade cognitiva em relação à função desempenhada pelos artefatos tecnológicos a fim de que a colaboração entre humanos e tecnologia não apenas amplie a justificação epistêmica de nossas ações, mas também preserve os fundamentos da responsabilidade individual e coletiva.

Referências

ALONSO, Bernardo Gonçalves & RAMOS, Ronaldo de Oliveira. A marca do Cognitivo e Cognição 4E. **Princípios: Revista de Filosofia**, Natal, v.29, n.58, jan./abr. 2022, p.24-48.

BURGE, Tyler. Individualism and the Mental. In: LUDLOW, P & MARTIN, N. (Eds.). **Externalism and self-knowledge**. Stanford: CSLI Publications, 1998.

CARVALHO, Eros Moreira de. A tese da mente estendida à luz do externismo ativo: como tornar Otto responsivo a razões? **Trans/Form/Ação**, 43(3), Jul-Sep. 2020, p.143-166.

CARTER, Adam J.; CLARK, Andy; KALLESTRUP, Jesper; PALERMOS, Orestis S. & PRITCHARD, Duncan (Eds.). **Extended Epistemology**. Oxford: Oxford University Press, 2018.

CASH, Mason. Extended cognition, personal responsibility, and relational autonomy. **Phenomenology and the Cognitive Sciences**, 9 (4):645-671, 2010.

CLARK, Andy. **Being there: putting brain, body, and world together again**. Cambridge: The MIT Press, 1997.

CLARK, Andy & CHALMERS, David. The Extended Mind. **Analysis**, v. 58, n.1, jan. 1998, p.7-19.

FARKAS, Katalin. Two versions of the Extended Mind Thesis. **Philosophia**, v.40, 2012, p.435-7.

GIERE, Ronald. Human moral responsibility is moral responsibility enough: A reply to F. Allan Hanson. **Phenomenology and the Cognitive Sciences**, 7, 2008, p. 425-7.

HANSON, F. Allan. The anachronism of moral individualism and the responsibility of extended agency. **Phenomenology and the Cognitive Sciences**, 7, 2008, p.415-24.

HAYLES, N. Katherine. **How we became posthuman: virtual bodies in cybernetics, literature, and informatics**. Chicago: University of Chicago Press, 1999.

HEERSMINK, Richard. Extended Mind and Cognitive Enhancement: Moral Aspects of Cognitive Artifacts. **Phenomenology and the Cognitive Sciences**, 16(1), 2017, p.17-32.

HUTCHINS, Edwin. **Cognition in the wild**. Cambridge: MIT Press, 1995.

HUME, David. **Investigações sobre o entendimento humano**. Tradução André Campos. São Paulo: Escala Educacional, 2006.

PERUZZO Júnior, Léo. Intentionality, Conceptual Content, and Emotions. **Revista de Filosofia Aurora**, Curitiba, v. 31, n. 54, set./dez. 2019, p.833-47.

PERUZZO JÚNIOR, Léo; CANDIOTTO, Kleber & KARASINSKI, Murilo (Orgs.). **Tendências contemporâneas de filosofia da mente e ciências cognitivas**. Curitiba: PUCPress, 2023.

PROTEVI, John. Affect, agency and responsibility: The act of killing in the age of cyborgs. **Phenomenology and the Cognitive Sciences**, 7, 2008, p. 405-13.

PUTNAM, Hilary. It ain't necessarily so. In: Putnam, Hilary. **Mathematics, matter and method**: philosophical papers. New York: Cambridge University Press, v. 1, 1975.

SANTAELLA, Lucia. **A inteligência artificial é inteligente?** São Paulo: Edições 70, 2023.

VIERKANT, Tillmann. **The tinkering mind**: agency, cognition, and the extended mind. Oxford: Oxford University Press, 2022.





NOS ENTRELAÇAMENTOS DA COGNIÇÃO, A EXPERIÊNCIA ESTÉTICA



Monica Aiub¹

Resumo:

Considerando a circularidade mente-corpo-mundo e os entrelaçamentos entre filosofia, ciência e arte nos estudos sobre a cognição, o presente artigo trata do papel da experiência estética para a compreensão dos processos cognitivos. Em seu percurso, aponta contribuições de Zeki, Ramachandran, Changeux e Couchot na formulação e análise da neuroestética e propõe o paradigma dinâmico da cognição 4E como alternativa para evitar as tendências reducionistas da análise. Dentre os elementos destacados no papel da experiência estética, encontram-se, principalmente, a dimensão do sentir, a criatividade e a empatia – elementos fundamentais para que possamos ver o mundo e a nós mesmos de outros modos, conhecendo e recriando nossas formas de existência num mundo em constante movimento.

Palavras-chave:

Experiência estética; Neuroestética; Cognição 4E; criatividade; empatia.

Abstract:

Considering the mind-body-world circularity and the interweaving between philosophy, science and art in studies on cognition, this article deals with the role of aesthetic experience in understanding cognitive processes. Along the way, it points out the contributions of Zeki, Ramachandran, Changeux and Couchot in the formulation and analysis of neuroaesthetics, and proposes the dynamic paradigm of 4E cognition as an alternative to avoid the reductionist tendencies of the analysis. Among the elements highlighted in the role of the aesthetic experience, there are, mainly, the dimension of feeling, creativity and empathy - fundamental elements for us to see the world and ourselves in other ways, knowing and recreating our forms of existence in a world in constant motion.

Keywords:

Aesthetic experience; Neuroesthetics; 4E Cognition; creativity; empathy.

¹ Doutora em Filosofia (PUC-SP), atua como orientadora filosófica, professora e pesquisadora no Espaço Monica Aiub: Filosofia, Arte e Cultura. www.monicaaiub.com.br. Contato: monica_aiub@uol.com.br

“As ciências cognitivas têm por objetivo descrever, explicar e, se possível, ampliar as principais disposições e capacidades da mente humana: linguagem, reflexão, percepção, coordenação motora, planejamento, decisão, emoção, consciência, cultura... [...] Desde seu surgimento na década de 1950, as ciências cognitivas eram pensadas como um conjunto de programas de pesquisa de diferentes disciplinas. Nascidos em um contexto científico fortemente marcado pelo nascimento da informática e pelo desenvolvimento de conceitos e técnicas de processamento formal da informação, elas agora estão intimamente ligados às neurociências.”

(Andler, Sciences cognitives, *Encyclopaedie Universalis*)

A definição de Andler, na *Encyclopaedie Universalis*, remete à compreensão das ciências cognitivas como a junção de esforços de diferentes disciplinas, constituindo programas interdisciplinares, a fim de “descrever”, “explicar” – portanto, aproximando-se da ciência – e “ampliar” – para tal, tendendo à tecnologia – “as principais disposições e capacidades da mente humana”. Gardner (1995) defende que a filosofia, além de ser a mais antiga das ciências cognitivas, pautou, com questões epistemológicas, a agenda inicial dos cientistas cognitivos. Ele descreve seis disciplinas iniciais para a constituição das ciências cognitivas: filosofia, psicologia, linguística, antropologia, inteligência artificial e neurociência. Todas atuando com uma “*episteme* comum” que comporta, além do trabalho interdisciplinar, diferentes abordagens ao problema mente-corpo, investigando e discutindo as implicações tecnológicas, éticas e epistemológicas das teorias da mente.

Os diferentes momentos e abordagens das ciências cognitivas (VARELA, 1994) – cibernética, cognitivismo, conexionismo e enação – comportam uma rede na qual ciência, filosofia, arte e tecnologia se entrelaçam nas tentativas de descrever, explicar, simular e ampliar as capacidades da mente humana. Mas a rede pode ser ainda mais ampla, envolvendo humanos e não humanos, natureza e cultura, filogênese e ontogênese, pesquisa etológica e simulações algorítmicas, simulação e criação...

Os pares acima apontados, e outros em que se pode pensar, não correspondem a uma visão dual quando observados pela perspectiva enativista (VARELA, 1994; VARELA; THOMPSON; ROSCH, 2001), que considera as capacidades cognitivas entrelaçadas ao vivido, uma cognição incorporada:

Propomos como nome o termo *enação* para salientar a convicção crescente de que a cognição não é a representação de um mundo preestabelecido elaborada por uma mente predefinida mas é antes a atuação de um mundo e de uma mente com base numa história da variedade das ações que um ser executa no mundo. A abordagem enativa encara então seriamente a crítica filosófica da ideia de que a mente é um espelho da natureza mas vai mais além, debruçando-se sobre este tema a partir do núcleo da ciência (VARELA; THOMPSON & ROSCH, 2001, p. 32).

Nesta perspectiva, a circularidade – entre a estrutura cerebral e os comportamentos e experiências; a descrição científica e a estrutura cognitiva do pesquisador; a reflexão e as crenças e práticas biológicas, sociais e culturais; as práticas biológicas, sociais e culturais do meio e o corpo – precisa ser considerada. Ou seja, toda cognição é incorporada, situada em uma cultura, em uma sociedade, em uma história. Ao mesmo tempo em que a história, a sociedade, a cultura são constituídas por experiências corpóreas, num movimento circular.

Assim, linguagem, reflexão, percepção, coordenação motora, planejamento, decisão, emoção, consciência etc. não ocorrem fora de certas práticas sociais, culturais e biológicas. São, portanto, resultantes de uma construção mútua de sujeitos e de um mundo compartilhado:

[...] Desvendar os mistérios da consciência constitui, talvez, o maior desafio já enfrentado pela ciência; algo que nunca poderá ser feito sem o auxílio da neurociência. O que deve ser abandonado, contudo, é a ingenuidade filosófica dos neurocientistas – uma ingenuidade que frequentemente se expressa na sua postura espontaneamente reducionista ou eliminativista. A ciência do cérebro deve ser uma ciência de como nós representamos nosso próprio cérebro. Não se trata de uma circularidade fútil, mas a reconhecimento de que questões epistemológicas não podem ser ignoradas por aqueles que praticam a neurociência cegamente (TEIXEIRA, 2004, p. 118).

Na circularidade mente-corpo-mundo há uma tentativa de suplantar a dualidade instaurada pela herança do momento cartesiano que seccionou mente e corpo, natureza e cultura, espírito e matéria, tornando o problema mente-corpo “insolúvel”. Esta herança dual também é revelada nas separações ocorridas na modernidade entre filosofia, ciência e arte, deixando à ciência o papel de descrever e explicar o corpo, enquanto à filosofia cabe a compreensão da mente. Por sua vez, a arte, antes *techne* – ou técnica para se bem fazer algo – e, a partir do século XIX, considerada capaz de envolver as emoções humanas e a empatia, torna-se o espaço da criatividade.

No encontro entre as várias disciplinas que compõem as ciências cognitivas em sua vertente enativista, filosofia, ciência e arte se entrecruzam, constituindo uma rede cujos fios são entrelaçados de modo a compor os “nós” que sustentam a existência. E, exatamente, a existência, o estar no mundo e a experiência do mundo que constituímos e que nos constitui é o ponto central para descrever, explicar, simular, criar e ampliar a cognição.

Do objeto da arte nas ciências cognitivas: as condutas estéticas

Se a filosofia pautou as questões epistemológicas e éticas desde os primeiros momentos das ciências cognitivas – e prossegue pautando; se a ciência marca sua presença nas tentativas de descrição e explicação dos fenômenos cognitivos; qual o objeto da arte?

Antes de esboçar uma resposta para um possível objeto da arte no fazer das ciências cognitivas, é preciso pontuar que arte pode ter muitas acepções. Como afirmou Couchot (2018, p. 54):

[...] Desde o Renascimento, o que se denomina ‘arte’ na cultura ocidental é nada mais que o aspecto particular e em perpétuo devir de um conjunto de práticas temporariamente codificadas, destinadas, talvez, a desaparecer, a se dissolver em outras práticas, ou a ser substituídas por outras, por razões que já podemos começar a suspeitar. Se isso for procedente, seria possível arriscar, retomando a última frase de *As palavras e as coisas* [Foucault], substituir a palavra homem pela palavra arte e afirmar que a arte se desvaneceria ‘como um rosto de areia na orla do mar’.

Afirma ainda Couchot (2018, p. 55) que, em ciências cognitivas, o estudo da arte não se limita ao estudo dos artefatos denominados como “obras de arte”, mas abarca, sobretudo, a investigação dos processos cognitivos envolvidos na percepção e na criação dos objetos estéticos, considerando, entre eles, além das denominadas ocidentalmente obras de arte,

[...] artefatos próprios às culturas que ignoram a arte (máscaras, cantos, danças ou objetos de culto), certos artefatos industriais ou artesanais (ferramentas, instrumentos, máquinas), ou até mesmo certos objetos naturais (sítios arqueológicos ou paisagens).

Sendo assim, o objeto de estudo são as condutas estéticas, isto é, o que uma experiência estética é capaz de produzir em cada um de nós e quais traços comuns podem ser observados.

O primeiro ponto a ser destacado no que se refere às condutas estéticas é a atenção. Couchot (2018, p. 58) descreve a atenção cognitiva como uma espécie de “feixe de luz dirigido a um ‘alvo’”, que funciona como um filtro, no qual atuam, de um lado, estímulos externos que provocam a identificação do alvo e, de outro, a decisão individual, pautada na experiência do sujeito, em sua memória, nas alterações de humor, assim como nas influências do ambiente, num “duplo movimento: filtragem e amplificação das percepções de um lado, projeção realizada pelo cérebro no mundo de suas pré-percepções de outro”. Ele também a descreve como uma capacidade selecionada pela evolução para permitir ao sujeito investigar seu entorno e ter um grau de previsibilidade.

A atenção estética é um tipo de atenção cognitiva voltada para objetos específicos: a beleza das formas. Por isso, morfotrópica. Ela provoca um estado afetivo: o prazer estético, ou seja, uma espécie de recompensa pela atenção àquela forma. Apesar de cada objeto estético provocar, em cada sujeito, em cada momento, diferentes reações e intensidades, este prazer manifesta-se no corpo, com alterações verificáveis e comportamentos expressivos. O prazer estético não é associado a uma região específica do cérebro, ele provoca estados afetivos gerando o sentimento estético e suas ressonâncias cognitivas. Nas palavras de Couchot (2018, p. 61), “A emoção e o prazer são sempre fortemente entrelaçados com apreciações, ideias e julgamentos.” Além de morfotrópica, a atenção estética é, também, para Couchot (2018), autoteleológica, isto é, é um fim em si mesma.

A atenção voltada para objetos estéticos em suas formas, então, pode gerar alterações orgânicas, movimentos corporais, comportamentos, pensamentos, valores, etc. e estes podem implicar em criação de novas formas e objetos estéticos. Couchot (2018, p. 67-75) distingue entre condutas estéticas receptoras, vinculadas à percepção dos objetos estéticos, e condutas estéticas artísticas ou operatórias, relativas à criação. As receptoras são discriminatórias e ativam sensações, por isso estão relacionadas ao prazer da contemplação. As condutas estéticas artísticas ativam habilidades e são associadas ao prazer da criação. Ambas estão conectadas entre si pelo prazer estético e se retroalimentam.

A antiga pergunta sobre “o começo” se coloca: onde se inicia o processo? Qual o ponto de partida da dimensão estética? Seria ele uma propriedade do objeto, qual seja, sua forma? Seria uma propriedade do olhar do sujeito que experiencia a percepção estética? Ou do sujeito que cria o objeto estético? Emergiria ele da relação entre o objeto estético e aquele que o contempla? Ou da relação entre este objeto e aquele que o cria? Ou, ainda, da relação entre aquele que cria e a aquele que contempla, mediada pelo objeto? Haveria alguma contribuição ou interferência do espaço ou do tempo no qual a relação se dá?

As respostas sobre “o começo” variam conforme o paradigma que conduz a pesquisa. Para o cognitivismo, a dimensão estética é natural e existe na forma de representações, ou seja, símbolos formais. Por este paradigma, a natureza é a fonte de toda a beleza e a arte é um processo de criação a partir da imitação das formas naturais. Para o conexio-nismo, é o sistema cognitivo do sujeito quem cria, portanto, é o sujeito – seja ele o artista ou o espectador – o ponto de partida da dimensão estética. Por fim, para o enativismo, a dimensão estética emerge da relação entre dois sistemas autopoieticos que interagem – de um lado o sujeito e de outro o objeto estético –, criando um mundo coerente de percepções que provocam emoções, sentimentos, ideias. Esta relação pode se dar a partir de um acoplamento natural ou através de uma interface:

Sentir prazer diante da visão de um quadro, ou ao ouvir uma peça musical, não é mais extrair informações de um sinal, decodificar os símbolos para lhes dar significado, é fazer emergir a própria forma das obras e fazê-las existir como mundo autônomo. Perceber é agir, fazer (COUCHOT, 2018, p. 442).

Cria-se um mundo ao criar uma obra de arte, mas também ao contemplá-la. E não se trata apenas de criar um mundo fictício como uma forma de encontrar prazer em algo artificial ou numa fuga da realidade. Cria-se a própria realidade. Para além da metáfora que a afirmação anterior poderia suscitar, a percepção estética pode promover a consciência da natureza biológica, da cultura circundante, dos elementos que compõem a realidade do entorno, das necessidades e possibilidades existentes, da alteridade, da variação e da diversidade das formas possíveis, impulsionando à construção de uma realidade diferente, que resista às opressões que impedem a vida. Por isso, “perceber é agir”.

A ação presente na conduta estética, seja ela receptora ou artística, ocorre no corpo e é movida pelo prazer estético. Mas dizer que ela ocorre no corpo não significa situá-la nos limites da pele do corpo de um sujeito que percebe ou cria o objeto estético. Isto porque corpo não pode ser entendido, aqui, como um receptáculo, um “corpo-envelope”, um recipiente:

[...] Algumas informações do mundo são selecionadas para se organizar na forma de corpo – processo sempre condicionado pelo entendimento de que o corpo não é um recipiente, mas sim aquilo que se apronta nesse processo co-evolutivo de trocas com o ambiente. E como o fluxo não estanca, o corpo vive no estado do sempre-presente, o que impede a noção do corpo recipiente (KATZ & GREINER apud GREINER, 2005, p. 130).

O corpo em questão é o resultado das interações, das experiências, das influências do entorno. Um corpo que, ao mesmo tempo, constitui o mundo e é constituído por ele. Um corpo que se movimenta: sente, pensa, se emociona, age e, ao agir, provoca movimentos, sensações, emoções, pensamentos e o agir de outros corpos. Ele não armazena informações, ele é constituído delas:

[...] Nosso sistema conceitual, que é encarnado e de raiz metafórica, ocupa um papel central, definindo as realidades cotidianas. Não há nada que esteja em um pensamento que não tenha estado também no sistema sensorio-motor do corpo (GREINER, 2005, p. 127).

As condutas estéticas, apesar de terem sido objeto de reflexões filosóficas durante a história, passam a ser estudadas pelas ciências cognitivas em sua perspectiva neurofisiológica, ou seja, a partir do estudo do corpo, mais precisamente, do cérebro.

Neuroestética

Neuroestética refere-se ao estudo das bases neurais da conduta estética. A apresentação oficial deste termo deu-se em 2002, no primeiro congresso sobre a disciplina neuroestética, ocorrido em São Francisco (CHANGEUX, 2013; COUCHOT, 2018; RAMACHANDRAN, 2014), e sua referência inicial é o trabalho de Semir Zeki (1998), mas a tentativa de estabelecer bases neurofisiológicas para tratar da arte e da beleza são anteriores. Abraham Moles, em *Teoria da informação e percepção estética*, publicado em 1958, propõe que obras de arte são mensagens que podem ser quantificadas em informações de dois tipos: semântica e estética, sendo a semântica expressa em símbolos e relativa a decisões, e a estética relacionada a estados mentais e afetivos intraduzíveis em símbolos lógicos, o que insere na informação estética a questão dos *qualia* (COUCHOT, 2018, p. 101). Umberto Eco (1986) também antecipa a discussão apresentando a “abertura” da obra de arte à interpretação do espectador, gerando uma espécie de “dialética” entre o autor e o público.

Contudo, a neuroestética proposta por Zeki pretende identificar leis neuronais relacionadas à beleza e à conduta estética a partir da investigação dos processos neuronais e das áreas do cérebro correlacionadas à percepção e criação de obras de arte. Para ele, há formas neuronais inatas que são ativadas quando o sujeito entra em contato com formas do mundo a partir de uma espécie de empatia (COUCHOT, 2018, p. 105).

No texto “*Art and the Brain*”, de 1998, Zeki compara as funções do cérebro visual (região do córtex frontal) com as funções da arte. Nesta comparação, o cérebro visual interessa-se em identificar, nos objetos do mundo, características que possibilitem uma categorização, buscando “constâncias” que permitam obter um conhecimento sobre o mundo. Tal conhecimento é organizado em um processo de “abstração”, construindo representações que são formadas a partir da eliminação de uma grande quantidade de informações da memória. A arte, segundo ele, teria a mesma função, trabalhando, também, com “constâncias” e “abstrações”, a fim de “buscar conhecimento em um mundo em mudança” (ZEKI, 1998, p. 78).

Por isso, Zeki (1998) considera que, de alguma maneira, os artistas atuam como os neurologistas, estudando o funcionamento do cérebro diante da experiência estética para produzir, a partir de formas singulares, abstrações que permitam obter algum conhecimento sobre o mundo. Desta maneira, apresenta a neuroestética como a busca dos fundamentos biológicos da estética, compreendendo estética como um produto do cérebro. Consequentemente, para ele, a formulação de uma teoria estética só é possível a partir da compreensão do funcionamento do cérebro humano.

Ramachandran (2014) sugere que as leis da estética são universais, extrapolando as barreiras culturais e interespecies. Ele cita os *bowerbirds*, pássaros-arquitetos na Austrália e na Nova Zelândia, que constroem seus ninhos com uma sofisticada “arquitetura” de formas e cores para atrair as fêmeas; nestes ninhos é possível identificar alguns princípios estéticos. As leis da estética, apontadas por ele, são: Agrupamento; Efeito de deslocamento de pico; Contraste; Isolamento; Esconde-esconde, ou solução do problema perceptual; Aversão a coincidências; Ordem; Simetria; Metáfora:

Não basta arrolar essas leis e descrevê-las; precisamos de uma perspectiva biológica coerente. Em particular, ao explorar qualquer traço humano universal como humor, música, arte ou linguagem, precisamos ter em mente três questões básicas: grosso modo, O quê? Por quê? e Como? Primeiro, qual é a estrutura lógica interna do traço particular que estamos considerando (correspondendo aproximadamente ao que chamo de leis)? Por exemplo, a lei do agrupamento significa simplesmente que o sistema visual tende a agrupar elementos ou traços similares na imagem em grupos. Segundo, por que o traço particular tem a estrutura lógica que tem? Em outras palavras, para que função biológica ele se desenvolveu? E terceiro, como o traço ou lei é mediado pelo mecanismo neural no cérebro? Todas essas três questões precisam ser respondidas antes que possamos afirmar genuinamente ter compreendido qualquer aspecto da natureza humana (RAMACHANDRAN, 2014, p. 166).

Além de apresentar as leis da estética a partir dessas três questões básicas (o quê, por quê e como?), pautado em observações de obras de arte e comportamentos de pessoas humanas e não humanas, Ramachandran (2014) propõe testes experimentais, baseados em RGP (resposta galvânica da pele), rastreamento ocular (técnica proposta por Alfred Yarbus) e o registro de atividade de células ao longo das vias visuais em primatas, comparando os registros de atividade diante da arte e diante de fotos antigas.

Para Changeux (2013), as leis estéticas são fruto da filogenética, mas, além delas, há uma cultura epigenética, oriunda do confronto que travamos com o ambiente ao longo da ontogênese. A cultura epigenética consiste em regras que regem a produção artística e as condutas estéticas. As regras que ele identifica e nomeia são: *adequação ao real* ou *mímese*; *consensus partium* ou harmonia das partes com o todo; *parcimônia*; *novidade*; *tranquilidade*; *exemplum* (CHANGEUX, 2013, p. 97-152):

Se a ciência consagra-se à busca de ‘verdades objetivas’, universais e que acarretem um progresso cumulativo do conhecimento; se a ética dedica-se à busca de uma ‘vida boa’ do indivíduo com e para os outros do grupo social; a arte, por sua vez, diz respeito à busca de uma ‘comunicação intersubjetiva’ que implica motivações e emoções em harmonia com a razão, mas sem progresso evidente e em constante renovação. A meu ver, desses três principais campos de atividade cultural do homem, a atividade científica, a regulação ética e a criação artística, a última é a mais antiga, pois já está presente no mundo animal. Ela é essencial ao reforço do laço social, pela universalidade dos modos de comunicação intersubjetiva que engaja. Do meu ponto de vista, as funções cognitivas, em particular a consciência e a atividade artística, são associadas a um maior desenvolvimento da organização cerebral, que se manifesta principalmente pela expansão do córtex cerebral e mais particularmente dos córtices de associação pré-frontal, parietotemporal e cingulado, em relação estreita com o sistema límbico (CHANGEUX, 2013, p. 97-8).

Changeux (2013, p. 91) apresenta a empatia como uma emoção moral, fundamental e universal, constituída por um dispositivo cognitivo inibidor da violência e, ao mesmo tempo, constituinte de laços sociais. Na arte, “a empatia intervém como ‘diálogo intersubjetivo entre as figuras, empatia do espectador com as figuras e entre o artista e o espectador, que emprega capacidade de atribuição, teoria da mente” (CHANGEUX, 2013, p. 125-6). Citando Theodor Lipps (1897), apresenta o conceito de “empatia estética”, que se refere à imitação, ou seja, ao apreciar uma obra de arte, o espectador “projeta a sua personalidade sobre o objeto de contemplação” (CHANGEUX, 2013, p. 126). Discute, na sequência do texto, as bases neurais da empatia, destacando a importância do córtex temporal, córtex frontal, sistema límbico...

O projeto de pesquisa em neuroestética supõe, na perspectiva de Changeux (2013), a desconstrução de pressupostos ideológicos como a dualidade corpo-mente, as oposições natureza-cultura e espiritual-material. Tais pressupostos são considerados por ele como frutos de uma visão dualista que não permite compreender a complexidade das funções cerebrais em seus múltiplos aspectos. Ele propõe uma abordagem “neuro-histórica”, complexa, que abranja questões genéticas, epigenéticas, desenvolvimentais, culturais, sociais, ambientais etc.; e que considere a plasticidade epigenética. Se concebermos o cultural como um “traço neurobiológico”; se identificarmos as atividades intelectuais como uma “manifestação da organização de nosso cérebro”; se considerarmos que nossos cérebros são plásticos e se modificam diante das interações e do vivido, a visão dualista não tem lugar (CHANGEUX, 2013, p. 98-9). Ainda assim, e ainda que compreenda elementos como a empatia, o projeto da neuroestética tem uma forte inclinação a um reducionismo, considerando o cérebro como centro de toda a investigação e da experiência estética.

Entrelaçamentos

A música não está na cabeça de quem ouve e, portanto, não é uma representação; a música não está nos sons produzidos pelos instrumentos e, portanto, não é uma realidade objetiva. A música é uma propriedade emergente de um acoplamento informacional bastante específico, entre um ouvinte com sua história de interações (seus hábitos de escuta) e estímulos sonoros (os quais são produzidos por sujeitos com grande expertise) configurados dentro de certos padrões históricos e culturais. Música é experiência — uma experiência que transcende o indivíduo, o tempo e o local.

OLIVEIRA (2014, p. 31)

O tempo inteiro interagimos e vivemos experiências que nos afetam e modificam, assim como provocamos, com nossas reações e nosso agir no mundo, novas experiências. Há, como afirmou Dewey (1979, p. 26), o “princípio de continuidade de experiência [que] significa que toda e qualquer experiência toma algo das experiências passadas e modifica de algum modo as experiências subsequentes”. Sensações, emoções, desejos, ideias, crenças, pensamentos, observações, dispersões, ações compõem nossas experiências. Mas também os objetos, os lugares, o tempo, o ambiente, outras pessoas, outros seres vivos, os sons, as cores, as formas, os hábitos... E ainda nossa história e a história dos diferentes elementos sociais e culturais que nos envolvem. Acrescidos, também, os marcadores somáticos (DAMÁSIO, 2000), registrados em nossas células a cada experiência vivida e sinalizadores de nossas futuras decisões, ações, sensações... enfim, experiências.

São todos fatores de um movimento ininterrupto e altamente complexo, que corresponde ao entrelaçamento de nossa cognição incorporada, situada, enativa e estendida – 4E Cognition [*embodied, embedded, enactive, extended*] (NEWEN *et al.*, 2018) – ao vivido. Entrelaçamento que se dá em circularidade, complexidade e plasticidade, gerando o constante, ininterrupto e complexo movimento que é próprio do existir.

Do entrelaçamento entre filosofia, ciência e arte, no paradigma dinâmico da cognição 4E, a experiência estética se dá num mundo em movimento – que não é pré-existente, mas é criado pelo próprio experienciar entrelaçado a suas condições; por um sujeito que experiencia – também constituído no experienciar e a partir dele, entrelaçado a sua história e suas condições; produzida diante de objetos estéticos, que não são fixos, mas emergem em suas formas ou significados diante das condições da experiência e dos sujeitos que experienciam, entrelaçados ao movimento dos espaços e tempos de criação, percepção, recriação; entrelaçando sensações, prazeres, emoções, pensamentos e ações:

Temos que elaborar novas perguntas metodologicamente calcadas em uma visão de que o fenômeno musical é emergente do viver de um corpo que ouve. Não se trata de um A que causa ou implica um B, como uma melodia que gera uma emoção ou uma energia que é transmitida por um corpo e engendra uma representação, substrato básico do processo cognitivo. Trata-se, portanto, de saber-se separar o que é descrição de um determinado nível de análise e o que é corpo vivendo aquela experiência emotivo-sonora (TOFFOLO & OLIVEIRA, 2021, p. 61).

Não basta identificar quais as regiões de nossos cérebros são ativadas durante a experiência estética, pois, embora possamos ter um mapeamento interessante sobre nossa cognição diante dos objetos estéticos e uma maior compreensão acerca de suas funções e das hipóteses relativas a nossos comportamentos, não conseguimos, ao menos ainda, derivar deste mapeamento a compreensão sobre nossos sentimentos, afetos, sensações, hábitos, relações ou, ainda, as possíveis interpretações que realizamos naquele momento. Também não basta – mesmo que conseguíssemos – mapear formalmente todos os signos de nossa cognição, tentando reproduzi-la:

A não ser que consideremos que a ciência constitui o único domínio da cognição e que neguemos que qualquer experiência, que não se traduz por inferências, por raciocínios lógicos [...] possa ter um verdadeiro valor cognitivo, concordamos em reconhecer na arte a capacidade de nos proporcionar, em graus diversos, certos conhecimentos do mundo e de nós mesmos: esse valor representa uma aptidão de explorar o mundo, de fazer parte dele, nem que seja para desejar mudá-lo (COUCHOT, 2018, p. 430).

Diante da experiência de um fenômeno, saber distinguir diferentes níveis de descrição do mesmo, permitir-se aprender com a experiência do vivido, criar novas formas de viver são entrelaçamentos possíveis entre ciência, filosofia e arte. Entrelaçadas, nos permitem novas experiências, com novos movimentos, percepções e criações.

Nos entrelaçamentos da cognição, a experiência estética nos traz a dimensão do sentir – como sensação e como sentimento; do afeto – a maneira como somos afetados e afetamos, mas também nossas afeições, paixões e até afecções; da comunicação – como compartilhamos as diferentes compreensões, interpretações, reflexões e, inclusive, criações acerca dos fenômenos, sendo capazes de acessar, compreender e construir, junto com os outros, através da empatia, o mundo que provoca nossas experiências e que também é construído por elas.

Talvez seja por isso que necessitamos tanto da beleza, seja ela de uma paisagem, de uma cena cotidiana, de uma obra de arte, de um fenômeno natural, de um movimento cultural... Não uma beleza estabelecida em padrões convencionais que apenas reproduzem e repetem ininterruptamente o mesmo, mas aquela que nos toca, nos transforma e nos permite ver o mundo de outros modos e recriá-lo no movimento incessante da vida.

Referências

ANDLER, D. Sciences Cognitives, **Encyclopaedie Universalis**. Disponível em: <https://www.universalis.fr/encyclopedie/sciences-cognitives/> Acesso em: 30mar2021.

COUCHOT, E. **A natureza da arte**: O que as ciências cognitivas revelam sobre o prazer estético. São Paulo: UNESP, 2018.

CHANGEUX, J-P. **O homem neuronal**. Lisboa: Dom Quixote, 1991.

CHANGEUX, J-P. **O verdadeiro, o belo e o bem**: Uma nova abordagem neuronal. Rio de Janeiro: Civilização Brasileira, 2013.

DAMÁSIO, A. **O erro de Descartes**. São Paulo: Cia das Letras, 2000.

DEWEY, J. **Experiência e educação**. São Paulo: Editora Nacional, 1979.

DEWEY, J. **Arte como experiência**. São Paulo: Martins Fontes, 2010.

ECO, U. **A obra aberta**. São Paulo: Perspectiva, 1986.

FODOR, J. A. The mind-body problem. *Scientific American*, 244(1): 124-32, jan.1981. Disponível em: <https://www.scientificamerican.com/article/the-mind-body-problem/>. Acesso em: 30mar2021.

GARDNER, H. **A nova ciência da mente**: uma história da revolução cognitiva. São Paulo: Edusp, 1995.

GREINER, C. **O corpo**: pistas para estudos indisciplinados. São Paulo: Annablume, 2005.

MOLES, A. **Teoria da informação e percepção estética**. Rio de Janeiro: Tempo Brasileiro, 1969.

NEWEN, A.; GALLAGHER, S. & BRUIN, L. D. **The Oxford handbook of 4E cognition**. New York: Oxford University Press, 2018.

OLIVEIRA, L. F. O estudo da música a partir do paradigma dinâmico da cognição. *Percepta* 2(1), 17-36, 2014.

RAMACHANDRAM, V. **O que o cérebro tem para contar**. Rio de Janeiro: Zahar, 2014.

TEIXEIRA, J. F. **Filosofia e ciência cognitiva**. Petrópolis: Vozes, 2004.

TOFFOLO, R. B. G. & OLIVEIRA, L. F. O conceito de mente, o fato musical e a questão metodológica em cognição musical. In: GERMANO, N. G. & RINALDI, A. (Orgs.). **Música, mente e cognição**: A pesquisa em Cognição Musical no Brasil. Curitiba: ABCAM, 2021.

VARELA, F. **Conhecer**: As ciências cognitivas – tendências e perspectivas. Lisboa: Instituto Piaget, 1994.

VARELA, F; THOMPSON, E & ROSCH, E. **A mente corpórea**: ciência cognitiva e experiência humana. Lisboa: Instituto Piaget, 2001.

ZEKI, S. Art and the Brain. *Daedalus*, vol. 127, no. 2, 1998, pp. 71–103. Disponível em: *JSTOR*, www.jstor.org/stable/20027491. Acesso em: 1 Aug. 2021.





FILOSOFIA DA PSICOLOGIA DE LUDWIG WITTGENSTEIN: AS PALAVRAS NO FLUXO DA VIDA



Arturo Fatturi

Resumo:

Apresentamos uma tentativa de esclarecer alguns pontos básicos a filosofia da psicologia de Ludwig Wittgenstein. O objetivo é esclarecer os pressupostos da investigação de Wittgenstein nesta área. Buscamos mostrar os objetivos da filosofia da psicologia e sua relação com os usos dos conceitos psicológicos na vida cotidianas, a fim de compreender expressões que visam descrever emoções e sentimentos. Primeiramente, apresentamos algumas das dificuldades impostas pelo texto de Wittgenstein. Em seguida, propomos algumas soluções plausíveis. Tentamos mostrar que a classificação dos conceitos psicológicos como “expressivos” apenas pode trazer clareza quanto aos usos dos mesmos, se forem considerados nos jogos de linguagem que envolvem nossa compreensão das ações relacionadas às circunstâncias ordinárias da vida e uso dos mesmos e não como base de possíveis teorias explicativas. Os conceitos psicológicos devem ser analisados nas circunstâncias em que são utilizados que são, por sua vez, o fluxo da vida.

Palavras-chave:

Wittgenstein; linguagem; filosofia; psicologia; mundo interior.

Abstract:

We present an attempt to clarify some basic points of Ludwig Wittgenstein's philosophy of psychology. The aim is to clarify the assumptions of Wittgenstein's research in this area. We seek to show the objectives of the philosophy of psychology and its relationship with the uses of psychological concepts in everyday life in order to understand expressions that aim to describe emotions and feelings. First, we present some of the difficulties imposed by Wittgenstein's text. Next, we propose some plausible solutions. We try to show that the classification of psychological concepts as “expressive” can only bring clarity to their uses, if they are considered in the language games that involve our understanding of actions related to the ordinary circumstances of life and use, and not as a basis of possible explanatory theories. Psychological concepts must be analyzed in the circumstances in which they are used, which are, in turn, the flow of life.

Keywords:

Wittgenstein; Language; philosophy; psychology; inner.

Psychological concepts are just everyday concepts. They are not concepts newly fashioned by science for its own purposes, as are the concepts of Physics or chemistry.

Ludwig Wittgenstein (RPPii, §62)

1.

A filosofia da psicologia de Ludwig Wittgenstein é um tema complexo, uma vez que não apresenta teses, hipóteses ou explicações que possam facilmente ser estabelecidas. Antes, seus escritos sobre filosofia da psicologia são explorações dos usos dos conceitos psicológicos. Os conceitos, tais como “dor”, “pensamento”, “expectativa” são apresentados em suas aplicações em vários exemplos e Wittgenstein explora os resultados dos diferentes aspectos de cada exemplo. No entanto, antes de interpretar estes escritos como uma investigação inacabada ou que o autor abandonou, deve-se considerar que o texto de Wittgenstein é um retrato, por assim dizer, do estado de desenvolvimento de uma investigação conceitual que ele estava elaborando. Logo, os textos são o que se pode denominar como “obra em progresso”. Mesmo o texto que foi publicado sob o título de *Investigações Filosóficas* deve ser considerado desta forma: uma obra em progresso.

Este aspecto da filosofia de Wittgenstein sempre foi motivo de controvérsia entre os filósofos e as pessoas que, por alguma razão, se aventuram em sua filosofia. Isto se mostra no fato de que ele próprio apenas preparou para publicação o manuscrito do *Tractatus Logico-Philosophicus*, já que os demais textos publicados com seu nome são frutos do trabalho de editores, os quais ou foram seus alunos durante seu período como professor na Universidade de Cambridge ou eram professores em outras universidades que frequentavam suas aulas. Seja como for, uma alegação contra a filosofia de Wittgenstein é a de que sempre se necessita exercer algum esforço interpretativo quanto aos textos e quanto à continuidade dos pontos de vista discutidos em cada texto.

Wittgenstein não foi um filósofo tradicional, isto é, alguém treinado nas tarefas da filosofia acadêmica que exige certas regras de publicação, convívio e discussão. Seu estilo de escrever sempre foi mais literário que explicitamente argumentativo, tal como é o padrão acadêmico. Esta característica é acrescida pela complexidade de seus manuscritos e datiloscritos, uma vez que estava sempre organizando os grupos de textos escritos sob várias formas ou conjuntos. Assim, é mais adequado dizer que os títulos das publicações denominam certas coleções de determinados escritos.

Contudo, dado o que já se conhece de seu processo de escrita, se ele continuasse a revisar, certamente elaboraria uma coleção diferente de observações, mudaria a ordem

das seções, acrescentaria ou suprimiria seções completas e, com isto, se teria agrupamentos de textos diferentes. Alguns escritos que fazem parte das publicações foram acrescentados pelos editores, já que os encontraram no mesmo conjunto, ainda que não se tenha uma indicação de Wittgenstein de que deveriam ser publicados numa mesma obra ou em separado. Este é, por exemplo, o caso da Segunda Parte do texto das *Investigações Filosóficas*. Tal situação inspirou alguns projetos editoriais a fim de estabelecer alguma ordem nos escritos que Wittgenstein legou ao mundo filosófico. Atualmente, há um projeto editorial, localizado na Universidade de Bergen, (Noruega) que apresenta bastante sucesso para estabelecer a padronização nos escritos que até agora estão disponíveis.

Este contexto se deve, por um lado, à forma com que Wittgenstein organizou ou deixou seus escritos e, por outro, deve-se considerar que a forma com que escreveu sua filosofia apresenta um caráter também literário, em que várias formas ou estilos são usados. Ou seja, quanto a este último ponto, deve-se considerar que a leitura dos escritos exige bastante do leitor, de modo que somos apresentados a ideias através de exemplos e da discussão de resultados parciais ou a possibilidades de resultados com base nestes mesmos exemplos.

Normalmente, quando uma possível conclusão é apresentada, Wittgenstein passa a considerá-la sob outro ponto de vista; as seções não apresentam uma clara continuidade argumentativa e, muitas vezes, a linha de raciocínio que o conduziu até uma possível conclusão é retomada e toma outra direção. Obviamente que estes aspectos também apresentam dificuldades para o leitor que espera encontrar escritos acadêmicos e, na mesma medida, para o pesquisador que deseja considerar os argumentos de Wittgenstein a partir de certo ponto de vista. O caso do presente ensaio não será muito diferente disto e, como é comum nos estudos da filosofia de Wittgenstein, o leitor poderá encontrar outros pontos de vista que partem de bases interpretativas diferentes.

Esta introdução quanto às dificuldades do texto e do estilo de escrita de Wittgenstein não deve ser considerada como uma desculpa para a apresentação de um texto superficial, mas como a explicitação de um ponto de partida. Considera-se que o tema da Filosofia da Psicologia se apresenta como uma exploração ou investigação conceitual e não como um amontoado de escombros ou partes desconexas de uma construção teórica inacabada que devem ser colocados em certa ordem. Considera-se que a discussão de Wittgenstein sobre o tema e a forma como ele as apresenta são deliberadas.

Se notarmos as várias remissões que Wittgenstein faz em seus escritos ao tema da psicologia, fica claro que ele conhecia as discussões teóricas e experimentais neste campo de estudos. Exemplo disto é que, ao final da segunda parte do texto das *Investigações Filosóficas*, ele faz um comentário quanto à confusão conceitual da psicologia e a existência de vários

métodos experimentais. Ao mesmo tempo, também se sabe que ele demonstrava interesse pela obra *Principles of Psychology*, de William James, sobre a psicologia, ainda que muito mais pelas afirmações filosóficas de James - as quais considerava inspiradoras em razão de serem metafísicas - do que por suas teorias explicativas oriundas de experimentações.

Deve-se lembrar também que, durante a segunda grande guerra mundial, Wittgenstein prestou serviço de guerra como enfermeiro em um hospital que recebia soldados feridos nas batalhas. Estes soldados apresentavam vários problemas físicos e psíquicos devido aos combates e ao morticínio nas batalhas. Portanto, ele conhecia os procedimentos psiquiátricos e a aplicação de metodologia psicológica para o tratamento dos traumas e ansiedades vivenciados pelos soldados.

Será necessário traçar um certo caminho investigativo que permita a discussão dos conceitos psicológicos e não, como já deve estar claro até aqui, encontrar algum tipo de teoria na qual se possa inserir as considerações de Wittgenstein. Neste ponto, encontram-se alguns caminhos plausíveis: primeiramente, pode-se traçar a relação entre a investigação dos conceitos psicológicos e a investigação conceitual que Wittgenstein realiza nas *Investigações Filosóficas* e, com base nisso, buscar esclarecer como se dá a compreensão dos conceitos psicológicos; outro caminho plausível seria: partir de certas passagens, que se pode encontrar em conjuntos de textos, e tentar esclarecer como ele procede no tratamento dos conceitos psicológicos.

Neste segundo ponto encontramos certas observações nos dois volumes das *Remarks on the Philosophy of Psychology* (RPPi e RPPii) e no volume 2 de *Last Writings on the Philosophy of Psychology* (LWPPi e LWPPii). Estes dois conjuntos de textos apresentam certos desenvolvimentos quanto aos conceitos psicológicos que permitem compreender como Wittgenstein tratava o tema. No entanto, é no texto de *LWPP* que ele fornece uma investigação mais longa. Mesmo assim, não se deve esperar um tratamento teórico, e, sim, uma análise conceitual das nuances de uso dos conceitos psicológicos e o que aprendemos com a mesma. Outro aspecto a ser ressaltado é o de que estes volumes são publicações posteriores ao texto das *Investigações Filosóficas*, embora tenham sido escritos durante a preparação dos escritos que constituem a Parte 2 do texto. São, portanto, “observações” preparatórias que posteriormente Wittgenstein selecionou para publicação.

Uma outra opção consistiria em compreender como Wittgenstein trata a questão dos conceitos que fazem referência ou visam fazer referência aos estados interiores. Conceitos como os de “dor”, “pensar”, “intenção” e outros, que parecem indicar que algum processo interior ou mental está ocorrendo. Poder-se-ia começar analisando os resultados da discussão quanto à privacidade do significado que ocorre nas seções 243 até 315 do texto das *Investigações Filosóficas*. Nestas seções, Wittgenstein analisa a possibilidade

de que certos usos de nossos conceitos visam significar experiências e processos que, assim se concebe, ocorrem no mundo interior. A proposta surge logo após a análise do conceito de “seguir regras”, que se localiza entre as seções 185 e 242. O resultado desta análise, resumidamente, é que aplicamos nossas palavras através de regras de emprego e o significado das mesmas depende de como são empregadas em diferentes jogos de linguagem. Ao mesmo tempo, a prática de emprego inclui as ações que se realiza ao empregar estes conceitos.

Portanto, perguntar “qual o significado do conceito X?” implica buscar definir apenas uma parte das possibilidades de emprego do conceito. A prática de usar as palavras é uma ação e não a simples exibição e definição de um símbolo (no caso, o conceito X). Isto quer dizer que sentido, significado, compreensão e referência do conceito estão intrincados em suas diferentes aplicações. Ora, a discussão da possibilidade da linguagem privada diria respeito aos significados das palavras que visam comunicar algo do mundo interior e das sensações, a partir de um modelo de definição dos conceitos. A discussão diz respeito, então, à compreensão da aplicação de nossas palavras quando elas visam dizer algo, ou quando se pretende dizer algo, sobre o mundo interior e, nesse sentido, os significados dos conceitos apenas seriam acessíveis a quem pudesse observar ao que os conceitos se referem.

Entretanto, mesmo estas palavras devem seguir regras de emprego e estas, por sua vez, não são regras privadas, uma vez que o falante as aprendeu através dos mesmos métodos que todas as outras pessoas. Logo, os critérios de correção do emprego dos conceitos devem ser públicos, isto é, uma linguagem privada seria impossível, já que seu originador teria de seguir regras acessíveis a qualquer pessoa, sob o risco de sua linguagem ser incompreensível, para ele mesmo. Assim, o que força a ideia da privacidade do significado é, na verdade, a imagem que se adota quanto ao que seja o mundo interior de uma pessoa: um âmbito privado, particular e do qual apenas seu possuidor poderia afirmar qualquer coisa com conhecimento. Dessa maneira, a questão agora passa a ser explorar a imagem de mundo interior que serviu de critério para a consideração quanto à possível privacidade do significado. Mas o texto das *Investigações Filosóficas* não apresenta conclusões programáticas. Diferentemente disso, apenas diz que uma linguagem privada seria impossível.

Não obstante, deve-se considerar que o texto das *Investigações* não é uma investigação acabada, tal como foi dito no início deste ensaio, mas um conjunto de textos que reflete o estágio do pensamento de Wittgenstein sobre vários temas conexos naquele momento da preparação dos escritos. Ora, os textos sobre a filosofia da psicologia são desenvolvimentos daqueles mesmos temas iniciados nas *Investigações Filosóficas*, mas com uma orientação diferente, qual seja, a análise dos conceitos psicológicos.

2.

A principal dificuldade que se encontra nas discussões de Wittgenstein quanto ao que denominou por filosofia da psicologia é a ausência de sistematização, isto é, não se encontra um ponto básico que tenha sido teoricamente mostrado e do qual se pode apreciar todo um desenvolvimento posterior. No entanto, tal característica é proposital, visto que Wittgenstein não propõe teorização alguma quanto ao significado, para só então principiar alguma investigação quanto à psicologia. Este é um dos aspectos mais chocantes para o leitor contemporâneo de estudos e investigações de filosofia da mente: as referências de Wittgenstein, ao que parece, são de alguma versão do Behaviorismo, predominante em seu tempo (anos 1940), ou, como foi dito mais acima, o manual de psicologia de William James.

Entretanto, na afirmação do último capítulo das *Investigações Filosóficas*, através da qual distingue a confusão conceitual da existência de uma variedade de métodos experimentais, pode-se perceber qual o direcionamento de suas investigações na psicologia, a saber: a investigação deve ser conceitual e não empírica. Wittgenstein considerava que as investigações conceituais se constituem na orientação das investigações em filosofia. Mesmo assim, encontramos algumas tentativas de sistematização. Uma primeira tentativa aparece em *RPPi* no momento em que Wittgenstein se questiona se deve tratar todos os conceitos psicológicos como conceitos de fenômenos e, assim, todo o campo da psicologia como marcado por experiências. Contudo, ele abandona esta tentativa sem maiores desenvolvimentos. Uma segunda tentativa aparece em *RPPii*, no qual Wittgenstein considera a elaboração de alguma sistematização quanto a estes conceitos. Na seção 63, por exemplo, ele pondera o seguinte:

Plano para tratamento dos conceitos psicológicos. Verbos psicológicos caracterizados pelo fato de que a terceira pessoa do presente pode ser identificada por observação, a primeira pessoa não. Sentenças em terceira pessoa do presente: informação. Na primeira pessoa do presente, expressão.

No entanto, ao final da passagem ele agrega as seguintes palavras “não de todo correto”, e, a partir disto, passa a distinguir e analisar diversos jogos de linguagem com certos conceitos. Por exemplo, os conceitos de pensamento, emoção e sentimento. Esta estratégia é a mesma que se percebe na sua análise da possibilidade de uma linguagem privada, na qual não ataca diretamente a possibilidade em questão, mas, através da análise dos usos dos conceitos que fazem parte da possibilidade, demonstra que tal linguagem não

teria sentido. O que tal análise nos permite compreender, num primeiro momento, é que Wittgenstein não se utiliza da *introspecção*¹ como método de análise dos conceitos, uma vez que a introspecção apenas faria sentido num contexto em que certo tipo de definição é adotado, a saber, a definição ostensiva e que, em segundo lugar, se supõe a existência de um âmbito a ser conhecido ou observado pela introspecção. Nesse viés, definir o conceito de “intenção” por meio da introspecção implicaria definir ao que a palavra se refere, qual estado ou evento o sujeito introspecta em seu mundo interior. Ou seja, tal método, para ser plausível, consistiria em realizar uma definição ostensiva, na qual o objeto designado seria um “objeto mental”.

Ora, lembrando a ordem argumentativa, nas primeiras seções das *Investigações* ele já havia demonstrado que a definição ostensiva traz consigo vários problemas e que, portanto, não se deve tratá-la como padrão de definição. Normalmente, considera-se que a definição ostensiva é a “ligação” de um conceito com algum objeto da realidade, mas apenas apontar para um objeto que esta ou aquela palavra designaria não fornece clareza sobre como a palavra deve ser usada. A partir disso, seguindo o texto das *Investigações*, tem-se a discussão sobre a ideia de que se usa as palavras através de regras de aplicação para as mesmas e estas regras, por sua vez, estão contidas, por assim dizer, no próprio aprendizado da relação para com a linguagem e no fato de que o uso das palavras ocorre dentro de certos contextos de ação. Tais contextos de ação, nos quais os usos das palavras se entrelaçam com as ações, são os jogos de linguagem.

A partir dessas análises e direcionamentos, surge a discussão da possibilidade de uma linguagem privada, isto é, uma linguagem na qual apenas o originador da mesma poderia conhecer ou saber o que as palavras significam, uma vez que se tratariam de eventos interiores ao mesmo e que só ele teria acesso. Esta discussão, tradicional agora na filosofia, e que se inicia na seção 243 e segue até seção 315 do texto das *Investigações Filosóficas*, termina, como já se disse mais acima, por negar a plausibilidade da privacidade do significado, dado que as regras que o originador deveria empregar para construir sua linguagem devem ser públicas. Isto é, o originador ou possível privatista deverá empregar regras que aprendeu de forma pública, ou seja, que outras pessoas também poderiam aprender, e, portanto, a ideia de privacidade do significado não faz sentido. Ao mesmo tempo, a implausibilidade da privacidade do significado conduz a questionar a concepção de que o mundo interior poderia ser um âmbito separado do mundo exterior, devido ao

1 A introspecção é aqui definida como o exame de nossas próprias ideias e sentimentos. Por definição, este exame apenas pode ser realizado pela pessoa que tem acesso ao que vai examinar, isto é: a suposição básica da proposta de um exame introspectivo é a existência de um mundo interior privado e acessível apenas ao seu possuidor.

fato de que as palavras usadas para comunicar emoções e sentimentos devem seguir as regras de emprego para as mesmas.

Tais argumentos nos permitem compreender por quais razões a análise dos conceitos psicológicos realizada por Wittgenstein usa exemplos de emprego dos conceitos, mas não apresenta uma teoria psicológica e nem discute uma teoria deste tipo. Sua análise consiste em tratar dos casos de uso dos conceitos psicológicos em jogos de linguagem. Ou seja, ele analisa os conceitos em suas várias relações com as ações que estão ligadas ao emprego destes conceitos. Aqui, pode-se fornecer uma tentativa de explicação quanto ao que parece insatisfatório nesta análise: por um lado, somos impelidos a buscar, nas análises apresentadas por Wittgenstein, a mesma concepção de objetividade que se encontra em textos científicos ou de psicologia, isto é, uma hipótese ou teoria é apresentada, críticas são desenvolvidas ou, também, busca-se mostrar que uma determinada conclusão não se segue; por outro e, mais especificamente, no que diz respeito à palavra “psicologia”, busca-se encontrar certas afirmações acerca do que se passa na mente ou no interior de uma pessoa, isto é, procura-se algumas afirmações sobre o que são as emoções, estados internos, sensações, ansiedade etc.

Essas duas características da nossa maneira de conceber “como” deve ser um texto investigativo são frustradas quando se acessa à investigação de Wittgenstein sobre os conceitos psicológicos. Nos conjuntos de textos em que trata destes conceitos, encontram-se apenas três possibilidades de classificação dos conceitos e todas são deixadas de lado. Além disso, Wittgenstein, já desde o *Caderno Azul*, havia condenado o costume dos filósofos em tentar “copiar”, para suas investigações, o mesmo tipo de trabalho elaborado pelo cientista, o qual consiste em criar hipóteses experimentais e testá-las. Portanto, ele não faria uma tentativa de investigação em que o método científico fosse o modelo a ser seguido. No que diz respeito à segunda maneira como se acessa as investigações na psicologia, encontra-se em Wittgenstein a busca por compreender a “imagem” de que seres humanos são compostos por dois âmbitos, por assim dizer: o âmbito mental ou interno, em contraposição ao âmbito externo ou físico, o qual incluiria o corpo. Esta segunda característica explica por qual razão Wittgenstein não busca, em sua filosofia da psicologia, desvendar ou revelar *que tipo* de atos, eventos ou estados estariam contidos nas definições de conceitos psicológicos. Antes, sua tarefa é compreender como a imagem de “interno/externo” se impõe ao entendimento em relação ao vocabulário psicológico e quais prejuízos ela traz em relação a se compreender os usos destes conceitos.

3.

Ao evitar o método da definição ostensiva e, conseqüentemente, não usar a introspecção como caminho para encontrar uma definição dos conceitos psicológicos, Wittgenstein corre o risco de admitir que os conceitos psicológicos apenas poderão ser compreendidos em seu uso, a partir do comportamento. No entanto, ainda que não negue o papel importante do comportamento na compreensão do uso destes conceitos, ele também não aceita o comportamentalismo como saída. No texto das *Investigações Filosóficas*, ele apresenta seu interlocutor acusando-o de ser um comportamentalista disfarçado (IF, §307) e, ao negar que seja este o caso, chama a atenção para os primeiros passos que esta atribuição pressupõe, isto é, de que se fala de estados e eventos mentais, os quais ainda não foram definidos (IF, §308). Chama a atenção para a expressão “conhecer um processo mais de perto” que leva a buscar a observação de que algo que acontece num meio que ainda não se consegue definir, apesar de se acreditar que já foi identificado, isto é, que se tem um mundo interior que distinguimos de nosso comportamento ou do mundo exterior.

Neste raciocínio, aplica-se uma imagem do que seja observar um processo: ele tem início, meio e fim; pode ser acessado através da observação. Ou seja, se assim se define processos, então, um “processo espiritual” (IF, §306) deve ter a mesma constituição, mas com outra qualidade, uma vez que ocorre no espírito. Ao se negar que tais processos ocorram, tudo se parece a dizer que o local onde ocorrem também é negado. No entanto, Wittgenstein não nega que exista mundo interior ou que a expressão “agora me lembro que...” tenha significado. O caso é que, como todos os conceitos psicológicos, a imagem de que fazem referência a algo paralelo ou concomitante ao mundo físico sugere a concepção de que devem ocorrer no mundo “não-físico”, isto é, na alma, mente ou mundo interior. Tal esclarecimento, dirá Wittgenstein, é o que “nos causa perplexidade, pois desmorona a comparação que deveria nos ajudar a tornar nossos pensamentos compreensíveis para nós” (IF, §308). Parece que se está negando os processos espirituais ou mentais, mas, diz ele, “é evidente que não se pretende negá-los”.

Uma possível rota de solução para o paradoxo que agora se apresenta é a consideração quanto ao que Wittgenstein considerou como critério para proferimentos em primeira e terceira pessoa, isto é: proferimentos em primeira pessoa são expressivos e em terceira pessoa são informativos. Isso leva à consideração de como se deve definir ou compreender o conceito “expressivo” no contexto da filosofia da psicologia de Wittgenstein, ainda que, quando apresentou a distinção, não tenha fornecido uma definição específica, mas apresentado casos em que os conceitos são usados de forma que se possa qualificá-

los como expressivos. Uma das principais características do uso expressivo dos conceitos é a de que não são usados a partir exclusivamente de seus significados, na medida em que suas situações de emprego, bem como os objetivos de seu uso são relevantes para a compreensão dos mesmos.

Assim, por exemplo, a afirmação “Eu sinto dores” possui certas condições de uso, as quais incluem o comportamento que a pessoa que a profere apresenta após enunciá-la, a situação em que é anunciada, e também o que a pessoa a quem se endereça poderá fazer. Na seção 310 das *Investigações Filosóficas*, Wittgenstein apresenta o caso de dizer para outrem que sente dores e descreve as possíveis atitudes que tal pessoa poderá ter: credulidade ou incredulidade, confiança ou desconfiança. Estas atitudes não são o resultado de operações intelectuais, mas elas próprias são a expressão da atitude da pessoa ou, como comenta Wittgenstein: “Sua atitude é a prova de sua atitude”. Isso significa que se tem de compreender uma atitude como sendo algo mais do que simplesmente uma manifestação física. A atitude é a expressão de como aquela pessoa considera a dor que está comunicando que tem. Nessa perspectiva, a expressão também pode ser considerada como parte de um gesto, uma “atitude para com” e, da mesma forma, incluir aspectos do que se chama de “semblante”.

4.

Como já apontado anteriormente, Wittgenstein tentou elaborar algumas classificações dos usos dos conceitos psicológicos com o intuito de encontrar alguma característica que os distinga. Após algumas tentativas, ele abandona a tarefa e concentra suas análises em dois tipos: conceitos psicológicos são expressivos no uso em primeira pessoa e informativos no uso de terceira pessoa. Contudo, não se deve considerar que ao qualificar uma sentença ou conceito como “expressivo” já se disse tudo sobre o mesmo, posto que isto apenas implicaria numa qualificação de tipos de conceitos. Tal classificação não diz como os mesmos devem ser usados. Quanto a esse raciocínio, considere-se o seguinte exemplo: a sentença “Eu tenho a intenção de ler o livro N”. Se a qualificamos como “expressiva”, a retiramos do jogo de linguagem para com a verdade ou falsidade, isto é, ela não visa descrever fatos. No entanto, sendo ela expressiva, isso implica que se deve considerá-la em sua relação para com as ações que estão interligadas no seu uso, ou se pode dizer, na sua situação de proferimento. Assim a sentença, a princípio, por sua forma, diz algo sobre uma intenção, mas, sendo ela expressiva, não se trata de um relato de uma determinada intenção, a qual se tem no momento em que é enunciada, isto é, não se trata de descrever a intenção. Antes, a própria sentença deve ser considerada como a expressão da nossa intenção.

Esse raciocínio remete ao que, mais acima, Wittgenstein disse quanto à atitude diante de quem “relata” suas dores: “sua atitude é a prova de sua atitude” (IF, §310). Existe aqui algo que se pode denominar como “relação interna” entre a sentença que expressa a intenção e a (própria) intenção, da mesma forma que há uma relação interna entre a atitude diante de quem relata ter dores (a pessoa que expressa compaixão) e a prova da mesma (isto é, que ela sente compaixão). Isso, por sua vez, leva a considerar que enunciados quanto ao mundo interior estão relacionados com nossas ações para com as mesmas. A partir dessa afirmação, da correlação entre o proferimento expressivo e a ação, pode-se ponderar que a classificação dos conceitos como expressivos ou descritivos ainda não seria suficiente, pois não explicaria como tais conceitos são realmente usados.

A fim de esclarecer um pouco mais o raciocínio apresentado, retome-se aquele caso exemplificado pela palavra “intenção”: pode ser o caso de que proferir aquela sentença não implique que, de fato, se irá ler aquele livro, mas se tenta causar algo na pessoa para quem é dirigida; suponha que a sentença foi direcionada para uma pessoa que não aprecia aquele livro ou aquele tipo de literatura; pode ser também que seja uma admissão de concordância como se fosse dito “Estou com você! Vou ler tal livro”, entre outras possibilidades. Ou seja, ainda que o uso em primeira pessoa (tenho a intenção de...) seja classificado como expressivo e que, na terceira pessoa, a sentença seja informativa, este passo classificatório ainda é inicial, no que diz respeito ao uso desses tipos.

Não se trata de minimizar a validade ou importância da concepção de uso expressivo de conceitos e sentenças, mas, sim, de indicar que tal aspecto diz respeito apenas a uma classificação dos mesmos. É preciso considerar ainda que as aplicações ou usos desses conceitos sofrem a influência da imagem aplicada ao mundo interior e que é sugerida pela própria linguagem quanto à existência de dois âmbitos no ser humano, a saber, o âmbito interno e o âmbito externo. Logo, classificar o conceito *intenção* como um conceito expressivo, quando usado na primeira pessoa do singular, e, por outro lado, informativo, quando usado na terceira pessoa, é dizer algo sobre as palavras. Talvez, por tal razão, Wittgenstein tenha inserido, no final da passagem citada mais acima, as palavras “não de todo correto”, abandonando, conseqüentemente, a busca por uma classificação dos conceitos psicológicos.

Certamente, neste ponto se tem por base uma interpretação dos argumentos de Wittgenstein a fim de lhes fornecer compreensibilidade. O mesmo tipo de atitude aparece no ensaio do filósofo Joachim Schulte, *Experience and Expression* (SCHULTE, 1995), que, ao tratar do tema, também considera que a falta de explicações para o abandono das tentativas de classificação dos conceitos psicológicos por parte de Wittgenstein se deve a sua insatisfação com a mera classificação, porquanto isto ainda não diz o que há

de essencial no seu emprego (SCHULTE, 1995, p. 33). A classificação não é nada mais que um meio para atingir a clareza quanto ao possível critério que se necessita e, nesse sentido, não tem por objetivo outra coisa que não servir como princípio orientador para outro caminho na investigação. Não é um fim em si, mas um meio que contribui para a descoberta de questões úteis e para nossa decisão quanto à validade de possíveis respostas. Ou seja, é necessário considerar que o uso expressivo dos conceitos psicológicos e seu uso informativo, na terceira pessoa, devem ser analisados em conjunto com as circunstâncias de seu emprego, as quais incluem atitudes e comportamentos das pessoas que os enunciam nas mesmas. Isso quer dizer que os conceitos são usados em situações de emprego e não como se fossem pronunciados num vazio de ações.

5.

Ao chegar a este ponto e considerando que existe relação entre uso expressivo e ação, alguém poderia argumentar o seguinte: ora, todas as tentativas de classificação foram abandonadas e, mais ainda, o próprio uso expressivo é tomado como uma classificação de palavras, pois faltaria sua relação para com as ações com que estão relacionadas. Portanto, não seria isso incidir num certo tipo de comportamentalismo? Não estaria Wittgenstein alegando que o importante, e relevante, para a compreensão dos conceitos psicológicos é o que a pessoa faz, como ela age quando alega ter dor ou que tem uma determinada intenção? Se as atitudes das pessoas, quando usam sentenças contendo conceitos psicológicos, é um elemento relevante para o significado do que dizem, não é isto eleger o comportamento como critério do âmbito psicológico? Ora, em momento algum Wittgenstein alega ser o comportamento de uma pessoa, ao usar conceitos psicológicos, o critério distintivo para o que se passa nela. Mas, ao mesmo tempo, não retira a relevância das atitudes e comportamentos. Isso significa dizer que o comportamento, as atitudes da pessoa, em conexão com o que diz, é um sinal para se compreender o que se passa com ela ou, também, para o que fazer em relação ao que ela diz sobre si mesma.

Nesse ponto, cabe lembrar as palavras de Wittgenstein na citação mais acima, quando responde ao seu interlocutor sobre a desconfiança de que, para ele, apenas o comportamento é relevante, isto é, o interlocutor desafia Wittgenstein a uma contraposição cujos termos não são explicados: comportamento em comparação com estados mentais. Ao que, especificamente, se está contrapondo o comportamento? Certamente que a processos e eventos psicológicos que não são comportamentais. No entanto, nessa oposição não foi oferecida uma definição do que sejam tais eventos e processos.

Surge novamente aqui a aplicação da imagem “interno - externo” que obriga ou impele a buscar por contrastes a diferença entre mundo interior e comportamento. As palavras iniciais da seção apresentam a influência dessa imagem no raciocínio de Wittgenstein: “Como é que se chega ao problema filosófico dos processos e estados psíquicos e o behaviorismo?” (IF, §308). Os primeiros passos são imperceptíveis, diz ele, “usamos os conceitos mas não temos claro para nós a que devemos aplicá-los”. Na verdade, o comportamento da outra pessoa, em relação ao que ela diz quanto a seus estados interiores, é um caminho para compreender tanto o que diz sobre si própria quanto a expressão do que ocorre nela. A imagem da dicotomia “interno - externo” sugere que aquilo que se pode observar, seu comportamento, se contrapõe a algo do mesmo tipo, isto é, ocorrências que não se consegue observar. Estes eventos estão escondidos.

Não obstante, por qual razão se deseja que sejam observáveis? Qual situação será resolvida se se conseguir observá-los? Por exemplo, se alguém diz que está triste ou tem tal e tal intenção, por qual razão saber se ele, de fato, tem aquela intenção acrescentaria algo ao que se pode fazer a partir do que ela nos diz? Será a certeza quanto ao que se passa nesta pessoa o que fará alguém agir a partir do que diz? Essas mesmas observações ocorrem a Wittgenstein nos escritos de *LWPPii* (p. 21-22) quando ele se pergunta sobre o significado ou a importância da afirmação “Tenho certeza de que ele tem dor”. Como se usa esta expressão? Qual a expressão de certeza no comportamento que permite que se esteja certo de que a pessoa tem dor? Ou seja, o comportamento da pessoa será o caminho, por assim dizer, para compreendê-la, mas, ao mesmo tempo, não é este o objetivo pelo qual é importante saber que ela tem dor e, sim, o que se fará a partir do que se compreende do que esta pessoa apresenta. Logo, o importante é saber o que significa, ou qual a relevância do que ela diz, e não o estado de conhecimento quanto ao que se pode saber sobre o mundo interior desta pessoa.

Ao comentar esse aspecto da argumentação de Wittgenstein, Schulte (1995, p. 50) lembra que Wittgenstein usa o conceito “expressão” em circunstâncias, num primeiro momento, bastante díspares: a expressão do comportamento e a expressão estética, tal como compreender uma passagem musical. Ou seja, assim como se tem de compreender a expressão da dor, a expressão de uma sensação, também se tem de compreender a expressão de uma passagem musical a fim de explicá-la para outra pessoa. Ora, ao ensinar alguém a compreender certa passagem musical, também se ensina algumas técnicas musicais, além de outras técnicas que ela já deve possuir. Mais ainda, como se lhe ensinará a elaborar uma certa avaliação ou interpretação daquela passagem, as palavras que serão usadas tendem a ser expressivas; se usará de certas comparações com situações, se pedirá que a pessoa compare a passagem com certo aspecto de uma paisagem ou com uma cor, por exemplo. E

o que tal pessoa adquirirá não será, especificamente, conhecimento, mas uma nova forma de ver aquela passagem em específico.

Neste viés, pode-se ter um sistema para lhe ensinar, mas este sistema não obedece a regras estritas e nem se apoia em hipóteses. Os conceitos, nestes casos, serão usados de maneira metafórica, mesmo que se lhe solicite que compare certo tom da melodia com algum aspecto de uma paisagem ou de uma pintura. Ao fim e ao cabo, não se falará de uma experiência específica, tal como uma vivência, mas de aspectos e nuances de algumas experiências que ajudarão a fazê-la compreender a passagem musical. Segundo argumenta Schulte (1995, p. 50 e segs), a lição que se deve retirar dessas análises de Wittgenstein é que qualquer aspecto que pareça específico nas experiências de compreensão da passagem musical, compreensão da expressão da dor e da intenção, terão de basear-se num jogo de linguagem relevante para a situação.

Ao mesmo tempo, a alegação de que não há objetividade nesses casos de compreensão baseia-se em algo ainda não esclarecido, pois o que leva a questionar se uma expressão é correta ou incorreta, apropriada ou não, depende de quais regras permitem construir julgamentos para estes casos. Por exemplo, assim como é possível não compreender certa passagem musical, também é possível não compreender certa expressão de dúvida ou de insegurança nas atitudes de alguém. Tal incompreensão tem base na falta de algum tipo de conhecimento objetivo atuando na situação? Mas que tipo de consideração “objetiva” poderia auxiliar ou ensinar a compreender corretamente nestes casos? Por exemplo, aprender teoria musical, para o caso da passagem musical, e alguma informação sobre comportamento, em outro? Mas se trata de algo específico que não é compreendido ou não se conhece? Aqui, argumenta Schulte (1995, p. 53), “tudo depende do jogo de linguagem e não de alguma definição essencialista, que não conheço, sobre algum tipo de fato específico”.

A partir de tudo que se expôs até agora, a filosofia da psicologia de Wittgenstein não visa descobrir aspectos ou, melhor, conhecimentos específicos quanto ao que seja o mundo mental, para só então explicar como se deve usar os conceitos psicológicos. A tarefa é, antes, descrever os usos que se faz destes conceitos nos jogos de linguagem em que aparecem ou em que são usados. Não é necessário que alguma ciência do cérebro ou neurociência forneça uma teoria com base na qual estes conceitos tenham seus significados explicitados, já que eles são efetivamente usados por todas as pessoas nas interações cotidianas com os mesmos. A incompreensão quanto ao que se passa no interior de outra pessoa é algo que interessa tanto quanto ter acesso ao seu diário pessoal, o qual ela, por motivos particulares, escondeu da visão pública.

Nessa perspectiva, a perplexidade não diz respeito à falta de conhecimento de algo específico ou essencial, mas às circunstâncias de sua atitude. A ideia de que aquela pessoa “esconde algo” – seus sentimentos, emoções, pensamentos e intenções – é algo que ocorre em certas situações. É possível descrever tais circunstâncias?, pergunta Wittgenstein. E se não se pode descrever, por quais motivos? Ora, é possível fornecer descrições mais ou menos claras destas circunstâncias nas quais se confia ou desconfia que se sabe que aquela pessoa está triste, por exemplo. “Alguém me ensinou a fazer isto”, pergunta o interlocutor de Wittgenstein. “Sim”, responde ele, “mas não se trata de ensinar a reconhecer algo específico, mas de ensinar a observar os sinais, tal como observamos os sinais externos da pretensa tristeza” (LWPPi, §971 - §972).

O resultado dessas observações não pode ser sistematizado como uma lista completa, pois é, desde o início, uma lista incompleta. Por exemplo, uma lista das possíveis atitudes frente a relatos de intenção, relatos de dor, entre outros. Ora, quem estaria interessado nesse tipo de lista, uma vez que se assume desde o princípio como incompleta? Teria tal lista algum uso prático? A resposta de Wittgenstein a estas questões é que “este jogo não funciona desta forma”. Não há algo faltando, pois nada está escondido, e se, por acaso, estivesse escondido e se isso fosse assumido desde o início, então não interessaria para o caso. Assustado com essa conclusão, o interlocutor alega: “posso esconder meus pensamentos de outra pessoa”. Ao que Wittgenstein responde: “Sim! Da mesma forma como pode esconder seu diário, e o fato de eu procurá-lo origina-se no interesse que tenho pelo que nele está escrito” (LWPPi, §972). Nesse sentido, “seus pensamentos estão escondidos para mim, é um pleonasma” (LWPPi, §975).

Por fim, a concepção de que se diz algo para si mesmo e que outra pessoa não tem acesso ou que está escondido dela, ou que é uma questão de processos mentais que ocorrem num âmbito privado interno, não resolve nada, pois, neste caso, também será um processo físico que se pode conceber como “escondido”. Considera-se que pensar em silêncio para si próprio está escondido do outro, mas isto pode apenas significar que ele não pode “adivinhar” o que seja o caso, por esta ou aquela razão. Isso não significa, todavia, que seja imperceptível para outra pessoa por estarem “escondidos em minha alma”. O exterior não deve ser considerado necessariamente como uma fachada (Wittgenstein usa aqui a palavra “façade”) por detrás da qual forças mentais operam. Ou, tal como escreveu Wittgenstein numa variante deste raciocínio, “mas você não tem de pensar desta forma. Se uma pessoa fala comigo, claramente sem esconder nada, então nem sempre sou tentado a pensar desta forma” (LWPPi, §978 e nota). Nesse sentido, o mundo interior se expressa no comportamento, ainda que não se possa, sem cometer enganos, afirmar que o mundo interior é o comportamento. Mesmo porque o comportamento humano é imprevisível.

E como seria se fosse previsível? “Como devemos imaginar este caso? Que ligações deveríamos pressupor?” (Zettel, §603 - §604; RPPii, §663 - §664).

Conclusão

Na filosofia da psicologia de Wittgenstein, não há uma busca por tentar estabelecer algum tipo de visão geral teórica sobre o mundo mental e nem uma tentativa inacabada de classificar os conceitos psicológicos em tipos, a saber: expressivos ou descritivos. O que Wittgenstein fornece é uma consideração quanto a vários aspectos do uso destes conceitos, indica o quanto a concepção de que o comportamento é apenas uma fachada, leva a postular um âmbito para lá das atitudes e que – conforme o ponto de vista – seria a origem do comportamento. A filosofia da psicologia é a tentativa de compreender a gramática dos conceitos ordinários sem que sejam expressão de teorias inacabadas e primitivas.

Certamente que o leitor que até aqui acompanhou este ensaio não se sentirá satisfeito se sua busca é por uma explicação teórica dos conceitos psicológicos ou uma teoria filosófica da psicologia. O que Wittgenstein nos oferece como investigação não pode ser classificado dentro dos critérios tradicionais, isto é, Dualismo, Materialismo ou Comportamentalismo. Sua investigação é essencialmente conceitual e seus conceitos devem ser compreendidos dentro dos jogos de linguagem em que são usados. Tais jogos consideram as circunstâncias de uso, as quais também são as circunstâncias da vida que levam alguém a ter confiança na crença de que a outra pessoa está triste, tal como afirma, e que, por outro lado, conduzem esse alguém a desconfiar de suas palavras e atitudes. Ou seja, os conceitos psicológicos são considerados dentro do fluxo da vida ou, como Wittgenstein afirmou nas *Investigações Filosóficas* (IF, §90a),

é como se tivéssemos de penetrar os fenômenos, mas nossa investigação não se dirige aos fenômenos, e sim, como poderia dizer, às possibilidades dos fenômenos. Isto quer dizer que meditamos sobre a “espécie de asserções” que fazemos sobre os fenômenos.

Essas asserções fazem parte do fluxo da vida tanto quanto os fenômenos que causam perplexidade e instigam a investigar. Acreditar que tais asserções são partes de teorias primitivas ou que apenas poderão ser claramente compreendidas se alguma teoria fornecer fundamentos unificados para o que seja o mundo interior será claramente a busca por especulações metafísicas que apenas confundirão as fronteiras entre aquilo que é gramatical e aquilo que é empírico.

Referências

Schulte, J. **Experience and expression**. Wittgenstein's philosophy of psychology, Oxford, Clarendon, 1995.

Wittgenstein, L. **Zettel**, Berkeley, Univ. of California Press, 1970.

Wittgenstein, L. **Remarks on the philosophy of psychology**, Vol. 1, Chicago, Chicago Univ. Press, 1980a.

Wittgenstein, L. **Remarks on the philosophy of psychology**, Vol. 2, Chicago, Chicago Univ. Press, 1980b.

Wittgenstein, L. **Investigações filosóficas**, São Paulo, Nova Cultural, 1980c.

Wittgenstein, L. **Last writings on the philosophy of psychology**, Vol. 1, London, Blackwell, 1982.

Wittgenstein, L. **Last writings on the philosophy of psychology**, Vol. 2, London, Blackwell, 1992.





Dinos Alegria

LINGUAGEM E MENTE: O PROBLEMA DA INTELIGÊNCIA ARTIFICIAL DE UM PONTO DE VISTA WITTGENSTEINIANO



Mirian Donat¹

Resumo:

A ideia de máquinas que possam desenvolver algum tipo de inteligência, assim como a inteligência humana, sempre intrigou a humanidade. Desde a criação dos modernos computadores essa possibilidade parece ser apenas uma questão de tempo e de desenvolvimento científico-tecnológico. Num sentido comum, entende-se que a inteligência artificial será capaz de falar e pensar tal como um humano e também interagir com humanos de tal forma que não seria mais possível distinguir homens e máquinas. A referência clássica dessa possibilidade é o teste de Turing, em que um julgador humano deve questionar independentemente um computador e outro humano. Se ao final do questionamento o julgador não for capaz de dizer qual é o computador e qual é o humano, então a máquina passou no teste e somos levados à conclusão de que, se não é possível diferenciar o comportamento da máquina do

Abstract:

The idea of machines capable of developing some form of intelligence, just like human intelligence, has always intrigued humanity. Since the creation of modern computers this possibility seems to be only a matter of time and scientific-technological development. In a common sense, it is understood that artificial intelligence will be able to speak, think like a human and also interact with humans in such a way that it would no longer be possible to distinguish between humans and machines. The classic reference to this possibility is the Turing test, in which a human judge must question a computer and another human independently. If at the end of the questioning the judge is not able to say which is the computer and which is the human, then the machine has passed the test and we are led to the conclusion that if it is not possible to differentiate between the behavior of the machine and the human, then the ma-

¹ Professora do Departamento de Filosofia e do Programa de Pós-graduação em Filosofia da Universidade Estadual de Londrina. E-mail: donat@uel.br

comportamento humano, então a máquina deverá ser considerada inteligente. Os atuais dispositivos de conversação, *chatbots* tais como ChatGPT e outros, trazem em si as questões levantadas pelo teste de Turing. Tais dispositivos devem interagir em conversações com humanos e responder às demandas de forma a simular a linguagem natural humana. A questão que se coloca é se tais máquinas realmente têm uma linguagem tal que não se poderia distinguir entre essa linguagem e a linguagem humana e se por tal característica deveríamos também aceitar que não seria mais possível distinguir homem e máquina. Para a análise desse problema, desenvolvemos o artigo em dois momentos: em primeiro lugar, avaliamos o que significa “ter uma linguagem”, partindo da noção fundamental de Wittgenstein de que toda linguagem se desenvolve e institui a partir de práticas humanas no mundo e, em segundo, analisamos a noção wittgensteiniana de “ser humano”, que apresenta uma crítica radical às concepções tradicionais, em especial o corte realizado entre o “corpo” e a “mente”, afirmando a unidade fundamental do ser humano.

Palavras-chave:

Inteligência artificial; linguagem; mente; corpo; jogos de linguagem; formas de vida.

chine should be considered intelligent. Current conversational devices, chatbots such as ChatGPT and others, bring up the issues raised by the Turing test. Such devices must interact in conversations with humans and respond to demands in a way that simulates human natural language. The question which arises is whether such machines truly have such a language that one could not distinguish from human language, and whether by this characteristic we should also accept that it would no longer be possible to distinguish between man and machine. For the analysis of this problem, we developed the article in two stages: firstly, we evaluate the meaning of “having a language”, starting from the fundamental notion of Wittgenstein that all language develops and establishes through human practices in the world; and secondly, we examine the Wittgensteinian notion of “human being”, which presents a radical critique of traditional conceptions, in particular the distinction made between “body” and “mind”, asserting the fundamental unity of the human being.

Keywords:

Artificial intelligence; language; mind; body; language games; forms of life.

I.

Desde a criação dos modernos computadores, a questão acerca da possibilidade da inteligência artificial tem se mostrado cada vez mais instigante. Num sentido comum, é a ideia de que a inteligência artificial será tanto capaz de pensar tal como um humano como interagir com humanos de tal forma que não seria possível distinguir homens e máquinas. Essa noção está assentada em um ponto de vista que enfatiza as analogias entre o funcionamento do cérebro e o funcionamento de computadores e que, assim, “o cérebro é somente um computador digital e a mente é somente um programa de computador”, nas palavras de John Searle (1994, p. 33), que chamou tal ponto de vista de IA forte, segundo a qual “a mente é para o cérebro o que o programa é para o *hardware* do computador”.

Como consequência desse ponto de vista temos a tese de que ter uma mente não depende de ter um cérebro (orgânico), uma vez que qualquer sistema físico pode ter uma mente desde que tenha o programa de regras com os *inputs* e *outputs* corretos. Ou seja, a mente é compreendida como um conjunto de funções que são realizadas por sistemas de regras em redes, sistemas esses que podem ser circuitos orgânicos, tais como o cérebro humano, ou circuitos de silício, por exemplo. O que importa é que tais sistemas tenham as configurações corretas para realizar as funções previstas. Se o resultado de um circuito e outro, em termos das funções realizadas, não puder ser distinguido, então podemos afirmar que temos uma inteligência artificial.

A referência clássica dessa possibilidade é o teste de Turing, em que um julgador humano deve questionar independentemente um computador e outro humano. Se ao final do questionamento o julgador não for capaz de dizer qual é o computador e qual é o humano, então a máquina passou no teste e somos levados à conclusão de que, se não é possível diferenciar o comportamento da máquina do comportamento humano, então a máquina deverá ser considerada inteligente.

Os atuais dispositivos de conversação, *chatbots* tais como ChatGPT e outros, trazem em si as questões levantadas pelo teste de Turing. Tais dispositivos devem interagir em conversações com humanos e responder às demandas de forma a simular a linguagem natural humana. Os resultados alcançados têm sido de tal magnitude que, muitas vezes, parecem interagir como se de fato falassem e pensassem. Podemos fazer perguntas e receberemos respostas, escritas ou faladas, sobre assuntos infinitamente diversos. Somos assim levados a pensar que tais máquinas realmente têm uma linguagem, e que, em vista disso, também teriam uma mente e uma inteligência que podem ser comparadas com a inteligência humana.

Em certas concepções filosóficas e antropológicas se acreditava que a linguagem seria uma característica eminentemente humana, que seríamos humanos por termos uma linguagem e, com a linguagem, a capacidade de representar o mundo, pensar e refletir sobre ele, enfim, ter consciência. Mas as nossas atuais máquinas parecem nos fazer acreditar que tal privilégio já não é mais tão aceitável, pois elas em tudo demonstram ter uma linguagem como a nossa, sendo, portanto, seres de linguagem “tal como nós”. Assim, volta a questão acerca do que é afinal que nos possibilitaria (ou não) distinguir o humano das máquinas.

O problema se coloca como uma pergunta acerca do que nos permite – ou não – diferenciar a linguagem humana e a linguagem das máquinas e se, em decorrência disso, poderíamos também encontrar uma forma de diferenciar – ou não – homem e máquina no que diz respeito à presença de processos mentais. Tendo como base a filosofia da linguagem do segundo Wittgenstein, vamos desenvolver a análise desse problema em dois momentos. Em primeiro lugar, avaliamos o que significa, para Wittgenstein, ter uma linguagem, o que será feito a partir da sua concepção de significado como uso da linguagem, a qual se fundamenta na noção de que toda linguagem é resultado das práticas humanas no mundo. Será preciso avaliar, com base nisso, se esta concepção se adequaria de alguma maneira aos pressupostos de que a máquina pode desenvolver uma linguagem tal como a linguagem humana. Depois disso, num segundo momento, para avaliar se máquinas poderiam realizar processos mentais, discutimos a noção wittgensteiniana de “ser humano”, através da qual faz duras críticas aos modelos filosóficos e psicológicos que separam mente e corpo, defendendo em seu lugar que o “humano” é uma unidade fundamental, além de enfatizar que esse conceito se baseia numa certa atitude que temos para com esse ser, o que ele chamou de “atitude para uma alma”.

II.

A ideia de que computadores ou outras máquinas possam falar e pensar está ligada ao desenvolvimento cada vez maior da linguagem destas máquinas. Atualmente, temos máquinas e dispositivos que efetivamente parecem dialogar conosco, como se pensassem e refletissem sobre o que nos dizem, deixando a impressão de algo como uma consciência por detrás de seu comportamento.

Ficamos cada vez mais envolvidos com essas tecnologias e passamos a acreditar que se, até o momento, não podemos provar que são capazes de algum tipo de processos

mentais, é só uma questão de tempo para que isso aconteça. Isso nos impele a colocar questões tais como o que é, afinal, ter uma mente ou processos mentais. Mas antes da pergunta sobre o que é ter mente e processos mentais, parece ser mais urgente responder à pergunta sobre o que é ter uma linguagem, pois é a partir do comportamento linguístico que se supõe a possibilidade de atribuição de estados mentais.

Os problemas relacionados com a linguagem e a significação, bem como o que significa ter uma linguagem, são questões que formam o núcleo fundamental da filosofia de Wittgenstein, desde seus primeiros escritos até suas últimas observações antes de sua morte. Para essa análise, tomo as *Investigações Filosóficas* como guia, obra maior de sua segunda fase. Com Wittgenstein, queremos mostrar que a linguagem não é o resultado de um processo de organização de sistemas de regras em um suporte (seja ele orgânico ou não), mas sim das práticas humanas e da interação entre sujeitos em uma forma de vida. A linguagem surge de um tipo de reação natural ao ambiente e aos outros sujeitos.

Na passagem das *Investigações Filosóficas* comumente conhecida como argumento da linguagem privada, encontramos a seguinte questão colocada por Wittgenstein: “Como uma pessoa aprende o significado dos nomes das sensações? Por exemplo, da palavra ‘dor’”, ao que ele mesmo responde:

Eis uma possibilidade: Palavras são conectadas à expressão originária, natural, da sensação, e postas em seu lugar. Uma criança se machuca e grita: e então os adultos falam com ela e lhe ensinam exclamações e, mais tarde, frases. Eles ensinam à criança um novo comportamento de dor (WITTGENSTEIN, 2022, § 244).

À primeira vista pode parecer que a passagem estaria tratando de uma nova temática no conjunto do livro – uma discussão acerca da mente e suas características e propriedades –, mas na verdade o que Wittgenstein demonstra com essa argumentação é que toda a linguagem tem a mesma raiz, até mesmo a linguagem que usamos para falar acerca dos nossos estados mentais. As questões relativas ao que são estados mentais ou o que seria ter consciência parecem como se tivessem um estatuto próprio e que descobrindo algo sobre tais coisas saberíamos também diferenciar seres que têm consciência de seres que não a têm ou estabelecer quando uma máquina é capaz de pensar. Pareceria como se fosse uma questão empírica acerca do que é ter estados mentais e consciência, o que resolveria de uma vez o problema.

Entretanto, a perspectiva de Wittgenstein se opõe fortemente a essa ideia. Para ele, a questão não é empírica, mas sim conceitual, é uma questão acerca do modo como usamos nossos conceitos, no caso nossos conceitos psicológicos. E a (dis)solução do problema envolve uma consideração acerca dos modos de uso dos conceitos, voltando-se para como eles surgem e como se desenvolvem e se instituem em um jogo de linguagem específico.

A linguagem não é um sistema de regras ou símbolos localizados no cérebro dos falantes, sistema tal que ordenaria os significados das palavras a partir de algum tipo de ligação da palavra com o objeto em questão. Se isto vale para a linguagem psicológica é porque está de acordo com a concepção desenvolvida por Wittgenstein ao longo das *Investigações*, sendo o argumento da linguagem privada uma demonstração de que também a linguagem psicológica obedece aos mesmos princípios de qualquer outro jogo de linguagem, visto que ela está enraizada em nossas práticas no mundo.

Dois pontos que aparecem na argumentação sobre a linguagem privada são fundamentais para a compreensão não só da linguagem psicológica, mas também da totalidade da linguagem. O primeiro ponto é a crítica de Wittgenstein às concepções referencialistas da significação, que tomam como base para a significação a relação entre palavra e objeto, tendo a definição ostensiva como o ato que liga a palavra e o objeto e institui a significação. Ora, Wittgenstein mostra que o significado da linguagem psicológica não se baseia em uma relação da palavra com o objeto, sendo mesmo tal relação um impedimento para a própria possibilidade desse jogo de linguagem. O exemplo do besouro na caixa deixa claro que, quando tomamos a experiência subjetiva como a referência para o significado da palavra, o que acontece é que o próprio “objeto desaparece, como irrelevante, de nossas considerações” (WITTGENSTEIN, 2022, § 293).

A crítica de Wittgenstein à concepção referencialista da significação abre as *Investigações Filosóficas* com a passagem de Agostinho acerca do modo como aprendeu a linguagem na infância. Para Wittgenstein, discutindo a passagem de Agostinho, é um erro pensar que o significado da palavra seja o objeto a ela correlacionado, pois com as nossas palavras fazemos muito mais coisas do que nomear os objetos no mundo. Nomear objetos é uma das funções da linguagem, mas longe de ser a essência de toda a linguagem, não passa de mais uma das inúmeras funções que a nossa linguagem pode desempenhar:

Existem *incontáveis* tipos: incontáveis tipos diferentes de emprego de tudo aquilo que chamamos de “sinais”, “palavras”, “frases”. E essa multiplicidade não é nada de fixo, dado de uma vez por todas; mas novos tipos de linguagem, novos jogos de linguagem, como podemos dizer, surgem e outros envelhecem e são esquecidos (WITTGENSTEIN, 2022, § 23).

Com a noção de jogo de linguagem, Wittgenstein demonstra que a linguagem está ligada às ações e práticas humanas no mundo. Não é possível pensar uma linguagem que esteja desconectada de alguma prática, que não esteja ligada ao agir humano. Este é o sentido da ideia de que o significado é o uso das palavras em determinados contextos, contextos práticos nos quais se usa a linguagem para realizar certas ações, cada palavra ocupando as funções que lhe cabem dentro daquele jogo de linguagem específico. As palavras são como as ferramentas que usamos para fazer alguma coisa, e cada ferramenta tem sua função de acordo com aquilo que precisa desempenhar numa situação particular, do mesmo modo que cada palavra tem sua função dada por aquilo que pretendemos fazer com ela.

Sendo assim, precisamos nos voltar para o modo como cada jogo de linguagem se constitui e institui o significado das palavras. Sendo conectado às práticas humanas no mundo, cada jogo de linguagem é o resultado de um longo desenvolvimento que resulta na organização linguística de um certo campo da ação humana. Essa possibilidade se assenta no fato de que os humanos têm reações e comportamentos primitivos que são compartilhados:

O tipo de reação da qual a linguagem pode se desenvolver deve ser o comportamento *compartilhado* ou instintivo da humanidade; reações como chorar quando está com dor ou triste; sorrir quando se está feliz; pular quando se está assustado; ofegar ou gritar quando se está com medo; mas também reagir ao sofrimento de alguém (MOYAL-SHARROCK, 2018, p. 93).

De fato, em uma passagem de *Cultura e valor* encontramos a seguinte consideração:

A vida pode educar-nos para a fé em Deus. E são também as *experiências* que o fazem; mas não me refiro a visões e a outras formas de experiência dos sentidos que nos mostram a “existência desse ser”, mas, por exemplo, a sofrimentos de vários tipos. Estes nem nos revelam Deus do mesmo modo que a experiência dos sentidos nos revela um objeto, nem dão origem a quaisquer conjecturas a seu respeito. As experiências, os pensamentos – a vida pode impor-nos este conceito (WITTGENSTEIN, 2000, p. 125).

Conceitos tais como “Deus” nasceriam de uma reação diante de sofrimentos de vários tipos. É o resultado de uma tentativa de encontrar algum sentido para o que nos cerca e nos causa tais sofrimentos. Essa é uma forma tipicamente humana de relação com o mundo e com aquilo que nos rodeia. A realidade nos afeta de tal forma que não ficamos passivos diante dela. Ao contrário, tentamos ativamente encontrar um sentido em meio ao caos que nos rodeia. A linguagem é a ferramenta que nos serve para organizar esse caos e encontrar nele alguma regularidade e sentido. Diante disso surgem os conceitos e os jogos de linguagem, cada um deles tentando de alguma maneira organizar uma parte ou setor da nossa experiência no mundo.

Essas reações e comportamentos primitivos vão sendo aos poucos organizados em uma forma regular de comportamento que, com o auxílio da linguagem, torna-se um padrão de referência para aqueles que compartilham daquele jogo de linguagem. Esse padrão de referência, por sua vez, constitui o conjunto das regras do jogo de linguagem, sendo então tomadas como padrões de correção para o falar e o agir.

Importante perceber que há um comportamento natural mediado por reações e gestos que aos poucos vai sendo substituído por palavras, instituindo um novo tipo de comportamento, agora mediado pela linguagem. Esse enraizamento da linguagem em formas primitivas de comportamento é fundamental para Wittgenstein, pois a linguagem é, para os humanos, algo que faz parte de sua história natural: “Dar ordens, perguntar, contar histórias, jogar conversa fora pertencem à nossa história natural assim como andar, comer, beber, jogar” (WITTGENSTEIN, 2022, § 25). A linguagem não nasce de algum tipo de raciocínio, mas de algo “animal”, este animal que somos os humanos. Sendo assim, a linguagem é um tipo de reação humana natural ao seu meio ambiente, ao mundo e aos outros falantes com quem interage.

Nessa nova concepção de significado como uso, o aprendizado da linguagem ocupa um lugar central. Aprender o significado de uma palavra não é realizar a ligação dela com algum objeto do qual seria o nome, mas sim aprender a tomar parte nas diferentes práticas que estão envolvidas nos jogos de linguagem. Esse aprendizado revela a íntima ligação da linguagem com as práticas, pois uma criança, quando aprende a linguagem, aprende a tomar parte nas práticas das pessoas com quem interage. E a criança é capaz de aprender porque ela mesma compartilha com os outros as reações e comportamentos primitivos que levaram ao desenvolvimento dos jogos de linguagem.

Aos poucos as reações naturais e primitivas vão sendo substituídas pelas expressões linguísticas. O modo como isso acontece é como um adestramento, em que a criança vai sendo treinada para fazer algo de uma forma diferente do que fazia. O que se aprende quando se aprende um novo jogo de linguagem são novas técnicas para o uso das palavras, é o desenvolvimento de capacidades que nos permitem realizar novas práticas e participar de novas situações. Aprende-se a manejar as palavras, tais como aprende-se a manejar ferramentas, com os mais diversos propósitos e objetivos.

Para Wittgenstein, as regras que instituem o significado da linguagem e da ação humana estão dadas num espaço compartilhado de uma forma de vida. Isso significa que para uma criança aprender a linguagem e aprender a agir como os adultos, ela precisa ser adestrada nesse sistema regrado, de tal maneira que aprende a reagir como se espera que reaja em cada situação. Peter Winch (2020, p. 93 – Grifos do autor) mostra de forma interessante que a criança precisa aprender também o que significa agir de acordo com regras:

O ponto aqui é que *importa* que o aluno reaja ao exemplo do professor de uma forma e não de outra. Ele tem que adquirir não apenas o hábito de seguir o exemplo de seu professor, mas também entender que algumas formas de seguir aquele exemplo são permitidas e outras não. Isto é, ele tem que adquirir a habilidade de aplicar um critério; tem que aprender não simplesmente a fazer coisas da mesma forma que seu professor, mas também *o que conta* como a mesma forma.

Ou seja, a criança não nasce com um sistema de regras universais que vai desenvolvendo com seu crescimento. Ela precisa aprender até o que seja seguir regras, e só então poderá tomar parte em situações de linguagem e ação significativas.

III.

Ligadas à concepção de que a linguagem está enraizada nas práticas humanas estão as considerações de Wittgenstein em relação ao comportamento humano, o que o leva a uma particular concepção do que seja o próprio ser humano. Para compreender o sentido do humano, para Wittgenstein, é fundamental compreender as relações entre o comportamento humano e a linguagem.

Entretanto, comportamento humano, para Wittgenstein, não se reduz a meros movimentos corporais, mas remete à totalidade da ação humana no mundo, com todas as suas relações e o contexto em que acontecem. São as práticas compartilhadas em uma for-

ma de vida, aquilo que Wittgenstein chama de comportamento humano, pois este é moldado de acordo com a forma de vida na qual se desenvolve. É como se o comportamento, enquanto movimentos corporais, fosse aos poucos recebendo um sentido específico. É assim que se pode falar de comportamento de dor, pois a dor envolve certamente reações naturais daquele que tem dor; essas reações seriam algo como um comportamento natural diante da dor, como a criança que se machuca e chora ou grita. Mas esse comportamento natural vai aos poucos sendo substituído por um comportamento mediado por palavras, ele vai sendo organizado de tal maneira que temos agora um “novo comportamento diante da dor”.

Se a dor, enquanto sensação, é algo que pode ser descrito da mesma maneira para qualquer sujeito e para qualquer forma de vida, o sentido que a dor tem para cada sujeito e para cada forma de vida pode ser bastante distinto. Esses diferentes sentidos estão relacionados com os modos como, por exemplo, em cada forma de vida se considera adequada ou não certa reação diante da dor, se é aceito ou não provocar a dor em outras pessoas, e assim por diante.

Por outro lado, para que se possa dizer que alguém tem dores é preciso que este tenha um comportamento característico a partir do qual se possa atribuir uma sensação a ele. Para Wittgenstein (2022, § 281): “apenas do homem vivo, e do que lhe é semelhante (se comporta de maneira semelhante), pode-se dizer que tem sensações; que vê; que está cego; que ouve; que está surdo; que está consciente; ou inconsciente”.

Para Wittgenstein, apenas o ser humano poderia ter sensações, ver, ouvir, enfim ter consciência. Isso significa que apenas um ser que *se comporta* como ser humano é que poderia efetivamente ter consciência e, para Wittgenstein, isso se liga ao fato de que os humanos não são um corpo que tenha uma alma, ou seja, algo que estaria de alguma maneira ligado ao corpo e que efetivamente nos torna humanos. Não há tal separação, “pois o corpo humano é a melhor imagem da alma humana”.

Essa noção é fundamental para compreender o ponto de vista de Wittgenstein, visto que ele defende que um ser consciente não é um ser que possa meramente apresentar um comportamento consciente, mas sim um ser que é uma unidade de corpo e alma e, para a qual, temos uma certa atitude. Wittgenstein ataca com isso as concepções dualistas que separam o corpo e a mente, como se a mente fosse algo que é acrescido ao corpo e que explica porque, afinal, somos humanos. Sendo assim, poderíamos pensar que qualquer artefato que manifestasse comportamento consciente deveria ser tomado como uma criatura realmente consciente. Mas falta algo aqui, de acordo com Wittgenstein (2022, § 284), aquilo que podemos chamar de uma “atitude para uma alma”:

Da mesma maneira, também um cadáver nos parece totalmente inacessível à dor. – Não temos a mesma disposição em relação ao que vive e ao que está morto. Todas as nossas reações são diferentes.

As atitudes e reações que temos diante desse ser é o que demonstra a maneira pela qual o consideramos. Por um lado, essas atitudes e reações mostram que não podemos aceitar como consciente um ser que não tenha nenhuma forma de expressar em seu comportamento sua própria consciência. É quando Wittgenstein (2022, § 283) questiona se admitiríamos que uma pedra possa ter dores: “e será que se pode dizer, a respeito da pedra, que ela tem uma alma, e que *essa alma* tem dor? O que uma alma, o que é que ter dor, têm a ver com a pedra”. Ora, de algo como uma pedra não admitiríamos dizer que tenha dores ou que tenha consciência, pois ela não tem nenhuma forma de comportamento que possa expressar para nós que tenha dores e consciência, que faça com que tenhamos para com ela uma atitude *para uma alma*.

Por outro lado, não basta o comportamento para atribuímos sensações e consciência, sendo necessário ainda um comportamento que seja *como* aquele de um ser humano. E essa não é uma questão empírica, mas conceitual, diz respeito ao que consideramos que seja isso que chamamos “ser humano”. E para Wittgenstein é parte da nossa concepção do que seja humano que seja um “homem vivo”, ou seja, um ser que tenha um corpo tal como qualquer outro homem e que *seja vivo*. Ser vivo significa que é alguém que tem uma história, que nasceu em um certo momento, que tem um pai e uma mãe, que precisa se alimentar e ter satisfeitas outras necessidades vitais. É alguém que compartilha sua existência com outros seres iguais a ele, com os quais vai desenvolvendo formas de interação que têm em sua base a noção de que todos sejam “homens vivos”. Nessa forma de compreender o ser humano, não se separa corpo e mente, muito menos se investiga que tipo de relação poderia haver entre eles, pois não há aqui qualquer tipo de relação entre duas coisas distintas. Temos uma reação diante de uma pessoa, um ser humano, não uma de suas partes:

Mas será que não é absurdo dizer, a respeito de um *corpo*, que ele tem dor? – E por que sentimos aí algo absurdo? Em que medida não é minha mão que sente dor; mas eu em minha mão?

Que tipo de controvérsia é esta: Será que é o *corpo* que sente dor? – Como se pode decidi-la? O que é que faz prevalecer que *não* seja o corpo? – Bem, talvez o seguinte: Quando alguém tem dor na mão, a *mão* não diz isso (a não ser que escreva), e não consolamos a mão, mas a pessoa que está sofrendo: olhamos em seus olhos” (WITTGENSTEIN, 2022, § 286).

Ou seja, não temos uma atitude diante de um corpo que demonstra estar com dor, mas uma atitude para uma pessoa: “olhamos em seus olhos”. Nossa reação é para uma totalidade, uma unidade entre corpo e alma, não tomamos cada um como partes separadas, mas sim os dois simultaneamente como o que seja o ser humano. São esses tipos de reação diante de uma pessoa que está na base do nosso conceito de humano e que definem o tipo de atitude que temos para com ele:

O que tenho em mente aqui abrange uma ampla gama de fenômenos. Por exemplo, em certas circunstâncias, pensamos ser apropriado sentir pena dos outros e confortá-los - isto é, quando eles estão sofrendo. Sentimos gratidão ou ressentimento em relação aos outros quando eles agiram de determinada maneira; isto é, nos importamos com o modo como os outros nos tratam de uma maneira que não nos importa caracteristicamente com o comportamento de objetos inanimados e, talvez, da maioria dos animais (COCKBURN, 2001, p. 98).

Portanto, comportar-se como humano depende de que haja um corpo humano que atue, não é apenas demonstração de capacidades cognitivas e de linguagem. Essas observações fazem parte da “história natural” do ser humano, “fatos naturais extremamente gerais”, que muitas vezes não são considerados por sua generalidade, mas que são fundamentais para compreender a linguagem humana.

Em lugar de corpo e mente, a noção wittgensteiniana de ser humano não faz nenhuma forma de separação ou distinção, seja entre corpo e mente, interno e externo, ou o que for. Quando vemos uma pessoa, não vemos movimentos corporais dos quais deduzimos uma mente, ou movimentos corporais que não nos permitem afirmar que tal pessoa tenha realmente uma mente. O que vemos é um ser humano vivo no fluxo da vida, que age de tal forma que em sua ação expressa sua mente.

Portanto, na atribuição de estados mentais, efetivamente, precisamos assumir que aquela criatura realmente tem uma consciência por detrás do comportamento. Essa atribuição resulta das nossas reações naturais diante deste ser e que se mostra na nossa atitude diante dele. Raimond Gaita mostra como isso acontece também em nossas relações com os animais, como conta nessa bela passagem:

Quando os cães respondem a nossos humores, prazeres e medos, quando eles se antecipam às nossas intenções, ou esperam ansiosos para saber se vamos levá-los para um passeio, eles não supõem que somos seres sensíveis e com intenções. Imagino que tenha sido assim também conosco em estágios primitivos. A partir dessas interações não hesitantes entre nós e entre nós e os animais, não se desenvolveram crenças, suposições e conjecturas sobre a mente, mas

sim os nossos próprios conceitos de pensamento, sentimento, intenção, crença, dúvida, e assim por diante. Com um entendimento equivocado sobre essas coisas, fascinados por uma imagem de nós mesmos como espectadores certos de nossas próprias mentes, interpretamos mal a história do desenvolvimento do nosso conceito de mente (GAITA, 2011, p. 77).

“Nossa mente” não é uma suposição ou crença que precisa de algum tipo de justificativa epistêmica para que possa ser aceita, mas está baseada nessa atitude natural que temos diante do outro; não é uma opinião que temos de que o outro “tem uma alma”, mas uma atitude para com alguém que “tem uma alma”.

IV.

Como vimos, ao considerar a linguagem como intrinsecamente ligada às práticas humanas no mundo, Wittgenstein rejeita a ideia de que a linguagem seja um sistema de regras na mente de um sujeito, que o capacita para desenvolver a sua própria linguagem. Nesse sentido, seria absolutamente contrário às noções chomskianas de uma gramática universal inata, por exemplo. Uma criança só desenvolve a linguagem se participa ativamente das relações com outros sujeitos em situações comuns de uso das palavras. Ter uma linguagem, nesse viés, é alcançado por meio de um longo processo de interação de cada falante com outros falantes em um espaço comum de comunicação. É somente depois de um longo aprendizado que se pode afirmar que alguém efetivamente desenvolveu a capacidade da linguagem, e isso significa que é capaz de tomar parte nas diferentes situações e jogos de linguagem com os quais se depara. A linguagem e a humanidade são o resultado das experiências humanas no mundo; experiências individuais, mas também, e principalmente, experiências compartilhadas com outros de sua mesma espécie, que assim conformaram juntos a sua linguagem e a sua humanidade.

Com isso, podemos concluir que a linguagem não é o resultado de um processo de organização de sistemas de regras em um suporte (seja ele orgânico ou não), mas sim da interação entre diferentes sujeitos em uma forma de vida. A linguagem surge de um tipo de reação natural ao ambiente e aos outros sujeitos. Diante dessa constatação, ficaria bastante difícil considerar que máquinas tenham linguagem “tal como nós”, pois a interação que temos com elas é completamente diferente daquelas que os humanos têm entre si, assim como as práticas envolvidas em uma e outra situação são completamente diferentes.

Da mesma maneira, podemos perceber a dificuldade que se coloca sobre a possibilidade de atribuição de pensamento e processos mentais a máquinas, posto que isso não

envolve apenas a interação em um nível de troca de sinais. Ademais, outra dificuldade para aceitar que máquinas possam pensar se assenta no tipo de atitude que somos levados a aceitar como razoáveis diante de tais máquinas. Dessa maneira, a não distinção entre máquinas e humanos dependeria de uma mudança de atitude em relação a elas. Não é apenas a atribuição de pensamento e consciência, mas todo nosso modo de agir e reagir em relação a eles que precisa modificar.

Sendo assim, a questão não diz respeito à busca por alguma coisa que faça com que os humanos sejam humanos, alguma explicação em nível científico do que seja essa propriedade e se ela poderia ser replicada em uma máquina. Como vimos, para Wittgenstein, não há algo como uma mente separada do corpo ou um corpo que pudesse não ter mente. Tais situações nos levariam para um mundo totalmente diferente do nosso, uma forma de vida em que nossos conceitos não mais teriam aplicação. O nosso conceito de humano não permite que façamos tais suposições, porquanto isso levaria a uma situação de um completo estranhamento:

Mas será que não posso imaginar que as pessoas ao meu redor sejam autômatos, não tenham consciência, ainda que seu modo de agir seja o mesmo de sempre? – Quando agora – sozinho em meu quarto – imagino isso, vejo as pessoas seguindo seus afazeres com o olhar fixo (como que em transe) – a ideia talvez seja um pouco perturbadora. Mas tente, então, aferrar-se a essa ideia nas relações usuais com outras pessoas, andando na rua! Diga a si mesmo, por exemplo: “Aqueles crianças são meros autômatos; toda a sua vivacidade é meramente automática”. E essas palavras ou passarão a não dizer absolutamente nada para você; ou despertarão em você um tipo de sentimento perturbador, ou algo do gênero” (WITTGENSTEIN, 2022, § 420).

Essa mesma situação “perturbadora” acontece quando passamos a considerar a possibilidade de máquinas que falam e pensam. Não nos é simples aceitar uma tal afirmação, na medida em que isso não é algo próprio de nossa forma de vida, assim como nossos conceitos de falar e pensar não se aplicam, em nossa forma de vida, às máquinas. Teríamos que promover uma radical modificação nas nossas convicções para que pudessemos aceitar algo como máquinas que pensam e falam.

Fica claro então que é bastante discutível a possibilidade de pensarmos em uma máquina que tenha uma linguagem e uma consciência, pelo menos “tal como os humanos têm”. Dizer que uma máquina pode ser a base para processos mentais implica em dizer que a mente é algo separado do corpo e que pode ser replicada em qualquer outro tipo de base, como placas de silício. É o modelo funcionalista da mente que defende que a cognição humana se assemelha a processos computacionais, processando símbolos de

acordo com regras e algoritmos. Nesse sentido, poderíamos aceitar que cada um de nós é um composto de uma mente e de um corpo, e que nossa mente poderia ser separada de nosso corpo e que nós continuaríamos a ser o que somos.

Esta é a conclusão que Wittgenstein rejeita, pois o ser humano não é um composto de corpo e mente, é uma unidade. E concluímos com Peter Hacker (2001, p. 81):

Se em um futuro distante fosse possível criar em um laboratório eletrônico um ser que agisse e se comportasse como nós, exibindo percepção, desejo, emoção, prazer e sofrimento, bem como pensamento, seria sem dúvida razoável conceber como uma criatura animada, embora não biológica. Mas, nesse sentido, não seria uma máquina, ainda que fosse fabricada.

Referências

COCKBURN, David. **An introduction to the philosophy of mind**. New York: Palgrave, 2001.

GAITA, Raimond. **O cão do filósofo**. Rio de Janeiro: DIFEL, 2011.

HACKER, Peter Michael Stephan. **Wittgenstein: meaning and mind**. Volume 3 of an Analytical Commentary on the Philosophical Investigations. Oxford: Blackwell, 2001.

MOYAL-SHARROCK, Danièle. Wittgenstein: Sobre a aquisição da linguagem. In: CARMO, Juliano Santo do & SAUCEDO, Rogério Correa (Orgs.). **Linguagem e cognição**. Pelotas: NEPFIL Online, 2018, p. 90-118. Disponível em: <https://wp.ufpel.edu.br/nepfil/files/2018/11/LCVFinal.pdf>

SEARLE, John. **Mentes, cerebros y ciencia**. Madrid: Cátedra, 1994.

WINCH, Peter. **A ideia de uma ciência social e sua relação com a filosofia**. São Paulo: Editora da Unesp, 2020.

WITTGENSTEIN, Ludwig. **Cultura e valor**. Lisboa: Edições 70, 2000.

WITTGENSTEIN, Ludwig. **Investigações filosóficas**. São Paulo: Fósforo, 2022.





ANDRÉ DIAS 2017

A TEORIA DE FRANZ BRENTANO SOBRE A CONSCIÊNCIA: DEBATES ATUAIS



Tárik de Athayde Prata

Resumo:

O artigo aborda o argumento de Brentano para a sua concepção de consciência, de acordo com a qual (a) nossos fenômenos mentais estão *fundidos* (b) com nossa consciência desses fenômenos. Discuto as interpretações de Uriah Kriegel (2003) e de Mark Textor (2017) sobre esse argumento, que visa recusar a tese de que a negação de fenômenos inconscientes conduz a um regresso ao infinito. Defendo que Textor tem razão ao colocar que o argumento, diferente do que alega Kriegel, não inclui a tese do regresso como uma de suas premissas.

Palavras-chave:

Consciência; Inconsciente; Regresso ao infinito; Percepção interna.

Abstract:

The article approaches Brentano's argument for his account of consciousness, according to which (a) our mental phenomena are *fused* (b) with our consciousness of them. I discuss the interpretations of Uriah Kriegel (2003) and Mark Textor (2017) about this argument, which aims to refute the claim that the denial of unconscious phenomena leads to an infinite regress. I maintain that Textor is right when he states that the argument, unlike Kriegel claims, does not include the thesis of a regress as one of its premises.

Keywords:

Consciousness; Unconscious; Infinite regress; Inner perception.

1. Introdução

Ao longo do século XX, o filósofo alemão Franz Brentano (1838-1917) foi particularmente lembrado como aquele que redescobriu o tema da *intencionalidade*¹ – (cf. PORTA 2014, p. 7) –, a característica de fenômenos mentais em virtude da qual eles são *a respeito de algo*², exercendo influência tanto na tradição da fenomenologia quanto na tradição da filosofia analítica – cf. Jacquette (2004, p. 1). Porém, desde o início do século XXI, a obra de Brentano tem sido revisitada por causa de suas contribuições para o debate a respeito da *consciência*³.

No presente trabalho, pretendo discutir algumas interpretações contemporâneas da teoria de Brentano sobre a consciência ou, mais exatamente, interpretações de seu argumento contra a tese de que a negação do inconsciente conduz a um *regresso ao infinito*. Esse argumento é a base da concepção de Brentano sobre a consciência interna (nossa consciência de nossos próprios estados mentais). Nessa concepção, a consciência interna consiste, fundamentalmente, em uma *fusão* de representações.

Mais especificamente, irei abordar as reconstituições do argumento contra o regresso articuladas por Uriah Kriegel (2003) e Mark Textor (2017), dois importantes intérpretes contemporâneos da filosofia de Brentano.

De acordo com Textor (2017), em uma passagem do *De Anima* [425b 12 - 425b 17], muito discutida no debate sobre Brentano, Aristóteles, em sua tentativa de recusar a ideia de um “sentido comum”, sugeriu duas possibilidades de argumentação: (1) o argumento da *duplicação do objeto* e (2) o argumento do *regresso ao infinito*. Entretanto, para

1 Na mais famosa passagem de sua principal obra, a *Psicologia de um ponto de vista empírico* (1874), Brentano escreveu: “Todo fenômeno psíquico é caracterizado por aquilo que os escolásticos da idade média denominaram o ‘existir em’ [*Inexistenz*] intencional (...), o que nós denominaríamos, embora não com expressões totalmente precisas, a relação a um conteúdo, a direção a um objeto (com o que não se deve entender aqui uma realidade), ou a objetividade imanente. Todo fenômeno psíquico contém em si algo como objeto, embora não todos do mesmo modo. Na representação algo é representado, no juízo algo é reconhecido ou recusado, no amor amado, no ódio odiado, no desejo desejado, etc.” (Brentano, 1924, p. 124-5; Brentano, 1995, p. 88).

2 Uma caracterização bastante clara e completa da noção de intencionalidade é dada por Wilfrid Sellars (1958, p. 507): “O status de referência a objetos e estados de coisas, atuais ou possíveis, passados, presentes ou futuros, o qual está envolvido no próprio significado do vocabulário ‘mentalístico’ da vida cotidiana. Acreditar, desejar, pretender, amar, odiar, raciocinar, aprovar – na verdade todos os estados e disposições caracteristicamente humanos acima do nível da consciência meramente sensorial – não podem ser explicados sem envolvimento [*encountering*] com tal referência ou sobriedade [*aboutness*]”.

3 Cf., por exemplo, Thomasson (2000); Hossack (2002); Zahavi (2004); Zahavi (2006); Textor (2006); Textor (2013) e Textor (2017).

Textor, em sua *Psicologia de um ponto de vista empírico*, no capítulo sobre a consciência interna, Brentano desenvolveu apenas a primeira possibilidade de argumentação (o argumento da duplicação), ainda que seu objetivo fosse refutar o regresso ao infinito (que era usado na literatura como um argumento a favor da existência de fenômenos mentais inconscientes).

Entendo que a comparação cuidadosa da reconstituição que Textor (2017, p. 55-6) com o argumento do próprio Brentano (1924, p. 176-79; 1995, p. 126-7) *confirma* essa interpretação. O exame desse argumento não revela uma correspondência muito exata com a reconstituição de Kriegel (2003, p. 112-3), que inclui a própria noção de um regresso ao infinito como uma das premissas do argumento.

Após uma discussão sobre alguns dos inúmeros sentidos da palavra “consciência” (seção 2) – discussão que nos ajudará a entender o sentido específico de *consciência* que está em jogo no argumento de Brentano –, farei uma rápida exposição dos quatro argumentos de Brentano contra a existência de fenômenos mentais inconscientes⁴ (seção 3), para então reconstituir, de maneira mais ou menos literal, o argumento de Brentano contra o regresso ao infinito (seção 4). Em seguida, discutirei a reconstituição desse argumento feita por Kriegel (2003), reconstituição segundo a qual a própria tese de que há um regresso é usada como uma *premissa* do argumento de Brentano (seção 5), e discutirei a reconstituição de Textor (2017), que defende que o argumento de Brentano no parágrafo oito do capítulo sobre a consciência interna é uma reformulação do *argumento da duplicação do objeto*, proposto por Aristóteles (seção 6).

2. Consciência

De acordo com Anders Schinkel (2007, p. 79), a palavra “*syneidesis*”, que os latinos traduziram por “*conscientia*”, tinha sentido próximo do que hoje chamaríamos “consciência” *de* algo. Esse sentido se constituiu através de um longo processo. Nas palavras de Marietta Jr. (1970, p. 177):

O prefixo “syn” significava a princípio *comunidade de saber*, tal como uma testemunha, confidente ou cúmplice teria. Existem exemplos desse significado

⁴ Na verdade, esses quatro argumentos de Brentano são *réplicas* a quatro argumentos, encontrados na literatura, a favor da existência de fenômenos mentais inconscientes.

em Sófocles, Xenofonte e Plutarco. Desse uso se desenvolveu o significado de “*saber consigo mesmo*”, que é expresso com ou sem o pronome reflexivo (grifos meus)⁵.

Segundo Schinkel (2007), o significado peculiarmente *moral* da palavra “*syneidesis*”, assim como de termos cognatos, se desenvolveu gradualmente, e nunca foi seu uso exclusivo. Mesmo quando essa palavra designava a consciência moral, o significado fundamental era o sentido original de consciência ou *ciência* [*awareness*].

É interessante perceber que esses três sentidos da palavra “*syneidesis*”, a saber, (i) o sentido social, (ii) o sentido individual e (iii) o sentido moral, ainda estão presentes nos usos das palavras atualmente usadas para designar a consciência, em diferentes línguas – cf. Güzeldere (1997, p. 8, p. 49).

Porém, é importante ressaltar que, nos debates teóricos contemporâneos, a palavra “consciência” é usada de maneira extremamente ambígua, podendo significar, por exemplo, (1) nossa consciência de objetos e estados de coisas, não necessariamente reais (isto é, a própria intencionalidade), (2) o modo como sentimos nossas experiências, (3) a sequência de nossos fenômenos mentais conscientes, (4) a *propriedade de um fenômeno mental* de estar presente nessa sequência etc.

O próprio Brentano (1924, p. 141; 1995, p. 101, grifos meus), reconhecendo a profunda ambiguidade desse termo, afirmou que:

Não se pode falar de um significado amplamente aceito e exclusivo, que estivesse ligado a essa palavra. Nos convencemos disso imediatamente tão logo observamos as exposições gerais que fizeram Bain, na Inglaterra, e Horwicz, na Alemanha, sobre os diferentes usos dessa palavra. Ora se entende com ela a lembrança dos próprios atos passados, especialmente se eram de *natureza moral*, como quando se diz: não tenho nenhuma consciência de culpa. Ora se designa com ela todo tipo de *conhecimento imediato* dos próprios atos psíquicos, especialmente uma percepção que acompanha os atos psíquicos presentes.

Entre os quatro sentidos de “consciência” mencionados acima, antes da citação, o segundo sentido, a assim chamada “consciência *fenomenal*” (*phenomenal consciousness*)

5 Todas as citações em língua estrangeira foram traduzidas para o português pelo autor do presente artigo.

teve um lugar de grande destaque no campo da filosofia analítica desde a década de 1970⁶. Decisivas para esse destaque foram as contribuições de Thomas Nagel, que, no famoso artigo *Como é ser um morcego?* (1974), defendeu que o problema mente-corpo se torna desesperadoramente difícil⁷ em virtude daquilo que ele designou como *o caráter subjetivo* da experiência consciente. De acordo com ele:

O fato de um organismo ter, *seja lá como for*, uma experiência consciente significa, basicamente, que há algo que seja ser como aquele organismo. (...) [F]undamentalmente, um organismo tem estados mentais conscientes se e somente se existe algo que é como ser esse organismo – algo que é como ser *para* o organismo. (NAGEL, 1974, p. 436; NAGEL, 2005, p. 247).

Isso geraria, na visão de Nagel (1974, p. 442; 2005, p. 254), um problema desconcomunal para o fisicalismo, porque, na sua perspectiva, “Se os fatos da experiência – fatos sobre como é *para* o organismo que tem a experiência – são acessíveis apenas de um ponto de vista, logo, é um mistério como o verdadeiro caráter das experiências poderia ser revelado através das operações físicas do organismo”. Entretanto, o destaque da consciência fenomenal na atual filosofia da mente também pode ser visto como algo problemático. O filósofo israelense Uriah Kriegel defendeu que:

Esse foco na qualidade sensorial da experiência eclipsa outro aspecto da consciência que é tradicionalmente pensado como mais central e talvez mais problemático. Kant tentou capturar esse aspecto quando ele escreveu que todo pensamento consciente tem um ‘Eu penso’ atrelado a ele, um ‘Eu penso’ que ‘acompanha (...) todas as representações o tempo todo em minha consciência’ (Kant 1965, A362-63) (KRIEGEL, 2002, p. 518).

Em sua tentativa de desenvolver uma perspectiva adequada a respeito desse caráter subjetivo da experiência – isto é, a característica da experiência de *ser para* o sujeito dessa experiência⁸ – Kriegel (2002) recorreu às contribuições de Brentano para a investi-

6 “A literatura filosófica sobre a consciência é focada principalmente no *caráter qualitativo* [grifo meu] da experiência. (...) Discussão do caráter subjetivo da experiência, e das perspectivas de naturalização e termos funcionalistas e/ou *representacionistas* [grifo meu], é muito mais limitada” (KRIEGEL, 2005, p. 24).

7 Sobre a ideia da subjetividade da experiência como fonte do *problema difícil* (*hard problem*) a respeito da consciência, cf. Chalmers (1995); Chalmers (2020).

8 Ele propõe uma distinção entre (a) o caráter *qualitativo* e (b) o caráter *subjetivo* da experiência nos seguin-

gação sobre a estrutura dos fenômenos conscientes. Na seção 5, a seguir, examinarei a reconstituição que Kriegel (2002) oferece sobre o argumento de Brentano para a tese central de sua teoria da consciência, a tese de que todo fenômeno mental é baseado em uma *fusão* de representações: a representação (primária) de um fenômeno físico e a representação (secundária) dessa primeira representação. Essa tese é uma inspiração para a teoria *auto-representacionista* que Kriegel (2002) articula a respeito da consciência.

Um outro importante pensador contemporâneo, o filósofo dinamarquês Dan Zahavi, entende a contribuição de Brentano para o debate a respeito da consciência em termos de sua capacidade de oferecer uma alternativa à chamada “perspectiva de ordem superior” (*higher-order view*). Nas palavras dele:

Por algumas décadas, teorias de ordem superior sobre a consciência tem desfrutado de grande popularidade, mas elas têm enfrentado recentemente uma crescente insatisfação. Muitos têm começado a procurar por alternativas viáveis, e nos últimos anos alguns têm redescoberto Brentano (ZAHAVI, 2004, p. 66).

De acordo com a perspectiva de ordem superior (que entende a consciência como uma *propriedade* de estados mentais), um dado estado mental exemplifica a propriedade de ser consciente quando ele é representado por *um outro* estado mental⁹. Brentano (1924), de fato, se opunha a esse tipo de visão. Na sua *Psicologia de um ponto de vista empírico*, ele afirmou:

Se a cognição [*Erkenntnis*] que acompanha um ato mental fosse um ato em seu próprio direito, *um segundo ato adicionado ao primeiro* [grifo meu], se a sua relação com o seu objeto fosse simplesmente aquela de um efeito com a sua causa, similar àquela que existe entre uma sensação e o estímulo físico que a produz, como poderíamos estar seguros por um instante de sua infalibilidade? (BRENTANO, 1924, p. 196; 1995, p. 139).

tes termos: “Quando eu tenho uma experiência consciente do céu, há um modo azulado [*bluish way*] como é para mim ter ou vivenciar minha experiência. Eu sugiro distinguir dois aspectos nesse ‘modo azulado como é para mim’: (i) o aspecto azulado, que nós podemos chamar de caráter qualitativo da experiência, e (ii) o aspecto para-mim, que nós podemos chamar de seu caráter subjetivo” (KRIEGEL, 2005, p. 23).

9 Cf. Rosenthal (1986, p. 335); Rosenthal (2008, p. 835).

Aqui, Brentano (1924) rejeita que a nossa consciência de nossos próprios estados mentais decorra da ação de *um outro* estado mental, o que equivale à perspectiva que hoje é designada como sendo de “ordem superior”. Essa rejeição é baseada em sua tese da *infallibilidade* dessa nossa consciência interna – cf. Brentano (1924, p. 198; 1995, p. 140). Entendo que tal infalibilidade é extremamente problemática – cf., por exemplo, as críticas de Armstrong (1968, p. 109-10) e Rosenthal (2018) a esse tipo de concepção –, mas esse tema não será discutido no presente trabalho.

O fato é que Brentano (1924) articula a sua argumentação a favor da *fusão* entre (a) nossos fenômenos mentais e (b) a nossa consciência desses fenômenos – fusão que garantiria a infalibilidade – em sua discussão de um dos quatro tipos de argumentos (que ele identifica na literatura de seu tempo) a favor da existência de fenômenos mentais inconscientes.

3. Os argumentos de Brentano contra a existência de fenômenos mentais inconscientes

No capítulo sobre a consciência interna, Brentano (1924, p. 147-8; 1995, p. 105) identifica quatro argumentos a favor da existência de fenômenos mentais inconscientes:

- (1) “Primeiro, poderíamos tentar provar que certos fatos dados na experiência requerem a hipótese de um fenômeno mental inconsciente como a sua *causa*”.
- (2) “Segundo, poderíamos tentar provar que um fato dado na experiência tem que produzir um fenômeno mental inconsciente como seu *efeito*”.
- (3) “Terceiro, poderíamos tentar mostrar que no caso de fenômenos mentais conscientes *a força da consciência concomitante é uma função de sua própria força*, de modo que, por causa dessa relação, em certos casos nos quais a última [a força do fenômeno mental] tem uma magnitude positiva, a primeira [a força da consciência concomitante] tem de carecer de um valor positivo”.
- (4) “Finalmente, poderíamos tentar provar que a hipótese de que todo fenômeno mental é objeto de um outro fenômeno mental leva a uma infinita complexidade de estados mentais, a qual é tanto intrinsecamente impossível quanto contrária à experiência”.

Cada um desses argumentos, de acordo com ele, teria uma série de condições, as quais não seriam satisfeitas na maneira como os argumentos foram efetivamente articulados¹⁰. Para provar que certos fatos requerem fenômenos inconscientes como suas *causas*, seria necessário, segundo Brentano, que:

- (1) Esses próprios fatos fossem *suficientemente estabelecidos* (o que os próprios defensores do inconsciente reconhecem que não é fácil) – cf. Brentano (1924, p. 148-9; 1995, p. 106);
- (2) Que a hipótese do inconsciente *realmente explicasse* esses fatos – cf. Brentano (1924, p. 149-53; 1995, p. 106-9);
- (3) E também que o fenômeno não possa ser entendido em termos de *outras hipóteses* – cf. Brentano (1924, p. 153; 1995, p. 109).

Nas passagens referenciadas aqui, ele nega que os defensores desse primeiro argumento consigam satisfazer essas condições. Além disso, para provar que um fato dado na experiência tem que produzir um fenômeno mental inconsciente como seu *efeito*, seria necessário que:

- (1) Fosse excluída a possibilidade de que esse fato causou um fenômeno mental consciente como efeito, e que esse fenômeno consciente foi *esquecido* – cf. Brentano (1924, p. 163; 1995, p. 116);
- (2) Também seria necessário que fosse provado que a suposta causa inconsciente era do *mesmo tipo psicológico* que as causas conscientes usuais – cf. Brentano (1924, p. 163; 1995, p. 116);
- (3) Teria que ser provado que as causas que evitam o aparecimento da consciência *não evitam* o aparecimento do próprio fenômeno mental – cf. Brentano (1924, p. 163; 1995, p. 116).

¹⁰ É importante ressaltar que, como coloca Kriegel (2018, p. 19): “Brentano sustentou a visão problemática de que toda a vida mental é consciente – não existem estados mentais inconscientes. Desse modo, sua filosofia da mente é, mais precisamente, uma filosofia da consciência”.

Brentano também não aceita que essas condições tenham sido satisfeitas pelos defensores desse segundo argumento. No que diz respeito ao argumento de que certos fenômenos mentais são *fracos demais para serem conscientes*, Brentano afirma que nenhum de seus defensores conseguiu provar o nexos entre (a) a intensidade de um fenômeno mental e (b) a intensidade da consciência correspondente, e afirma que a falta de acurácia das medições psicológicas pode ser vista como um obstáculo para se estabelecer essa correlação – cf. Brentano (1924, p. 168; 1995, p. 120).

Por outro lado, ele afirma que, em nossas experiências conscientes, nós constatamos uma correlação entre (a) a intensidade dos fenômenos representados (conscientemente) e (b) a intensidade das representações (conscientes) desses fenômenos. Com base na (suposta) infalibilidade de nossas percepções internas – nossas percepções de nossos próprios fenômenos mentais –, Brentano sustenta que temos de aceitar essa correlação e afirma também que tal correlação mostra que todos os fenômenos mentais têm que ser conscientes, pois “onde quer que exista um ato mental de maior ou menor intensidade, é necessário atribuir uma intensidade igual à representação que o acompanha” (BRENTANO, 1924, p. 170; 1995, p. 121) – cf. Textor (2006, p. 421). Em outras palavras, Brentano vira esse terceiro argumento contra os defensores do inconsciente.

Porém, esses três primeiros argumentos não foram muito influentes no debate sobre a filosofia de Brentano¹¹. Dos quatro argumentos a favor da existência de fenômenos mentais inconscientes recusados por Brentano, sem dúvida, o mais debatido na recepção à sua obra¹² foi o argumento de que a recusa da existência de fenômenos inconscientes desencadeia um *regresso ao infinito*, regresso que mostraria que essa recusa é insustentável.

4. A réplica de Brentano para o argumento do regresso ao infinito

No segundo capítulo do segundo livro de sua *Psicologia*, Brentano articula o seu argumento contra a tese de que a *recusa* da existência de fenômenos mentais inconscientes conduz a um regresso ao infinito. A ideia dos defensores do inconsciente é que, mesmo no caso de um fenômeno mental extremamente simples – o exemplo usado por Brentano é

11 Para algumas discussões a respeito desses três primeiros argumentos, cf. Krantz (1990, p. 746) e Tassone (2012, p. 124-8).

12 Cf., por exemplo, Kranz (1990, p. 746); Kriegel (2003, p. 112-5); Zahavi (2004, p. 71-2); Zahavi (2006, p. 3) e Textor (2006, p. 424-30).

a *escuta* (*Hören*) de um som –, a negação de que esse fenômeno possa existir inconscientemente exigiria a presença de um segundo fenômeno mental que o representasse para o sujeito portador desse fenômeno. No entanto, se a existência de fenômenos inconscientes é negada, esse segundo fenômeno teria que ser representado por um terceiro que, por sua vez, teria que ser representado por um quarto fenômeno mental, e assim *ad infinitum*.

Em sua tentativa de refutar esse argumento, Brentano articula um argumento, cuja formulação irei citar de maneira completa:

Existem indubitavelmente ocasiões nas quais estamos conscientes de um fenômeno mental enquanto ele está presente em nós; por exemplo, enquanto temos a representação de um som, estamos conscientes de tê-la. Pode-se perguntar: em um caso como esse, nós temos diversas representações heterogêneas ou apenas uma única? Antes de responder a essa pergunta, temos de ter clareza sobre se nós queremos determinar o número e a variedade de representações segundo [A] o número e a variedade de *objetos* [grifo meu], ou se segundo [B]¹³ o número dos *atos psíquicos* [grifo meu], nos quais nós representamos os objetos. No primeiro caso, então nós teríamos que dizer que nós temos diversas representações de diferentes tipos, de tal modo que uma formaria o conteúdo da outra, enquanto ela mesma teria um fenômeno físico como conteúdo. Se isso for correto, então o fenômeno físico tem que, de certo modo, pertencer ao conteúdo de ambas as representações, a uma como objeto explícito [*ausschließlicher*], à outra, por assim dizer, como objeto implícito [*eingeschlossener*]¹⁴. Pareceria, portanto, como Aristóteles já notou, resultar que o fenômeno físico teria que ser representado *duas vezes* [grifo meu]. Porém, esse não é o caso. Na verdade, a experiência interna parece provar de modo indubitável que a representação do som está conectada com a representação da representação de uma maneira tão peculiarmente íntima que o seu ser ao mesmo tempo contribui internamente para o ser da outra. Isso indica o peculiar entrelaçamento [*eigentümliche Verwebung*] do objeto da representação interna com essa representação mesma e [indica] o pertencimento de ambas ao mesmo ato psíquico (BRENTANO, 1924, p. 176-9; BRENTANO, 1995, p. 126-7).

Parece-me que, a partir da leitura atenta dessa passagem, o argumento pode ser reconstituído – de forma mais próxima às formulações do texto – nos seguintes termos:

(P1) Algumas vezes, estamos conscientes de um fenômeno mental enquanto ele está presente em nós;

13 Incluí essas letras para destacar as duas opções no texto.

14 Textor (2006, p. 417) também traduz o termo “*ausschließlicher*” por *explícito* e o termo “*eingeschlossener*” por *implícito*, assim como esses termos foram traduzidos na edição britânica da *Psicologia* de Brentano.

(P2) Se estamos conscientes de um fenômeno mental, então ou (a) nós temos *várias representações*, de diferentes tipos, ou (b) nós temos *apenas uma representação*;

(P3) Podemos determinar o número e variedade das representações ou (I) por meio do número e variedade dos *objetos*, ou (II) por meio do número e variedade dos *atos mentais*;

(P4) Se nós determinarmos o número de representações *através do número de objetos*, então há uma multiplicação de representações, pois o objeto tem que ser representado duas vezes;

(P5) Nossa experiência interna mostra que (i) a representação do objeto (representação primária) contribui internamente para o ser da (ii) representação da representação do objeto (representação secundária);

(C1) A (i) representação primária e a (ii) representação secundária formam *um único fenômeno mental*;

(C2) Não há regresso ao infinito.

Essa é uma reconstituição bastante simples, na medida em que permanece muito próxima à própria formulação de Brentano. Gostaria de destacar que, nesta reconstituição mais literal do argumento, não encontramos entre as premissas nenhuma referência ao regresso. Encontramos tal referência apenas na conclusão. Esse fato me parece corroborar a interpretação de Textor, que será discutida a seguir.

De todo modo, as tentativas de reconstituição do argumento que serão examinadas nas próximas seções são mais sofisticadas, na medida em que tentam formular as suas premissas com base numa interpretação do *contexto mais amplo* no qual Brentano articula suas teses.

Porém, enquanto Kriegel parece se ater ao fato de que o argumento foi, explicitamente, articulado para refutar o regresso ao infinito (supostamente decorrente da negação da existência de fenômenos inconscientes), Textor, por outro lado, expande seu olhar ao

contexto histórico-filosófico a partir do qual Brentano elaborou suas ideias. Mais exatamente, Textor leva em conta o fato de que Brentano, ao articular o argumento contra o regresso ao infinito (supostamente desencadeado pela recusa do inconsciente), estava dialogando com certas teses defendidas por Aristóteles no seu tratado *De Anima*.

Considerarei primeiro a reconstituição de Kriegel.

5. A reconstituição de Kriegel

Na tentativa de fundamentar a sua própria visão *autorepresentacionista* da consciência de estado (isto é, a consciência como uma *propriedade* de estados mentais – cf. Kriegel, 2009, p. 26), Uriah Kriegel (2003, p. 112) examina a teoria de Brentano sobre a consciência, usando-a como um degrau a ser superado na articulação da sua própria teoria, e reconstitui o argumento para a fusão de representações nos seguintes termos:

(1) Um estado mental M , de um certo sujeito x é consciente no tempo t somente se x está ciente de M no tempo t ;

(2) Se x está ciente de M no tempo t , então a ciência de x a respeito de M ou (i) é parte do próprio M , ou (ii) é parte de um estado mental M^* , onde $M^* \neq M$;

(3) A opção (ii) leva a um regresso ao infinito [como mostrará a argumentação exposta a seguir];

(4) A opção (i) é o caso, ou seja, a ciência de x a respeito de M é parte do próprio estado mental M [decorrente de (2') e (3')];

(5) Um estado mental M de um sujeito x no tempo t é consciente somente se M constitui (parcialmente) ciência de si mesmo.

Um ponto que me parece digno de nota é que, enquanto a premissa 1, tal como formulada por Kriegel, corresponde às ideias de Brentano, a premissa 2 não me parece corresponder ao argumento propriamente dito, na medida em que Brentano (1924, p. 176-9; 1995, p. 126-7), na passagem citada acima, não faz uso do conceito de *parte*. A formulação desta premissa só pode ser fruto de interpretação, mas justamente essa interpretação me parece impor ao argumento uma perspectiva que não encontramos nele de fato.

Como vemos, a terceira premissa do argumento, segundo a reconstituição de Kriegel, envolve a noção de um regresso ao infinito, regresso este que Kriegel (2003, p. 112-3) explica¹⁵ nos seguintes termos:

- (1) Para todo estado mental M_i , se M_i é consciente, então o sujeito x tem – simultaneamente – um estado mental M_{i+1} tal que M_{i+1} representa M_i ;
- (2) Se M_1 é consciente no tempo t , então x tem um M_2 em t , tal que M_2 representa M_1 [decorrente de (1)];
- (3) M_1 é consciente (*ex hypothesi*), portanto:
- (4) x tem no tempo t um M_2 , tal que M_2 representa M_1 [decorrente de (2) e (3)];
- (5) Todos os estados mentais são conscientes, portanto:
- (6) M_2 é consciente [decorrente de 5];
- (7) Se M_2 é consciente, então x tem no tempo t um estado mental M_3 , tal que M_3 representa M_2 [decorrente de (1)], portanto:
- (8) x tem no tempo t um estado M_3 , tal que M_3 representa M_2 [decorrente de (6) e (7)];
- (9) M_3 é consciente [decorrente de (5)];

Ad infinitum...

15 Uma outra explicação para o surgimento do problema do regresso ao infinito – no caso dos fenômenos mentais conscientes – é oferecida por Denis Fisette (2015, p. 25):

- (1) Todo fenômeno mental é sobre um objeto (a escuta de um som) (TESE I).
- (2) Todo fenômeno mental é, ele mesmo, objeto de uma consciência que o acompanha (TESE II).
- (3) A representação que acompanha o estado mental inicial é numericamente distinta do estado mental representado.
- (4) Se, porém, a representação também tem que ser consciente (TESE II), e a representação que a faz consciente tem que ser, por sua vez, consciente, então a série é infinita.
- (5) Portanto, ou a representação do estado inicial é inconsciente (e a TESE II é, então, falsa), ou tem que haver um número infinito de estados mentais.

Neste caso, a premissa (5) – a tese de que todos os estados mentais são conscientes – é exatamente a tese que os defensores do inconsciente querem refutar, argumentando que essa tese conduz a um regresso “intrinsecamente impossível e contrário à experiência”, como diz Brentano (1924, p. 148; 1995, p. 105). Mas essa é, justamente, a tese que Brentano quer *defender*. É para simultaneamente (i) defender essa tese e (ii) recusar o regresso que Brentano recorre à tese de que há uma *fusão* de representações, isto é, a tese de que todo fenômeno mental, que é representação de um fenômeno físico, é sempre representado por uma representação que está *internamente unida* a ela (a primeira representação).

O próprio Kriegel (2003) quer articular um autorepresentacionalismo – a tese de que estados mentais são conscientes por incluírem uma representação deles mesmos – sem se comprometer com a tese de que todos os fenômenos mentais são conscientes¹⁶. Mas esse não é o tema do presente trabalho.

Sem dúvida, não é absurdo supor que a noção de um regresso ao infinito desempenhe um papel no argumento de Brentano, até porque é bem sabido que esse argumento é baseado em uma passagem onde Aristóteles, uma referência filosófica importantíssima para Brentano – cf. George & Koehn (2004) –, discute esse regresso. Como explica Smith (1986, p. 149):

Revivendo *um antigo argumento* [grifo meu], Brentano raciocinou que a representação secundária não pode ser uma segunda representação, uma reflexão ou um juízo sobre a representação primária. Pois isso levaria a um *regresso ao infinito* [grifo meu]. (...) Assim, Brentano observou, a consciência [*awareness*] secundária que torna uma experiência consciente tem que ser interna à experiência, e não um segundo ato mental que acompanha ou segue a experiência.

Kriegel (2003, p. 111) é ainda mais explícito a esse respeito:

O principal argumento de Brentano (...) também é emprestado do *De Anima* III.2 (425^bll-7) de Aristóteles. (...) Esse argumento é desenvolvido e defendido por Brentano através do segundo livro da *Psicologia*. Sua estrutura é bastante direta. Uma percepção consciente é ela mesma percebida; portanto, ou é percebida por si mesma ou por outra percepção; mas ela não pode ser percebida por outra percepção, pois isso levaria a um regresso infinito de percepções; portanto ela tem que ser percebida por si mesma¹⁷.

16 “Brentano sustenta que todos os estados mentais são conscientes. Como nós veremos, essa alegação bastante implausível pode ser extirpada do seu sistema sem muita repercussão. Mas ela significa que sua filosofia da mente é, no fundo, realmente apenas uma filosofia da consciência.” (KRIEGEL, 2018, p. 8).

17 A respeito da inspiração aristotélica do argumento de Brentano, cf. também Textor (2013, p. 466) e Textor

Todavia, me parece que um exame mais detido das considerações de Aristóteles, exame que é, de fato, empreendido por Textor, mostra que a relação entre o argumento de Brentano e a noção de regresso ao infinito é *mais complexa* do que é suposto por Smith e Kriegel. Ainda que o regresso ao infinito – supostamente implicado pela negação da existência de fenômenos mentais inconscientes – seja, sem dúvida, o tema do argumento, temos razões para crer que, inspirando-se em Aristóteles, Brentano articulou o argumento em termos do problema da *duplicação do objeto* – cf. Textor (2006, p. 417-21).

6. A interpretação de Textor

De acordo com Textor, o argumento de Brentano para a tese da *fusão* de representações (fusão que estaria na base de todos os nossos fenômenos mentais) precisa ser entendido no contexto das considerações de Aristóteles sobre o problema do *sentido comum*, isto é, “o sentido por meio do qual nós fazemos comparações ‘através’ de diferentes sentidos, mais específicos” (TEXTOR, 2017, p. 52).

Em sua tese de doutorado, *A psicologia de Aristóteles – em especial a sua doutrina no Nous Poiétikos* (1867), Brentano baseou-se no critério aristotélico para a individuação de sentidos para defender a tese de que há um sentido extra, através do qual percebemos nossas percepções sensoriais, atreladas aos cinco sentidos. Textor (2017, p. 50-1) reconstituiu o argumento de Brentano para essa tese nos seguintes termos:

(P1) Nós percebemos ver quando vemos;

(P2) Ou nós percebemos ver pela visão ou por outro sentido;

(P3) O objeto distintivo da visão é a cor; por meio da visão, podemos perceber também forma e tamanho junto com a cor;

(P4) Ver não é colorido, nem tem forma ou tamanho que possa ser visto junto com a cor;

(C1) Logo, nós não percebemos ver pela visão;

(2017, p. 55).

- (C2) Logo, nós percebemos ver por um outro sentido;
- (P5) Ver não é o objeto distintivo de nenhum dos outros sentidos;
- (C3) Portanto, nós não percebemos ver por nenhum dos cinco sentidos conhecidos;
- (C4) logo, nós percebemos ver por um sentido extra.

Brentano entendia que Aristóteles não foi claro quanto a se concordava ou não com a existência de um sentido interno. Mas se em sua tese de doutorado ele coincidiu com a maioria dos intérpretes ao atribuir essa tese a Aristóteles – cf. Brentano (1924, p. 185-6; 1995, p. 132, nota de rodapé) –, em seu mais conhecido livro, *A psicologia de um ponto de vista empírico* (1874), o filósofo alemão acabou revisando sua interpretação. Ele entendeu que uma importante passagem do *De Anima* de Aristóteles implica a rejeição da tese de um sentido interno, através do qual nós perceberíamos nossas percepções dos fenômenos externos. Nessa passagem, o filósofo grego escreveu:

Uma vez que é através dos sentidos que nós estamos cientes de que estamos vendo ou ouvindo, tem que ser ou através da vista que estamos cientes de estar vendo, ou através de algum outro sentido que não a vista. Mas o sentido que nos dá essa nova sensação tem que perceber tanto a vista quanto o seu objeto, por exemplo, uma cor, de modo que ou vão existir dois sentidos que percebam o mesmo objeto sensível, ou o sentido perceberá a si mesmo. Além disso, mesmo se o sentido que percebe a vista fosse diferente da vista, nos vemos diante de duas opções: ou caímos em um regresso ao infinito, ou temos de aceitar que o sentido é ciente de si mesmo (ARISTOTLE, 1995, p. 677 [425b 12 - 425b 17]).

De acordo com Textor, aqui, Aristóteles sugere tanto (A) o argumento da *duplicação do objeto*¹⁸ quanto (B) o argumento do *regresso ao infinito*. Mas ele argumenta que Brentano não recorre a esse último para articular a sua refutação do regresso. Nas palavras desse intérprete:

18 Não está completamente claro para mim em que sentido Aristóteles, nessa passagem [425b 12 - 425b 17], está sugerindo o argumento da duplicação do objeto. Mas suponho que seja no momento em que ele diz que “vão existir dois sentidos que percebam o mesmo objeto sensível”, pois isso implica que esse objeto externo terá que ser representado nos *dois* diferentes sentidos, sendo, assim, duplicado.

A ameaça do regresso é um obstáculo que precisa ser removido para limpar o caminho para a aceitação da perspectiva de que todo ato mental é consciente; mas ela não é usada por Brentano em um argumento a favor da sua visão positiva. O argumento de Brentano para a sua nova visão, no meu entendimento, é um desenvolvimento do argumento da duplicação de Aristóteles (cf. BRENTANO, 1874, p. 176-9; 1995a, p. 97-8). (TEXTOR, 2017, p. 55).

Textor (2017, p. 55-6) reconstitui o argumento de Brentano, de forma esquemática, nos seguintes termos:

- (P1) Assuma-se, para *reductio*, que S percebendo S percebendo x é distinto de S percebendo x;
- (P2) S percebendo S percebendo x é diretamente de S percebendo x, e indiretamente de x;
- (C1) Logo, quando S percebe S percebendo x, x é representado duas vezes: uma vez na percepção de S percebendo x, e uma segunda vez em S percebendo x;
- (P3) Não há dupla representação de x quando S percebe S percebendo x;
- (C2) Logo, S percebendo S percebendo x é o mesmo ato que S percebendo x.

Penso que podemos elucidar o sentido dessa reconstituição do argumento nos seguintes termos: se perceber a percepção de um fenômeno físico ($S_2 \rightarrow S_1 \rightarrow x$) for suposto como algo distinto da percepção desse fenômeno ($S_1 \rightarrow x$), então a percepção da percepção (S_2) do fenômeno físico será *diretamente* a respeito da percepção (S_1) do fenômeno e *indiretamente* a respeito do próprio fenômeno (x). Disso decorreria que o fenômeno físico teria que ser representado duas vezes, uma vez em S_2 e outra vez em S_1 . Porém, como os dados de nossa experiência interna não mostram uma dupla representação, podemos concluir que a percepção da percepção do fenômeno físico e a simples percepção desse fenômeno pertencem *ao mesmo ato mental*.

7. Considerações finais

No meu modo de entender, a reconstituição do argumento proposta por Textor (2017, p. 55-6) se encaixa melhor com a argumentação efetivamente oferecida por Brentano (1924, p. 176-9; 1995, p. 126-7) no parágrafo oito do capítulo sobre a consciência interna. Para mostrar isso, vou comparar as premissas da reconstituição de Textor com as premissas da minha reconstituição mais literal do argumento, exposta na seção 4 acima, bem como com o próprio texto de Brentano.

Quando Textor formula a premissa P1 em termos da diferença entre perceber a percepção de um fenômeno físico ($S_2 \rightarrow S_1 \rightarrow x$) e perceber esse fenômeno ($S_1 \rightarrow x$), isso me parece corresponder a um elemento da premissa P2 da minha reconstituição, onde eu reconstruo o questionamento de Brentano sobre se (a) nós temos *várias representações*, de diferentes tipos, ou (b) nós temos *apenas uma representação*, quando estamos conscientes de um fenômeno mental. A formulação da premissa P1 por Textor corresponde à opção (a), que é justamente a opção que Brentano pretende *refutar*.

Quando Textor formula a premissa P2 em termos da percepção da percepção (S_2) sendo *diretamente* sobre a percepção (S_1) e indiretamente sobre o fenômeno físico (x), isso corresponde fielmente às considerações de Brentano, quando ele afirma que, caso nós determinemos o número e a variedade de *representações* segundo o número e a variedade de *objetos*, então teremos “diversas representações de diferentes tipos, de tal modo que uma formaria o conteúdo da outra, enquanto ela mesma teria um fenômeno físico como conteúdo” (BRENTANO, 1924, p. 177-8; 1995, p. 127). A percepção (S_1) do fenômeno físico seria o conteúdo da percepção dessa percepção (S_2), enquanto que o fenômeno físico (x) seria o conteúdo da primeira percepção (S_1).

Quando o intérprete tira dessas premissas P1 e P2 a conclusão C1 de que o objeto tem que ser representado *duas vezes*, isso também corresponde fielmente à passagem na qual Brentano alude à ideia da duplicação do objeto em Aristóteles¹⁹, a qual está representada na premissa P4 da minha reformulação do argumento.

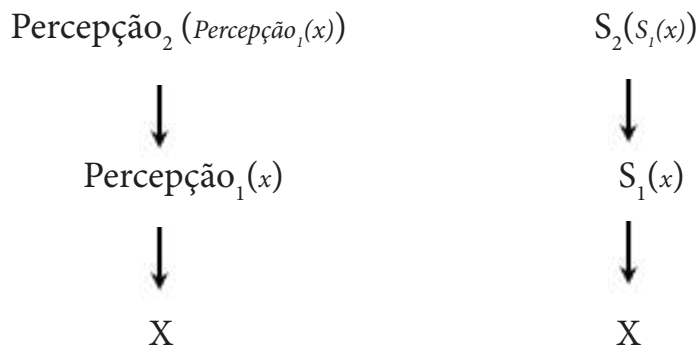
Por fim, quando Textor formula a premissa P3 em termos da negação de uma dupla representação do objeto, e disso tira a conclusão C2 de que a percepção da percepção do fenômeno físico é o mesmo ato que a simples percepção desse fenômeno, ele está sendo

¹⁹ “Pareceria, portanto, como Aristóteles já notou, resultar que o fenômeno físico teria que ser representado duas vezes” (BRENTANO, 1924, p. 178-9; 1995, p. 127).

claramente fiel ao recurso de Brentano à nossa experiência interna²⁰, bem como à conclusão de que as duas representações estão peculiarmente entrelaçadas²¹. Esses dois aspectos estão formulados na premissa P5 e na conclusão C1 da minha reconstituição mais literal do argumento de Brentano.

Com base na *fusão* [*Verschmelzung*] – cf. Brentano (1924, p. 183); Brentano (1995, p. 130) – da representação da representação do fenômeno físico com a representação desse fenômeno, Brentano entende que o regresso ao infinito está evitado, pois não há necessidade de uma segunda representação para que a representação do objeto seja consciente. Assim, obviamente, o regresso é evitado, pois não há necessidade de uma terceira, de modo que não há necessidade de uma quarta, e assim por diante.

A concepção da consciência interna que Brentano quer negar pode ser representada graficamente como se segue:



Nessas duas figuras, vemos que o objeto intencional x – correspondente ao fenômeno físico X – aparece duas vezes na estrutura complexa de percepções postulada na hipótese que Brentano quer refutar²². Porque nossa experiência imediata não nos revela dois

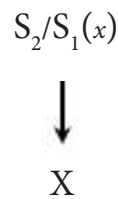
20 “Na verdade, a experiência interna parece provar de modo indubitável que a representação do som está conectada com a representação da representação de uma maneira tão peculiarmente íntima que o seu ser ao mesmo tempo contribui internamente para o ser da outra” (BRENTANO, 1924, p. 179; 1995, p. 127).

21 “Isso indica o peculiar entrelaçamento do objeto da representação interna com essa representação mesma e ao pertencimento de ambas ao mesmo ato psíquico” (BRENTANO, 1924, p. 179; 1995, p. 127).

22 É importante ressaltar que Brentano trabalha com uma noção peculiar de *objeto*, noção oriunda da tradição escolástica, influenciada por Aristóteles. Como esclarece Mauro Antonelli (2022, p. 310): “Na tradição aristotélica-escolástica o objeto (*obiectum*, ἀντικείμενον) é sempre por definição o *objeto de* uma faculdade ou de uma função, ou *de* uma atitude epistêmica ou intencional. O objeto *não* é uma coisa ou entidade; mas

objetos na experiência, ele entende que não pode haver uma duplicidade de percepções envolvida na simples percepção de um fenômeno físico.

Portanto, o resultado seria que toda percepção de um objeto inclui uma percepção dessa percepção, que torna a percepção do objeto consciente para o sujeito. Penso que essa perspectiva pode ser representada graficamente através da seguinte figura:



Aqui, a barra “/” representa uma relação de *fusão* entre essas representações perceptivas, de tal maneira que quando representamos um fenômeno físico apreendemos simultaneamente a própria representação psíquica desse fenômeno – cf. Brentano (1924, p. 179; 1995, p. 127).

Se essa concepção da consciência interna – nossa consciência de nossos próprios fenômenos mentais – é uma concepção *aceitável* , é outra questão. No presente trabalho, defendi a interpretação de que a reconstituição desse argumento por Mark Textor é mais fiel às intenções de Brentano com sua argumentação.

Referências

ANTONELLI, M. Consciousness and intentionality in Franz Brentano. *Acta Analytica*, Vol. 37, p. 301-22, 2022.

ARISTOTLE. On the soul. In: Barnes, J. (Org.). *The complete works of Aristotle: The revised Oxford translation*. Princeton: Princeton University Press, 1995 p. 641-92 (Vol. 1).

sim *se refere* a uma coisa ou entidade.” Sobre a peculiaridade da noção de objeto na tradição escolástica, e a maneira como essa noção foi usada por Descartes, cf. Cottingham (1995, p. 138-9).

ARMSTRONG, A. **A materialist theory of mind**. London: Routledge & Kegan Paul, 1968.

BRENTANO, F. **Die Psychologie des Aristoteles, insbesondere seine Lehre vom Nous Poietikos**. Mainz: Verlag von Franz Kirchheim, 1867 [edição facsimilar].

BRENTANO, F. **Psychologie vom empirischen Standpunkt**. 2ª Edição. Leipzig: Felix Meiner, Vol. I, 1924.

BRENTANO, F. **Psychology from an empirical standpoint**. London: Routledge & Kegan Paul, 1995.

CHALMERS, D. Facing up to the problem of consciousness. **Journal of Consciousness Studies**, Vol. 2, Nº 3, p. 200-19, 1995.

CHALMERS, D. Enfrentando o problema da consciência. **Prometheus: Journal of Philosophy**, Nº 33, maio-agosto, p. 319-52, 2020 [Dossiê Linguagem e Cognição]. Tradução de Pedro H. G. Muniz. Disponível em: <https://periodicos.ufs.br/prometeus/issue/view/861>

COTTINGHAM, J. **Dicionário Descartes**. Rio de Janeiro: Jorge Zahar, 1995.

FISSETTE, D. Target paper: Franz Brentano and higher-order theories of consciousness. **Argumentos**, Ano 7, Nº 13, p. 9-39, 2015 [Special issue: Brentano and philosophy of mind]. Disponível em: <http://periodicos.ufc.br/argumentos/issue/view/523>

GEORGE, R.; KOEHN, G. Brentano's Relation to Aristotle. In: JACQUETTE, D. (Org.). **The Cambridge companion to Brentano**. Cambridge: Cambridge University Press, 2004, p. 20-44.

GÜZELDERE, G. Introduction – The many faces of consciousness: A field guide. In: BLOCK; Flanagan & GÜZELDERE (Orgs.). **The nature of consciousness**. Cambridge (Massachusetts): MIT Press, 1997, p. 1-67.

HOSSACK, K. Self-knowledge and consciousness. **Proceedings of the Aristotelian Society**, Vol. 102, Nº 2, p. 163-181, 2002.

JACQUETTE, D. Introduction: Brentano's philosophy. In: JACQUETTE, D. (Org.). **The Cambridge companion to Brentano**. Cambridge: Cambridge University Press, 2004, p. 1-19.

KRANZ, S. Brentano on 'unconscious consciousness'. **Philosophy and Phenomenological Research**, Vol. 50, Nº 4 (Jun.), p. 745-53, 1990.

KRIEGEL, U. Consciousness, permanent self-awareness, and higher-order monitoring. **Dialogue**, 41, p. 517-40, 2002.

KRIEGEL, U. Consciousness as intransitive self-consciousness: Two views and an argument. **Canadian Journal of Philosophy**, Vol. 33, Nº 1, p. 103-32, Mar., 2003.

KRIEGEL, U. Naturalizing subjective character. **Philosophy and Phenomenological Research**, Vol. 71, Nº 1, p. 23-57, 2005.

KRIEGEL, U. **Subjective consciousness: A self-representational theory**. Oxford: Oxford University Press, 2009.

KRIEGEL, U. **Brentano's philosophical system: mind, being, value**. Oxford: Oxford University Press, 2018.

NAGEL, T. What is it like to be a bat? **The Philosophical Review**, Vol. 83, No. 4, p. 435-50, 1974.

NAGEL, T. Como é ser um morcego? **Cadernos de História e Filosofia da Ciência**, v. 15, n. 1, jan./jun., p. 245-62, 2005. Tradução de Paulo Abrantes e Juliana de Orione Arraes Fagundes. Disponível em: <https://www.cle.unicamp.br/eprints/index.php/cadernos/article/view/617>

PORTA, M. A. G. **Brentano e sua escola**. São Paulo: Loyola, 2014.

ROSENTHAL, D. Two concepts of consciousness. **Philosophical Studies**, Vol. 49, p. 329-59, 1986.

ROSENTHAL, D. Consciousness and its function. **Neuropsychologia**, Vol. 46, p. 829-40, 2008.

ROSENTHAL, D. Misrepresentation and mental appearance. **Trans/Form/Ação**, Vol. 41, p. 49-74, 2018 [Número especial: Consciência]. Disponível em: <https://revistas.marilia.unesp.br/index.php/transformacao/issue/view/503>

SELLARS, W. Intentionality and the mental. In: FEIGL, H.; SCRIVEN, M. & MAXWELL, G. (Orgs.). **Minnesota studies in the philosophy of science**. Minneapolis: University of Minnesota Press, p. 507-20, 1958 (Volume II).

SMITH, D. W. The structure of (self-)consciousness. **Topoi**, Vol. 2, p. 149-56, 1986.

TASSONE, B. G. **From psychology to phenomenology: Franz Brentano's psychology from an empirical standpoint and contemporary philosophy of mind**. New York: Palgrave MacMillan, 2012.

TEXTOR, M. Brentano (and some neo-brentanians) on inner consciousness. **Dialectica**, Vol. 60, p. 411-32, 2006.

TEXTOR, M. Brentano on the dual relation of the mental. **Phenomenology and Cognitive Science**, Vol. 12, p. 465-83, 2013.

TEXTOR, M. Brentano on consciousness. In: KRIEGEL, U. (Org.). **The Routledge handbook of Franz Brentano and the Brentano school**. New York, London: Routledge, 2017, p. 49-60.

THOMASSON, A. L. After Brentano: A one-level theory of consciousness. **European Journal of Philosophy**, Vol. 8, N° 2, p. 190-209, 2000.

ZAHAVI, D. Back to Brentano? **Journal of Consciousness Studies**, Vol. 11, N° 10-11, p. 66-87, 2004.

ZAHAVI, D. Two takes on a one-level account of consciousness. **Psyche**, Vol. 12, N° 2, (May), p. 1-9, 2006.





BIAS

WILLIAM JAMES, PRAGMATISMO, INTENCIONALIDADE, 4ES¹



Arthur Araújo

Departamento de Filosofia/UFES - Brasil

Resumo:

Nos termos gerais deste artigo, procuro situar o pragmatismo de William James em conjugação com sua forma de empirismo. O que resulta dessa conjugação é uma compreensão anti-representacionista de intencionalidade. Se considerada como sinônimo de intencionalidade no chamado contexto 4Es em filosofia da mente, em particular, dois sentidos de *sense-making* evidenciam uma distinção entre fenomenologia e empirismo. Enquanto o enativismo representa a perspectiva fenomenológica, chamarei de extensio-nismo radical minha abordagem da intencionalidade a qual tributo ao pragmatismo e ao empirismo James. Como conclusão da minha estratégia de situar a questão de *sense-making* (como sinônimo de intencionalidade), a ideia é que ela resulta de uma dinâmica de campo inerente ao próprio fluxo da experiência e,

Abstract:

In the general terms of this paper, I look for to situate William James' pragmatism in conjunction with his form of empirism. What results from this conjugation is an anti-representationalist understanding of intentionality. If considered as a synonym of intentionality in the so-called 4Es context in philosophy of mind, in particular, two meanings of *sense-making* indicate a distinction between phenomenology and empiricism. While enactivism represents the phenomenological perspective, I will call my approach to intentionality radical extensionism which I pay homage to James' pragmatism and empiricism. As a conclusion of my strategy of situating the issue of *sense-making* (as a synonym for intentionality), the idea is that it results from a field dynamic inherent in the very flow of experience and, therefore, incorporating an essen-

¹ Neste artigo, de modo resumido, retomo algumas partes do meu trabalho *William James, Jakob von Uexküll et la critique des notions de représentation mentale et d'intentionnalité*, publicado em *Philosophical Enquiries: Revue des Philosophies Anglophones* (2016).

portanto, incorporando um sentido essencialmente empírico de modo alternativo ao enativismo.

Palavras-chave:

William James; pragmatismo ; intencionalidade; 4ES.

tially empirical sense as an alternative to enactivism.

Keywords:

William James; pragmatism; intentionality; 4Es.

Introdução

Nos contextos de filosofia da mente e ciências cognitivas, notadamente, o pragmatismo tem sido visto como uma alternativa às concepções tradicionais da cognição fundada em na noção de representação (ver, por exemplo, *Pragmatism and Embodied Cognitive Science* editado por Roman Madzia & Matthias Jung, 2016). Em muitos aspectos, o pragmatismo incorpora a crítica ao representacionalismo no que se refere à relação mente e mundo. Em filosofia da mente e ciências cognitivas, o representacionalismo corresponde à visão de que “há estados mentais que funcionam para codificar estados do mundo” (Chemero, 2009, p. 17). Em oposição direta ao representacionalismo, encontramos “visões do tipo defendidas pelos pragmatistas americanos” (Chemero, 2009, Gallanger, 2014; Johnson, 2016; 2017).

Considerando o pragmatismo de William James na esteira da sua forma de empirismo, a crítica ao representacionalismo está notoriamente presente: “as teorias representacionais colocam uma ‘representação’ mental, ‘imagem’ ou ‘conteúdo’ na lacuna [entre pensamento e mundo] como um tipo de intermediário” (James [1907], 1977, p. 61). Em sentido amplo, pode-se tomar o pragmatismo de James como uma resposta às teorias representacionais da mente na medida em que ele deflaciona a noção de representação como condição de sustentação da intencionalidade do pensamento. Em relação aos contextos da filosofia da mente e ciências cognitivas, pode-se argumentar que o pragmatismo de James consiste em uma forma de anti-representacionalismo.

Embora o rótulo ‘anti-representacionalismo’ corresponda à versão de neo-pragmatismo de Richard Rorty em *Philosophy and the Mirror of Nature* ([1979] 2009), em que ele nega que a linguagem não seja “sobre” nada, tal perspectiva de pensamento remonta às origens do pragmatismo. Em *What pragmatism means* ([1907] 2000, p. 25), James retoma o princípio pragmático de Peirce e enfatiza o sentido de atividade do pensamento como

um processo do qual se desenvolve uma forma dinâmica de intencionalidade. Segundo James, uma concepção pragmática de intencionalidade descredita o compromisso com as ideias de representação e referência do pensamento. Do fato de que um pensamento é intencional, não se segue que ele seja representação ou referência a alguma coisa. Nesse sentido, a concepção pragmática de intencionalidade de James se afasta da tradição representacionista herdeira de Brentano (Jackman, 1998). Segundo o pragmatismo de James, a intencionalidade consiste mais em atividade do pensamento do que em representação e, desse modo, converge com a perspectiva 4E em filosofia da mente (Chemero, 2013; Chemero & Käufer, 2016; Hufendiek, 2016; Kühle, 2017; Di Francesco, Marraffa & Pater-noster, 2022).

Na vasta literatura sobre as chamadas 4Es (Menary, 2010; Rowlands, 2010; Ward & Stapleton, 2012; Newen, de Bruin & Gallagher, 2018), essa expressão resume um movimento de pensamento alcançando um domínio de estados e processos, inclusive pessoas, que, literalmente, se estendem ao ambiente” (Robbins & Aydede, 2009, p. 1); “pensar e conhecer podem (às vezes) depender direta e não instrumentalmente do trabalho contínuo do corpo e/ou do ambiente extra-orgânico” (Clark, 2011, p. xxviii); e “a cognição é “constituída” por processos corporais e ambientais” (Newen, de Bruin & Gallagher, 2018, p. 4)². De acordo com a perspectiva 4E, como parte da atividade cognitiva, a intencionalidade se estende além do limite da estrutura corporal e incorpora uma dinâmica organismo-ambiente. Uma tal forma de intencionalidade deveria ser considerada como “não-derivada” no sentido de que se trata de um engajamento prático entre organismo e ambiente (Newen, de Bruin & Gallagher, 2018, p. 10). Um engajamento prático que se pode traduzir nos termos do pragmatismo de William James.

No desenvolvimento do seu itinerário intelectual, James alia pragmatismo e empirismo e, no entanto, a aliança não transige com as formas tradicionais de empirismo (Bacon, 2012, p. 36). A radicalização da noção de experiência assinala a concepção do pragmatismo como consequência conceptual do empirismo radical: as distinções e divisões entre a mente e o mundo são fatos operacionais, instrumentais ou convencionais (ver James, [1905] 1977, p. 194). Se o pragmatismo de James nos ensina uma lição, certamente, é que não precisamos de entidades ou representações mentais a fim de preencher a suposta lacuna entre o interior (subjetivo) e o exterior (objetivo). Para James, de fato, não há lacuna alguma ou duplicação interna na experiência.

2 A expressão 4E significa “a ideia de que os processos mentais são (1) incorporados [embodied], (2) situados [embedded], (3) enativos [enacted] e (4) estendidos [extended] (Rowlands, 2010, p. 3).

Segundo o não-representacionalismo de James, na medida em que a intencionalidade se refere à ação e aos efeitos práticos do pensamento, ela se desenvolve no *continuum* da experiência da qual emerge a diferenciação operacional, instrumental ou convencional entre mente e mundo. Nesses termos, segundo James (1978 [1907], p. 245), a intencionalidade é mais um processo transicional (ou “ambulatorial”) do que um “salto” do pensamento em direção ao objeto. Com um traço nitidamente anti-dualista, o princípio do pragmatismo de William James e Peirce pode enriquecer os debates em torno da noção de intencionalidade não-derivada no contexto 4E na filosofia da mente. Este é o aspecto central que procurarei desenvolver última parte deste artigo.

Como assinala Mark Rowlands (2009), de acordo com a visão que se tornou conhecida como ‘mente estendida’ no contexto 4Es em filosofia da mente, alguns processos mentais se estendem ao ambiente do organismo cognoscente na medida em que são compostos (parcialmente) de processos executados por esse sujeito em ambientes adequados. Em sentido convergente com essa visão, os enativistas entendem que os processos mentais são também (parcialmente) constituídos pela capacidade sensório-motora do organismo de agir, de forma apropriada, em determinados ambientes. Dadas as semelhanças evidentes entre enativismo e extensionismo, é comum considerá-los variações de um mesmo tema. No entanto, segundo Rowlands (2009), as semelhanças entre enativismo e extensionismo são relativamente superficiais e as divergências se mostram mais profundas do que se pensa. Nos termos em que estou desenvolvendo este artigo, partilho dessa compreensão de Rowlands.

Do meu ponto de vista, se considerada o tema da intencionalidade no contexto 4Es em filosofia da mente, o que os enativistas entendem como *sense-making* (Di Paolo, Rohde & De Jaegher, 2010), a transação organismo-ambiente, teria um caráter derivado da rede de relações dinâmicas integrando o trinômio ‘corpo-mente-ambiente’³. Partindo da conjugação entre o pragmatismo e a forma de empirismo de William James em paralelo com a noção de mente estendida, chamarei de *extensionismo radical* minha abordagem da intencionalidade⁴. O que está em questão é um embate entre enativismo e extensionis-

3 Não traduzirei aqui o termo ‘sense-making’ por ‘fazer-sentido’. Entendo que, assim como em inglês, esse termo já tem uma significação técnica instituída e, portanto, ele dispensa tradução.

4 Não somente ao enativismo se outorga ser radical. Ver, por exemplo, *Radicalizing Enactivism* (Hutto & Myin, 2013). No contexto das 4Es em filosofia da mente, por ‘extensionismo radical’, entendo a conjunção entre o pragmatismo e o empirismo de James. Isso significa que, por exemplo, a dissolução do chamado problema da constituição: ou seja, “se, modo bastante geral, a questão é se, e em caso afirmativo, como somos capazes de decidir (seja empiricamente, pragmaticamente ou a priori) se um processo cognitivo particular é constituído por ou meramente dependente de processos extracranianos ou processos extracorporais” – tra-

mo em torno da noção de intencionalidade e, conseqüentemente, entre dois sentidos de *sense-making*. Se, em termos enativistas, a atividade de *sense-making* expressa o sentido fenomenológica da autonomia do organismo, segundo minha compreensão de um extensionismo radical, ela significa um aspecto derivado da dinâmica da experiência.

Como ideia geral deste artigo, o que tenho em mente é deflacionar o sentido fenomenológico de *sense-making* e outorgar à intencionalidade um sentido pragmático e empírico nos termos de James. Para isso, na última parte do artigo, exploro a noção de ‘campo’ tal qual ela é entendida por James como núcleo da sua forma radical de empirismo. Em linhas gerais, considerando o sentido de *sense-making* como sinônimo de intencionalidade, como background filosófico, o que se apresenta é um embate entre fenomenologia e empirismo⁵.

Intencionalidade e representação

Desde Franz Brentano, historicamente, a noção de intencionalidade seguiu duas linhas de desenvolvimento: 1) a partir de uma perspectiva fenomenológica, a intencionalidade é considerada a marca do mental (Husserl e Sartre); e 2) em uma retomada do termo por filósofos analíticos (em particular, Roderick Chisholm e Wilfrid Sellars nas décadas de 1960 e 1970), a intencionalidade é a expressão de uma propriedade linguística das sentenças (ver Jacob, 2004, particularmente, a Segunda Parte). Embora sendo duas linhas de desenvolvimento distintas, ambas sustentam a co-extensividade entre intencionalidade e representação: *sendo sobre (aboutness)* objetos e estados de coisas no mundo, os estados mentais possuem um conteúdo representacional e podem ser expressos por atitudes proposicionais.

dução minha (Newen, de Bruin & Shaun Gallagher, 2018, p. 7). Segundo o pragmatismo e a forma radical de empirismo de James, nem constituição nem dependência! A ideia é que, em um primeiro momento, se diferenciam as partes ‘interior’ e ‘exterior’ (ao organismo) do próprio fluxo da experiência; e, em um segundo momento, por um processo de transição contínua entre essas partes, se desenvolvem as “relações cognitivas”.

5 “A cognição ou *sense-making* é um engajamento intencional e normativo do sistema com seu ambiente. Uma das proposições básicas da abordagem enativa é que ser autônomo é uma condição necessária para o sistema incorporar intencionalidade e normatividade originais” (Thompson & Stapleton, 2009, p. 6 – tradução minha). E “de acordo com a visão enativista, os seres vivos são seres *sense-making* em íntima relação com seus ambientes. Aqui está como [Francisco] Varela coloca a esta ideia: “o modo no qual os seres vivos fazem sentido” – estas palavras têm um duplo sentido. De um lado, elas se referem a como os seres vivos procedem com suas atividades de *sense-making* [...] De outro, elas se referem a como entendemos que os seres vivos fazem sentido a nós (Thompson, 2011, p. 114).

Em sua forma canônica, como os estados mentais *são sobre*, o verbo ‘crer’ expressa a estrutura representacional da atitude proposicional. Em ‘A crê que P’, o verbo indica o *ser sobre* (*aboutness*) de um agente cognitivo (A) em relação a uma proposição (P) que representa o conteúdo mental do agente: por exemplo, ‘A crê que as águias são aves de rapina eficientes’ (P). Nesse caso, como uma proposição é semanticamente avaliável, ela pode representar corretamente a relação entre estado mental e mundo. No contexto da chamada ‘psicologia popular’ (*folk psychology*), por exemplo, os objetos dos estados mentais são entidades linguísticas como proposições, propriedades, relações, etc. O que determina a significação de um conteúdo mental é a proposição que o representa cuja propriedade é ‘intensional’ (com s) ou ‘não-extensional’: “sentenças que descrevem estados intencionais são intencionais” (Zalta, 1988, p.13; ver também Jacob, 2004, p. 120)⁶. Seguindo o desenvolvimento das duas linhagens brentanianas entre fenomenologia e filosofia analítica da mente, se evidencia uma relação intrínseca entre as noções de representação e intencionalidade: ou seja, explicar a intencionalidade responde à questão de como os estados mentais podem *ser sobre* e ter conteúdo representacional.

Em um paralelo entre o representacionalismo em filosofia da mente e a segunda geração das ciências cognitivas dos anos 70 e 80, segundo o que Francisco Varela chama de ‘hipótese cognitivista’, a reciprocidade entre intencionalidade e representação se apresenta nos seguintes termos:

A noção em jogo aqui é representação, ou intencionalidade, o termo filosófico para a qualidade do que é “sobre algo”. O argumento cognitivista é que o comportamento inteligente pressupõe a capacidade de representar o mundo de alguma forma [...] Essa compreensão da representação é [...] a seguinte: a hipótese cognitivista afirma que a única maneira de explicar a inteligência e a intencionalidade é postular que a cognição consiste em agir com base em representações que têm uma realidade física na forma de código simbólico no cérebro ou em uma máquina (Varela, 1988, p. 37-8 – tradução minha).

Convergindo com as raízes brentanianas, em termos cognitivistas, as noções de representação e intencionalidade significam:

6 Em *Perceiving* (1957), de modo exemplar, Roderick Chisholm propõe uma interpretação linguística do sentido brentiano de intencionalidade em que sentenças descrevendo estados intencionais (com c) são intencionais (com s) ou não-extensionais. Segundo Chisholm, trata-se de um critério de distinção em que as sentenças de uma linguagem são usadas para descrever os fenômenos intencionais. Como consequência, Chisholm deriva o exame da intencionalidade (com c) do exame das sentenças descritivas e atributivas de intencionalidade (com s).

- (a) Intencionalidade: propriedade de um estado mental de *ser sobre* um objeto ou estado de coisas no mundo. O estado mental torna-se significativo de acordo com essa propriedade.
- (b) Representação: uma realidade interna (pensamentos, ideias, conceitos, etc.) representa uma realidade externa

A princípio, a questão da reciprocidade entre representação e intencionalidade pode ser entendida em dois sentidos. Enquanto o primeiro sentido está relacionado ao estudo científico da cognição, o segundo é filosófico. Os dois sentidos estão inter-relacionados. O problema é, contudo, explicar como um tipo de realidade interna pode *ser sobre* alguma coisa no mundo. Face a tal problema, do ponto de vista representacionista em filosofia da mente e na segunda geração das ciências cognitivas dos anos 70 e 80, a solução consiste em adicionar um tipo especial de intermediário entre o pensamento e o objeto. Segundo a chamada hipótese cognitivista clássica (acima descrita por Varela), para dar conta de explicar a cognição é necessário postular, de forma simbólica, *representações mentais* (na constituição das atitudes proposicionais) que sejam a base da própria cognição: ou seja, representações que constituam a base da relação entre o interior do organismo (ou máquina) e o mundo exterior. O problema consiste, de novo, em explicar como representações mentais podem *ser sobre* alguma coisa no mundo (Chemero, 2011, p. 30).

Uma alternativa de lidar com esse problema talvez seja, como assinala Francisco Varela (1988, p. 98), assumirmos que *exeunt* as representações. Ou seja, assumindo um movimento de mudança na perspectiva filosófica, reconsiderar pragmaticamente nossa compreensão da intencionalidade fora da sombra da representação.

O princípio do pragmatismo de James e Peirce

Segundo Peirce, James apresenta o pragmatismo como um “método” cujo objetivo é tornar as ideias claras determinando suas consequências práticas:

O termo [pragmatismo] é derivado de $\pi \alpha \gamma \mu \alpha$, que significa ação, de onde vêm nossas palavras ‘prática’ e ‘prático’. Foi introduzido pela primeira vez na filosofia pelo Sr. Charles Peirce em 1878. Em um artigo intitulado ‘Como tornar nossas ideias claras’ [...] o Sr. Peirce, depois de apontar que nossas crenças são

realmente regras para a ação, disse que, para desenvolver a significação de um pensamento, precisamos apenas determinar qual conduta ele está apto a produzir: essa conduta é, para nós, a única significação [...] Este é o princípio de Peirce, o princípio do pragmatismo James, [1907] 2000, p. 25 – tradução minha).

Ao ampliar o escopo do princípio de Peirce, apresentado em *How to make clear our ideas* ([1878] 1966, p. 123) ([1878] 1966, p. 123), em *What Pragmatism means* ([1907] 2000, p. 25), William James considera seus efeitos psicológicos na vida cotidiana além de ser unicamente uma restrição lógica em nossa compreensão de um conceito. Apesar das divergências quanto à compreensão do princípio do pragmatismo por Peirce e James, ambos parecem concordar que a própria noção de significação não depende da exigência de representação mental (em termos de conteúdo proposicional) e a significação consiste antes em um processo baseado em efeitos práticos orientando nossos pensamentos⁷. Para James ([1907] 2000, p. 25-6), em particular, o pragmatismo (como método) procura “interpretar cada noção traçando sua consequência prática”.

Em se considerando o princípio do pragmatismo de James e Peirce, em particular, a ideia é que os termos psicológicos (como crenças, por exemplo) possam ter suas significações alteradas temporariamente e, portanto, não funcionariam como codificação dos estados mentais⁸. Para James, o princípio do pragmatismo reside em seus efeitos práticos: “para desenvolver a significação de um pensamento, precisamos apenas determinar qual conduta ele está apto a produzir: essa conduta é para nós a única significação” (James

7 É importante, contudo, não se confundir as respectivas compreensões de Peirce e James sobre o pragmatismo (ver Wiener, 1966, p. 181). Por muitas vezes, o próprio Peirce insistiu que sua compreensão do pragmatismo tinha sido mal compreendida. Para expressar a definição original, Peirce (CP 5: 414) propõe o uso do termo ‘pragmaticismo’ na medida em que, segundo ele, esclarecer o significado de um conceito é identificar o hábito que ele produz. No que diz respeito à incorporação de James do princípio do pragmatismo de Peirce, de fato, ele concede a ele um uso mais abrangente no curso da experiência humana. Seguindo essa linha de pensamento e incorporando a incorporação de James do princípio pragmático de Peirce, neste artigo, usarei a expressão ‘o princípio pragmático de Peirce e James’.

8 Na filosofia da mente, as teorias representacionais assumem que as atitudes proposicionais são codificações dos estados mentais e, em termos gerais, expressam termos da psicologia popular. O termo ‘crença’ é considerado o exemplo canônico de atitude proposicional. Em ‘A acredita p’, por exemplo, o verbo ‘crer’ expressa a relação de um agente cognitivo (A) e uma proposição (P) que representa o conteúdo mental atribuído ao agente: ‘João (A) acredita que as águias são aves de rapina eficientes’ (P). Como as proposições são semanticamente avaliáveis, elas podem representar corretamente o mundo. Levando em conta o princípio pragmático de James, dificilmente se pode pensar na significação dos termos mentais como sendo representados por atitudes proposicionais. No caso das crenças, pragmaticamente, a significação está longe de expressar um conteúdo representacional e é muito mais uma referência à ação e aos efeitos práticos das nossas ações de pensamento. Para Peirce, aliás, uma crença estabelece uma ‘regra de ação’ ou um ‘hábito’: “crença é uma regra de ação [e] a essência da crença é o estabelecimento de um hábito, e diferentes crenças se distinguem pelos diferentes modos de ação a que dão origem” (Peirce, [1878] 1966, p. 121 – tradução minha).

[1907] 2000, p. 25). De acordo com a concepção de James de “método pragmático”, cada noção deve ser interpretada traçando suas respectivas consequências práticas” (James [1907] 2000, p. 25). Em *Will to Believe*, por exemplo, esse aspecto metodológico do pragmatismo é sistematicamente posto em prática por James.

Seguindo o princípio pragmático de Peirce, James assinala a relevância da noção de consequências práticas no contexto das comunidades humanas. Em *A vontade de crer*, dedicado a Peirce e no qual James apresenta o método pragmático (como teoria da significação), ele afirma que, em algumas circunstâncias, alguém está justificado a acreditar em uma proposição, mesmo que haja poucas evidências para isso. Nesse caso, o mais importante é que a crença adquira uma significação dinâmica uma vez que é socialmente justificada em uma comunidade. Na visão de James, portanto, as crenças incorporam valor efetivo como sendo ferramentas úteis na vida humana que permitem lidar com a realidade: “uma ideia é ‘verdadeira’ desde que acreditemos que é benéfica para nossas vidas [e] seria melhor para nós acreditar nessa ideia, a menos que a crença nela incidentalmente colidisse com outros benefícios maiores da vida” (James, [1907] 2000, p. 37-8).

De acordo com o princípio pragmático de James e Peirce, uma vez que a significação dos nossos pensamentos adquire uma forma dinâmica, as crenças estão longe de serem representações estáticas do mundo e indicam um sentido muito mais prático. Se alguém acredita que ‘as águias são aves de rapina eficientes’, isso certamente tem “consequências práticas satisfatórias que se espalham holisticamente em toda uma comunidade” (Malachowski, 2013, p. 43): isto é, a crença “as águias são aves de rapina eficientes” está holisticamente mesclada em um *continuum* experiencial de fato, significação e contexto.

No seu capítulo seminal *The Stream of Thought in Principle of Psychology*, antecipando o princípio pragmático, James ([1890] 1983, p. 250) torna explícito uma forma de pensamento-processo sobre a significação: “Pois o importante sobre uma linha de pensamento é a sua conclusão. Essa é a significação ou, como dizemos, o tema do pensamento”. Nesse mesmo capítulo, ademais, James ([1890] 1983, p. 255) afirma que a ‘significação’ de uma palavra, tomada dinamicamente, em uma sentença se torna bem diferente quando tomada estaticamente ou fora de contexto. Em ‘Colombo descobriu a América em 1492’, conforme observado por James, a significação está em toda a frase nem mais nem menos (James, [1890] 1983, p. 265). Ao criticar a psicologia associacionista, em particular, James argumenta que a significação de ‘Colombo descobriu a América em 1492’ não pode ser reduzido a ideias individuais como ‘Colombo’, ‘América’ ou ‘a descoberta da América’: ou seja, a significação está no processo de desenvolvimento da sentença como um todo.

Como uma importante lição tirada do princípio pragmático de James (e Peirce), uma vez que não há significação intrinsecamente determinada nos nossos pensamentos, a noção mesma de significação se insere e se desenvolve em diferentes contextos das nossas condutas. Se considerado o princípio pragmático, os termos psicológicos ('crença', por exemplo) ganham significação de acordo com possíveis diferenças práticas no curso das nossas ações de pensamento ao invés de serem representações codificadas por atitudes proposicionais.

Intencionalidade como processo: William James e a deflação pragmática da reciprocidade entre intencionalidade e representação

No contexto 4Es em filosofia da mente, notadamente, observa-se uma distensão entre duas perspectivas: de um lado, o enativismo; de outro, a visão de mente estendida. Desde Francisco Varela (*et al.*, 1993), enquanto o enativismo se desenvolve segundo uma perspectiva fenomenológica, o extensionismo tende a gravitar em torno de um núcleo dinamista (Clark, 2011; Chemero & Käufer, 2016; Rowlands, 2018). Embora os enativistas reconheçam a cognição como parte de um processo dinâmico que cria e sustenta tanto o organismo quanto o ambiente ao qual ele está respondendo, eles insistem na diferença entre “incorporação” e “a simples extensão na relação corpo-mente-ambiente” (Thompson & Stapleton, 2009, p. 24). Segundo a perspectiva enativista em linha com a tradição fenomenológica de Merleau-Ponty, a atividade de ‘sense-making’ seria sinônimo de enação. A ideia é que os organismos transformam o mundo em valores e significações ambientais (Thompson & Stapleton, 2009, p. 25).

Latente nessa ideia de transformação do mundo em valores e significações ambientais, está uma compreensão fenomenológica de intencionalidade: uma forma de intencionalidade que se designa “não-derivada” no sentido de ser uma construção entre movimentos corporais hábeis em conjunto com as demandas ambientais (Newen, de Bruin & Gallagher, 2018, p. 10). Do ponto de vista extensionista, contudo, o trinômio ‘corpo-mente-ambiente’ integra uma rede de relações em que intencionalidade resultaria mais da dinâmica de um processo do que a simples extensão da mente em relação ao ambiente. Na minha compreensão, o que os enativistas entendem como ‘sense-making’ (Di Paolo, Rohde & De Jaegher, 2010, p. 39) a transação organismo-ambiente teria um caráter derivativo da rede dinâmica de relações integrando o trinômio ‘corpo-mente-ambiente’. Somente na superfície enativismo e extensionismo são similares e as divergências são mais profundas do que se supõe (Rowlands, 2009). Em convergência com o extensionismo em

filosofia da mente, o aspecto derivativo da intencionalidade me parece estar no cerne do pragmatismo de William James.

Segundo James, se considerado que o termo ‘pragmatismo’ significa ‘ação’ e que a significação do pensamento reside em seus efeitos práticos, deve-se notar que o pragmatismo se opõe a uma perspectiva representacionista da intencionalidade na medida em que é uma filosofia da ação (Kilpinen, 2009). Em consequência, as diferenças são claras entre a compreensão de intencionalidade de James e o representacionismo de Husserl (incluindo seu mestre Brentano). Na teoria da significação de James, conforme apresentada nos seus textos sobre o pragmatismo, a crítica à noção de representação mental é flagrante:

Ao longo da história da filosofia, o sujeito e seu objeto foram tratados como entidades absolutamente descontínuas [...] As teorias representativas colocam uma “representação”, “imagem” ou “conteúdo” mental na lacuna, como uma espécie de intermediário. (James, [1907] 1978, p. 61 – tradução minha).

Se considerada a questão da intencionalidade, fica evidente que uma abordagem pragmática de pensamento recusa a condição da representação mental: tendo como ponto de partida a dissolução de uma concepção internalista de mente, a representação mental já não é mais entendida com a condição de ser da intencionalidade.

É a título de uma teoria da significação que o pragmatismo de James pode ser visto como uma ferramenta de reconsideração conceitual da intencionalidade. Além disso, James não negligencia a intencionalidade. O que ele desenvolve é uma deflação da noção britânica da inexistência intencional (James, [1909] 2000, p. 144). De acordo com o princípio do pragmatismo de James (e Peirce), explicar a intencionalidade do pensamento consiste em explicar como ele se dirige à realidade e também em mostrar que a significação da intencionalidade reside nos efeitos práticos aos quais o pensamento conduz. Fica claro aqui que uma concepção pragmática da intencionalidade estabelece uma clara diferença entre significação e representação: na medida em que a significação da intencionalidade reside nos efeitos práticos do pensamento, ela não tem como condição ser representação em qualquer sentido.

Cabe aqui reconsiderar a aplicação do princípio do pragmatismo de James (como teoria da significação) e analisar em que medida ela fornece elementos para uma concepção pragmática da intencionalidade. O primeiro ponto que merece ser enfatizado é que a significação de um pensamento está em seus efeitos práticos e, portanto, os

pensamentos não têm significação intrínseca. O segundo elemento que vale destaque é que a significação da intencionalidade é uma referência à ação do pensamento. Nesse sentido, de acordo com o princípio do pragmatismo de James, a intencionalidade é mais uma questão de dinâmica do pensamento do que representação. Em *Some Metaphysical Problems Pragmatically Considered*, apresentação do método pragmático, James explora a propriedade disposicional do pensamento e exclui a ideia de significação intrínseca. Na citação de James a seguir, vemos a dimensão disposicional dos pensamentos como modos de significação:

O giz, a madeira e a lã mostram novamente, apesar de suas diferenças, propriedades comuns, e até agora eles próprios são contados como modos de uma substância ainda mais primordial, a matéria, cujos atributos são a ocupação do espaço e a impenetrabilidade. Da mesma forma, nossos pensamentos e sentimentos são afeições ou propriedades de nossas várias almas, que são substâncias, mas novamente não totalmente em si mesmas, pois são modos da substância ainda mais profunda 'espírito' [...] Um grupo de atributos é o que cada substância aqui é conhecida como, elas formam seu único valor efetivo para nossa experiência real (James, [1907] 2000, p. 41 – tradução minha).

Na medida em que os pensamentos são modos, eles se mostram como disposições e não episódios de representação. Enquanto os episódios indicam uma realidade interna, as disposições são referências a relações externas de significação. A noção de disposição, portanto, consiste em um modo de relação externa de significação entre pensamento e objeto. Assim, do ponto de vista de James, a intencionalidade indica mais uma dinâmica do pensamento do que um episódio interno de representação. Aqui, oportunamente, podemos comparar a citação anterior de James com o que Gilbert Ryle chama de 'disposição' em contraste com a ideia de um episódio interno de representação da mente. Em seu *The concept of mind* ([1949] 2009), Ryle argumenta que a mente não é uma coisa dotada da propriedade de representar o mundo. Para ele, o que determina a significação dos conceitos mentais é uma disposição de acordo com um modo particular em uma circunstância apropriada e não como indicação da existência de episódio interno:

Quando descrevemos o vidro como quebradiço ou o açúcar como solúvel, estamos usando conceitos disposicionais, cuja força lógica é esta. A fragilidade do vidro não consiste no fato de ele estar em um determinado momento sendo realmente pressionado. Ele pode ser quebradiço sem nunca ter sido trepidado. Dizer que é quebradiço é dizer que, se alguma vez for, ou se alguma vez tiver sido, golpeado ou forçado, voaria, ou teria voado, em fragmentos. Dizer que o açúcar é solúvel é dizer que ele se dissolveria, ou teria se dissolvido, se imerso em água [...] O mesmo é verdade sobre disposições especificamente humanas, como qualidades de caráter. O fato de eu ser um fumante habitual não implica

que eu esteja fumando neste ou naquele momento; é minha propensão permanente fumar quando não estou comendo, dormindo, dando palestras ou participando de funerais, e não tenho fumado recentemente (Ryle, [1949] 2009, p. 31 – tradução minha).

Tendo apresentado o método pragmático também como uma teoria da significação, James o aplica à explicação de questões metafísicas tradicionais como, por exemplo, a de saber se a realidade é matéria ou espírito. James se esforça para mostrar que essas questões devem ter significação prática e diferença na vida humana. Seguindo essa linha de raciocínio, podemos aplicar o método pragmático à explicação do conceito de intencionalidade. Se não podemos ver no conceito de intencionalidade nenhuma consequência ou diferença prática nos nossos pensamentos, então ela não significa nada concretamente, o que equivale a dizer que é ela apenas pura abstração intelectual.

Aqui, novamente, segundo o princípio do pragmatismo de James (e Peirce), a ideia de que a significação de um conceito reside em consequência prática dos nossos pensamentos encontra um equivalente na concepção de Ryle sobre a significação conceitual:

Quando um objeto é descrito como duro, não queremos dizer apenas que ele resistiria à deformação; queremos dizer também que ele emitiria, por exemplo, um som agudo se atingido, que nos causaria dor se tocássemos em contato com ele, que objetos resistentes ricocheteariam nele e assim por diante indefinidamente (Ryle, [1949] 2009, p. 32 – tradução minha).

Considerando que a significação do pensamento reside em disposição, é importante indicar aqui que, em *The tigers in India*, James empreende uma deflação do conteúdo representacional da intencionalidade. Esse texto é a segunda parte de *The meaning of truth* ([1909] 1978) que se seguiu às conferências sobre pragmatismo (1907). James não nega a intencionalidade do pensamento. O que ele procura mostrar é que a noção britânica de inexistência intencional carece de sentido visto que a intencionalidade seria apenas um intermediário entre pensamento e objeto. Para James, a aceitação da intencionalidade não significa que a relação entre pensamento e objeto pressupõe um intermediário como um tipo de representação mental. Embora de forma latente, a primeira parte de *The tigers in India* mostra uma concepção pragmática da intencionalidade que, por exemplo, se opõe às teorias da representação mental em filosofia da mente.

Na primeira parte de *The tigers in India*, apesar da referência indireta a Brentano (2005 [1874]), James (2000 [1909], p. 142-143) argumenta que a intencionalidade não implica as noções de ‘presença na ausência’, ‘autotranscendência’ ou ‘inexistência intencional’ dos

conteúdos mentais. James rejeita qualquer intencionalismo representacionista. Do ponto de vista representacionista na filosofia da mente, a ideia significa que os pensamentos são representados por atitudes proposicionais e são referências a objetos ou estados de coisas no mundo. James defende, ao contrário, uma concepção de intencionalidade segundo a qual, mesmo no caso em que os pensamentos se dirigem a objetos, sua significação não é, de forma alguma, um tipo de representação que possa ser representada por atitudes proposicionais (Jackman, 1998, p. 1-2).

Se considerarmos a intencionalidade sob um ângulo pragmático, certamente, ela não terá uma significação brentiana ou neobrentiana. Na medida em que James deflaciona drasticamente o sentido transcendental da intencionalidade, ela resulta mais da função de ajustamento [*fitting*] entre pensamento e objeto segundo a dinâmica da experiência: ou seja, a transição entre estado mental e objeto consiste em um processo intra-experiencial:

O que queremos dizer com conhecer os tigres [na Índia] é apontar mentalmente para eles enquanto nos sentamos aqui [...] Em tudo isso, não há autotranscendência em nossas imagens mentais tomadas por si mesmas. Eles são um fato fenomenal; os tigres são outros; e apontar para os tigres é uma relação intra-experiencial perfeitamente corriqueira (James, [1909] 2000, p. 142-143 – tradução minha).

Para James, portanto, a intencionalidade denota os modos pelos quais o pensamento pode conhecer objetos fora da mente. Nesse sentido, a intencionalidade é mais uma ação do pensamento sobre o objeto do que uma representação. Que os pensamentos tenham uma representação subjacente a eles em termos da inexistência intencional, isso não parece determinar uma condição de significação da intencionalidade. A significação de visar os tigres na Índia reside nos efeitos práticos das proposições que não contradigam nossos pensamentos sobre tigres em termos gerais:

O apontar de nossos pensamentos para tigres é conhecido simples e unicamente como uma procissão de associações mentais e consequências motoras que seguem o pensamento e levariam harmoniosamente, se seguidas, a alguns contextos ideais ou reais, ou mesmo à presença imediata de tigres. É conhecido como nossa rejeição a uma onça, se essa besta nos fosse mostrada como um tigre; como nosso consentimento para um tigre genuíno, se assim for demonstrado. É conhecida como nossa capacidade de proferir todos os tipos de proposições que não contradizem outras proposições que são verdadeiras para os tigres reais (James, [1909] 2000, p. 142-3 – tradução minha).

Nos termos de James, a intencionalidade significa um processo de ajustamento [fitting] do pensamentos ao objeto que não se determina por representações mentais. É uma concepção de intencionalidade que segue o princípio do pragmatismo segundo o qual a significação (intencional) dos pensamentos é mais a referência a um processo do que a representações internas. Segundo James, a relação intencional entre pensamento e objeto é uma operação accidental de ajuste externo e não significa um tipo misterioso de representação na mente:

Apontar significa aqui uma operação tão externa e accidental quanto qualquer outra que a natureza produz [...] Uma pedra em um campo pode “encaixar”, dizemos, em um buraco em outro campo. Mas, a relação de “encaixe”, desde que ninguém carregue a pedra até o buraco e a jogue, é apenas um nome para o fato de que tal ato pode acontecer. Da mesma forma com o conhecimento dos tigres aqui e agora (James, [1909] 2000, p. 136 – tradução minha).

Segundo uma concepção pragmática de intencionalidade, James deflaciona as noções Brentanianas e neobrentanianas de representação do conteúdo mental. Segundo ele, a intencionalidade é antes um “apontar a” [pointing to] do pensamento que dispensa a duplicação do conteúdo mental entre o ato de visar e o objeto visado. Para James, em resumo, na medida em que a intencionalidade se refere concretamente a um processo dinâmico de ajustamento pensamento-objeto, em nada ela tem a ver com representação. Em convergência com a perspectiva 4Es, a ideia é que uma concepção pragmática de intencionalidade dispensa a noção de representação mental como intermediário entre mente e mundo.

James, pragmatismo, intencionalidade, 4Es

O principal tópico que explorarei nesta última parte do artigo é a contribuição do pragmatismo de William James, bem como sua forma de empirismo, à questão da intencionalidade em relação à noção da mente estendida (Clark & Chalmers 1998; Rupert, 2009; Rowlands 2010; Clark 2011). Em sua concepção de pragmatismo, claramente, James desenvolve uma teoria da significação (como “método pragmático”) ao enfatizar que a significação do pensamento está na ação à qual ele conduz. Sendo uma teoria da significação, em particular, o pragmatismo de James expressa uma concepção de intencionalidade.

Neste último tópico, em resumo, vou desenvolver o que chamo de *extensionismo radical* minha abordagem da intencionalidade. O que está em questão é, assim como

entendo, um embate entre enativismo e extensionismo em o torno do sentido de intencionalidade no contexto 4Es em filosofia da mente⁹. Na minha compreensão, entre significativas diferenças, os respectivos sentidos de *sense-making* adotados por enativistas e extensionistas expressam uma distensão filosófica mais aguda. Segundo os enativistas (Di Paolo, Rohde & De Jaegher, 2010, p. 39), a transação organismo-ambiente consiste em uma forma de *sense-making* como “criação e apreciação de significação”. Latente nessa concepção enativista de *sense-making* desde Varela *et al.* (1993), encontra-se um sentido de intencionalidade descendente da fenomenologia de Merleau-Ponty.

O que vou desenvolver aqui, contrariamente ao enativismo, é uma abordagem do sentido de *sense-making* como derivado do processo dinâmico integrando uma rede de relações no trinômio ‘corpo-mente-ambiente’. Segundo o pragmatismo e a forma de empirismo de James, a ideia é que, como sinônimo de *sense-making*, a intencionalidade resulta de um processo seletivo emergindo da dinâmica da experiência. Como ilustração do embate entre enativismo e extensionismo, considerarei duas possíveis compreensões do sentido de *sense-making*.

Em *Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories* (2009), clara e precisamente, Evan Thompson e Mog Stapleton exploram as diferenças entre a abordagem enativista e a tese da mente estendida. Como noções nucleares no artigo, Thompson e Stapleton consideram autonomia e *sense-making*. Segundo, não se pode assimilar “incorporação” como simples extensão na relação corpo-mente-ambiente. Como ilustração dessa distinção, eles exploram a situação de bactérias móveis nadando e escalando um gradiente alimentar de açúcar:

As células giram até atingirem uma orientação que aumenta sua exposição ao açúcar, ponto em que nadam para frente, gradiente ascendente, em direção à zona de maior concentração de açúcar (Thompson & Stapleton, 2009, p. 24 – tradução minha).

Para Thompson & Stapleton (2009, p. 25), diante desse quadro, o açúcar não tem significação ou valor intrínsecos e, contudo, ele se torna significativo em função da organização autônoma do metabolismo bacteriano. A ideia é, em mundo abundante de elementos ambientais, somente alguns desses elementos serão transformados em itens

⁹ Sobre intencionalidade e 4Es em filosofia da mente, ver Rietveld, Denys & van Westen (2018); Rowlands (2018).

significativos pela atividade de *sense-making* do organismo (Thompson & Stapleton, 2009, p. 25): “*sense-making* é o lado interacional e relacional da autonomia”. Tomando como referência a fenomenologia de Merleau-Ponty, Thompson e Stapleton entendem que a atividade de *sense-making* é comportamento ou conduta na qual o organismo ele mesmo *enage* e cria um mundo de significação e valor. Desse modo, conforme a força da capacidade corporal, os organismos vivos incorporam processos ambientais em uma rede operacionalmente fechada que constitui sua própria autonomia. Na medida em que a atividade de *sense-making* representa um processo de significação incorporada, ela indica o sentido de uma intencionalidade não-derivada como o acoplamento orientado na transação organismo-ambiente. Este me parece ser o quadro que resume o sentido enativista de *sense-making*.

O que exploro a seguir é o sentido extensionista de *sense-making* segundo o pragmatismo e a forma radical de empirismo de William James. O que os enativistas consideram ser a atividade de *sense-making* a expressão fenomenológica da autonomia do organismo, na minha compreensão, significa um aspecto derivado da dinâmica da experiência e a que chamo de *extensionismo radical*. Partindo de um ponto de vista pragmático e empírico nos termos de James, o que tenho em mente é a deflação do sentido fenomenológico de *sense-making*.

Em *A world of pure experience* (1904), William James insiste em caracterizar a experiência como um processo no tempo. Assim como a forma do campo visual, James afirma que o campo da experiência (no sentido de mundo experienciado) é igualmente “franjado” e “sem limites bem definidos” (James, 1977, p. 207). Como muito bem observa Joel W. Krueger em *James on experience and the extended mind* (2006), uma vez que James descreve a experiência como processo temporal, um paralelo se estabelece com as concepções de mente no contexto das 4Es: a ideia de que, como parte de um processo, a mente se estende, de modo franjado com o ambiente, além do corpo do organismo. Derivado desse processo, nos termos do pragmatismo James, podemos compreender a intencionalidade como uma operação acidental de ajuste externo entre mente e mundo (como, segundo a imagem do próprio James, uma pedra, em um campo, pode “encaixar, em um buraco em outro campo). A ideia de ajuste externo me parece compreender o caráter extensivo da intencionalidade como um processo derivado da dinâmica ‘mente-corpo-ambiente’. Nos termos do que chamo de extensionismo radical no contexto das 4Es em filosofia da mente, em particular, essa ideia de ajuste externo sugere a dissolução do chamado “problema da constituição” (Newen, de Bruin & Shaun Gallagher, 2018, p. 7).

Embora não argumentarei aqui, entendo que não se trata de considerar se, empírica, pragmaticamente ou a priori, um processo cognitivo particular é constituído

ou meramente dependente de processos extracranianos ou processos extracorporais. Segundo o pragmatismo e a forma radical de empirismo de James, nem constituição nem dependência! Como mostrarei a seguir, James desenvolve um núcleo epistemológico em termos do caráter temporal e dinâmico da experiência. Nesse sentido, em um primeiro momento, se diferenciam as partes ‘interior’ e ‘exterior’ (ao organismo) a partir do próprio fluxo da experiência; e, em um segundo momento, por um processo de transição contínua entre essas partes, desenvolvem-se as “relações cognitivas”. Nos termos de James, essa transição significa a diferenciação cognitiva entre as partes da experiência: “são essas relações de contínua transição vividas que tornam cognitivas nossas experiências. Nos casos mais simples e completos, as experiências são cognitivas umas das outras” (James, [1904] 1996, p. 87-8).

Publicado originalmente em *Mind* (1879), o artigo *Are we automata?* será revisado e incorporado ao Capítulo V, *The Automaton-Theory*, em *The Principles of Psychology* (1890). Nesse artigo, em oposição a uma visão mecanicista, James explora uma compreensão de mente/consciência como um processo funcionalmente seletivo. No seminal Capítulo *The Stream of Thought*, ademais, como quinta característica do pensamento, James assinala que, na dinâmica do fluxo de pensamento, “ele está sempre interessado mais em uma parte do seu objeto do que em uma outra, recebe e rejeita, ou escolhe [e que, portanto,] o ato de pensar não é senão uma outra forma da atividade seletiva da mente” (James [1890] 1983, p. 273; 276).

Para James, contudo, a atividade seletiva da mente supõe um certo parâmetro de ordenação da experiência: “sem interesse seletivo, a experiência é um caos absoluto. O interesse por si mesmo dá acento e ênfase, luz e sombra, fundo e primeiro plano – perspectiva inteligível, em uma palavra” (James [1890] 1983, p. 381). Embora seja uma atividade da mente, o interesse não se refere a um mero produto da vontade; é antes uma exigência vital que se desenvolva de forma dinâmica entre organismo e meio. Nesse sentido, por exemplo, James expõe críticas contundentes ao darwinismo mecanicista de Herbert Spencer no sentido de que mostrar que muito mais do que se conformar ao meio, o interesse é um elemento essencial no processo de estruturação da mente em sua relação com as variações espontâneas do pensamento em cooperação com o ambiente: ou seja, como produto de uma exigência vital, a mente se desenvolve como um processo funcionalmente seletivo segundo uma dinâmica que ordena e transforma o caos da experiência em uma perspectiva inteligível.

Assim como o interesse corresponde a uma exigência vital, a própria noção de experiência também incorpora o mesmo sentido quando James afirma que a ‘experiência pura’ se refere ao ‘fluxo imediato da vida’: ou seja, “aquilo” [“that”] do qual se experencia

diretamente e que antecede ao “o que” [“what”] da reflexão (James, 1977, p. 171; 173-4). Nessa dimensão que se pode compreender como “proto-fenomenológica” (Heft, 2001, p. 26), a experiência é ‘pura’ no sentido de ser inqualificável e destituída de distinção mente e mundo ou consciência ou matéria. É a experiência apreendida segundo a ordem temporal dos acontecimentos imediatos da vida sem duplicação interna ou transcendência do conteúdo experienciado. Como campo instantâneo do presente, ademais, só virtual ou potencialmente sujeito e objeto emergem do fluxo da experiência enquanto atualidade ou existência inqualificada ou simplesmente um “aquilo” (James, 1977, p. 177). Nos termos da sua forma radical de empirismo, por ‘puro’, James designa o fato do que é concretamente experienciado:

Para ser radical, um empirismo não deve admitir em suas construções nenhum elemento que não seja diretamente experienciado, nem excluir delas qualquer elemento que seja diretamente experienciado. Para tal filosofia, as relações que conectam as experiências devem ser elas mesmas relações experienciadas, e qualquer tipo de relação experienciada deve ser considerada como “real” como qualquer outra coisa no sistema (James, 1977, p. 195 – tradução minha).

Como assinala Harry Heft (2001, p. 27), na medida em que todas as relações são elas mesmas relações experienciadas, segundo a dinâmica da experiência, é conforme essa dinâmica que se desenvolve o núcleo epistemológico do empirismo de James (Heft, 2001, p. 27): a característica que define o conhecer [knowing] é a ‘seletividade’, ou seja, conhecer é, desde o início, uma função seletiva em que porções da experiência são conectadas e diferenciadas. Nos termos de James, nesse processo, que engendra uma estrutura contínua, nenhum elemento transempírico exerce função e todos os elementos de cognição se derivam de um processo de seleção inerente à própria dinâmica da experiência.

Como ilustração desse processo de seleção empírica dos elementos de cognição, em *The Place of Affectional Facts in a World of Pure Experience* ([1905] 1977), James argumenta que, na experiência humana, as coisas podem possuir diferentes qualidades: a mesma coisa pode significar propriedades físicas (quando tomadas em um contexto da experiência) e propriedades de um valor psíquico (quando tomadas em outro contexto). Para James, contudo, o que determina a seleção de um contexto cognitivo (se físico ou psíquico) resulta da própria dinâmica da experiência: a preciosidade de um diamante em um contexto da experiência é uma qualidade da gema e, em outro, é um sentimento da nossa mente. Em *On a Certain Blindness in Human Beings*, complementarmente, afirma James:

Nossos julgamentos sobre o valor das coisas [...] dependem dos sentimentos que as coisas despertam em nós. Onde julgamos uma coisa preciosa [...] é apenas porque a própria ideia já está associada a um sentimento (James [1899] 2009, p. 1 – tradução minha).

Considerando os sentimentos [*feelings*] que as coisas despertam em nós, a preciosidade de um diamante expressa um “fato afetivo” (Heft, 2017, p. 122). Na medida em que se desenvolve uma relação dinamicamente constante conectando os contextos da experiência, a diferenciação de um sentimento (como ‘preciosidade de um diamante’) é derivada de um processo de seleção inerente ao próprio fluxo de experiência (mais do que de uma atividade intelectual de separação entre objetivo e subjetivo). Como a estrutura da experiência varia ao longo do tempo, a seleção se desenvolve na dinâmica dessa estrutura e define distintos contextos cognitivos.

Para James, dado o caráter temporal da experiência, não se pode afirmar que fatos objetivos e afetivos representam domínios cognitivamente isolados entre si. Em vez disso, quando considerada a afirmação ‘a preciosidade de um diamante é uma qualidade da gema’, se, em um contexto da experiência, o ‘é’ corresponde a uma descrição objetiva, em outro, ele expressa um sentimento de apreciação subjetiva¹⁰. Em função do caráter dinâmico e plural da experiência, não há a significação fixa de um conteúdo cognitivo (como ‘a preciosidade de um diamante é uma qualidade da gema’ ou ‘a preciosidade de um diamante um sentimento da nossa mente’). Segundo a forma radical do empirismo de James, as coisas são sentidas e selecionadas em certos contextos da experiência em que elas incorporam distintas significações objetivas ou subjetivas:

Como ‘subjetivo’ dizemos que a experiência representa; como ‘objetivo’ ela é representada. O que representa e o que é representado são aqui numericamente o mesmo; mas devemos lembrar que nenhum dualismo entre ser representado e representar reside na experiência per se. Em seu estado puro, ou quando isolado, não há autodivisão dela em consciência e no que a consciência é ‘de’. Sua subjetividade e objetividade são apenas atributos funcionais, realizados apenas quando a experiência é ‘tomada’, isto é, falada, duas vezes, considerada junto com seus dois contextos diferentes respectivamente (James, [1904] 1977, p. 177 – tradução minha).

10 Tomando como referência a conjugação entre o pragmatismo e a forma radical de empirismo de William James, pode-se compreender a intencionalidade como um processo dinâmico se estendendo em determinados contextos. Nesse sentido, sendo uma espécie de background, ‘contexto’ equivale aos aspectos ambientais que são dinamicamente incorporados ao processo transacional entre organismo e mundo. Nas palavras de Gregory Bateson (1979, p. 14), aliás, “nada tem significação exceto que seja visto como em algum contexto [...] E ‘contexto’ está associado a outra noção indefinida chamada ‘significação’” (tradução minha).

Nos termos do empirismo de James, diferentemente da interpretação de Thompson e Stapleton, se considerada a situação das bactérias (móveis nadando e escalando um gradiente alimentar de açúcar), se o açúcar se torna significativo, isso resulta de uma dinâmica em que as coisas são sentidas e selecionadas em certos contextos da experiência e assim incorporam distintas significações (objetivas ou subjetivas). Nesse sentido, como sinônimo de intencionalidade, entendo que a atividade de *sense-making* expressa um aspecto derivado da dinâmica da experiência no sentido de resultar de uma seleção acidental de ajuste externo entre organismo e ambiente.

Em sua excelente leitura do empirismo de James, Harry Heft (2001, p. 37) lembra que, como modos de ‘conhecer’ [*knowing*], ‘perceber’ [*perceiving*] e ‘pensar’ [*thinking*] são produtos de processos seletivos. Nesse sentido, como assinala Heft, na medida em que o mundo se apresenta com potencial cognitivo, todos os objetos possíveis de cognição estão baseados em seleção e interesse. E se o interesse organiza o caos da experiência em uma perspectiva cognitivamente inteligível, isso resulta de um processo seletivo inerente ao próprio fluxo da experiência. Como assinalado no parágrafo anterior em relação à significação do açúcar, se a significação resulta de um elemento de interesse, a matriz da qual o interesse emerge é derivada de uma atividade dinâmica em que seletivamente se diferenciam os distintos contextos cognitivos da experiência.

Do ponto de vista do empirismo de James (Heft, 2001, p. 54) em convergência com a perspectiva 4Es em filosofia da mente, os modos de cognição (perceber ou pensar) se ancoram na nossa condição corporal. É com nossos corpos que se desenvolve a seleção cognitiva e, como consequência, o centro do nosso engajamento com o mundo:

O núcleo objetivo da experiência de todo homem, seu próprio corpo, é, verdadeiramente, uma percepção contínua; e igualmente contínuo como uma percepção [...] é o ambiente material desse corpo, mudando por transição gradual quando o corpo se move (James, [1904] 1977, p. 204 – tradução minha).

O mundo experienciado (também chamado de ‘campo da consciência’) sempre vem com nosso corpo como centro, centro de visão, centro de ação, centro de interesse. Onde o corpo está “aqui”; quando o corpo age é ‘agora’; o que o corpo toca é ‘isto’; todas as outras coisas estão ‘lá’ e ‘então’ e ‘aquilo’. Essas palavras de posição enfatizada implicam uma sistematização das coisas com referência a um foco de ação e interesse que reside no corpo [...] Na medida em que ‘pensamentos’ e ‘sentimentos’ podem ser ativos, sua atividade termina na atividade do corpo, e somente despertando primeiro suas atividades podem começar a mudar as do resto do mundo [...] O corpo é o centro da tempestade, a origem das coordenadas, o lugar constante de tensão em todo essa corrente-experiência (James, [1912] 1977, p. 284 – tradução minha).

Como assinala James, segundo uma dinâmica de campo da experiência, o corpo instância núcleo de referência espacial e temporal ('aqui', 'lá', 'isto', 'aquilo', 'agora', 'em seguida' etc.) assim como as ações e os interesses dos nossos pensamentos e sentimentos em relação ao mundo. O uso do termo 'campo' não é gratuito. Nos termos do empirismo de James, compreender a seleção cognitiva do nosso engajamento com o mundo supõe uma dinâmica de campo própria à experiência.

Em suas notas preparatórias para um seminário em psicologia (1895-6), James descreve os dados imediatos da experiência como 'campo': "Campos' que 'desenvolvem', sob as categorias de continuidade uns com os outros [categorias como]: igualdade e alteridade [de] coisas [ou de] correntes de pensamento, realização da significação de um campo no conteúdo de outro campo, 'postulação' de um campo por outro, cognição de um campo por outro etc." (James *apud* Perry, 1976, p. 365). Para James, como a metáfora do campo é inerentemente espacial (Lamberth, 1999, p. 131), ela explicita uma distinção que se torna crucial na sua forma de empirismo em relação à distinção sujeito-objeto. A ideia é que as diferenciações epistemológicas sejam compreendidas em termos de um processo de cognição de um campo (sujeito) por outro campo (objeto).

Considerando a noção de dinâmica de campo, James entende que as categorias epistemológicas possam ser melhor descritas em termos de polos (subjetivos e objetivos) derivados da experiência. Por meio desse processo de cognição de um campo por outro, uma estrutura é selecionada e diferenciada da experiência imediata. Segundo a forma de empirismo de James, contudo, o que precisa ser explicitado é de que forma se desenvolvem os processos de seleção e diferenciação entre os polos da experiência.

Em *Are we automata?* (1879), e posteriormente incorporado com modificações em *The Principles* (1980), James comparou a natureza ativa do perceber com a atividade de um escultor: o que pode ser percebido/esculpido é compelido pelo mundo/mármore. É evidente que as intenções do escultor têm uma função ativa e, no entanto, é importante reconhecer que elas não são simplesmente impostas ao mármore. Em comparação, por mais potente que seja nossas capacidades perceptivas, sempre estaremos compelidos pela resistência do mundo diante de nós. Segundo Harry Heft (2001, p. 55), como um bloco de mármore (não esculpido), o mundo oferece resistência ao longo das linhas de uma estrutura que é lhe inerente (embora ainda não realizada). O que se deve mostrar é de que modo uma estrutura cognitiva é selecionada e diferenciada na dinâmica da experiência segundo uma resistência que é própria a ela.

Em uma longa anotação de trabalho entre 1905 e 1908 (*The Miller-Bode Objections*), James (1988) retoma a noção de campo segundo princípios do eletromagnetismo. Com essa noção, em particular, James sugere de que modo uma estrutura cognitiva é selecionada e diferenciada na dinâmica da experiência. Ele considera que “o efeito mental é um campo como um todo incluindo o eu, o corpo e tudo que está a volta” (James, 1988, p. 107). Embora seja um exemplo ilustrando a solução de James ao problema de como duas mentes podem conhecer a mesma coisa, ele se mostra adequado também à explicitação do modo como uma estrutura cognitiva é selecionada e diferenciada da dinâmica da experiência. Diz James (1988, p. 125): “No mundo físico, embora uma mensagem sem fio se espalhe por todos os lados, ela só é eficaz onde encontra um receptor adequado”.

Como observa Heft (2017, p. 120), se considerada a percepção de um objeto, evidentemente, o receptor adequado seria um percepiente. Na medida em que a mensagem irradia por todos os lados, ela representa a fonte potencial da experiência em direção a eventuais receptores. Retomando a metáfora do mármore, ele representaria a experiência como fonte de irradiação que se espalha dinamicamente por todos os lados e, no entanto, a irradiação ela mesma somente encontraria efetivação significativa diante de um polo receptor.

Por mais tentador que seja ver, na irradiação, um fenômeno de enatividade que ilustra a significação do açúcar como resultante do *sense-making* de bactérias móveis (nando e escalando um gradiente alimentar), não se pode perder de vista que estamos falando de uma dinâmica de campo. Com o advento das teorias de campo em meados do Século XIX, período contemporâneo a James, já não se trata de considerar a descrição da natureza em termos de coisas existindo separadamente (Sachs, 1973, p. 5). O que se apresenta é muito mais um movimento de pensamento que visa uma compreensão da natureza segundo a categoria da continuidade. Com estudo dos campos magnéticos, por exemplo, observou-se que eles são criados por uma dinâmica de cargas elétricas e linhas de força de objetos em interação contínua (Sachs, 1973, p. 6). Com os resultados de Michael Faraday, em particular, a descrição da matéria tende mais aos termos de uma linguagem de campos contínuos de força. Em *The Analysis of Matter* ([1927] 1992), Bertrand Russell assinala que, do ponto de vista do cenário na física contemporânea, nossa compreensão da realidade é mais relacional do que material. De acordo com Russell, com o conceito de campo, a realidade é considerada como processos e eventos mais do que substância e propriedades:

“Eventos” são definidos como o campo da relação de presença [e] existe o que pode ser chamado de “campo” de eventos causalmente conectados [...] Isso se tornou possível devido ao fato de que a “matéria” pode ser considerada como um sistema de eventos, não como parte do material do mundo” (Russell, [1927] 1992, p. 377; 399; 400-1 – tradução minha).

Quando se fala de campo, portanto, a realidade se converte em processos relacionais de presença de eventos ou simplesmente campos. Essa ideia me parece estar no centro do empirismo de James: ou seja, espalhando-se por múltiplas direções segundo a dinâmica de um campo, a experiência é um processo de irradiação em que se diferenciam eventos ou polos da realidade.

Como uma lição tirada das teorias de campo, o que enativistas consideram ser a significação (do açúcar) uma resultante da atividade *sense-making* (de bactérias móveis em um gradiente alimentar), nos termos do empirismo de James, ela se deriva de uma dinâmica análoga à dinâmica de um campo eletromagnético. De igual modo, podemos retomar a imagem do mármore (experiência) e compreender que ele se converte em um campo de emissão enquanto o escultor (perceptivo) assume a atividade de recepção. Entendendo *sense-making* como atividade derivativa de uma dinâmica de campo, de modo análogo, podemos considerar o sentido de intencionalidade e, em particular, intencionalidade não-derivada.

Uma vez que James assume a noção de campo como uma função nuclear na sua forma de empirismo, concluirei este artigo com um resumo breve do sentido de intencionalidade não-derivada segundo uma forma de extensionismo radical. Diferentemente de uma perspectiva enativista no contexto 4Es em filosofia da mente, se considerarmos a noção de ‘intencionalidade não-derivada’ como uma construção conjugando movimentos corporais hábeis em conjunto com demandas ambientais (Newen, de Bruin & Gallagher, 2018, p. 10), ela resultaria da dinâmica da experiência (análoga à dinâmica de um campo eletromagnético) mais do que da simples incorporação de elementos externos à estrutura física do organismo. Retomando a imagem do mármore, enquanto se pode ver nela a representação da experiência como uma fonte de irradiação que se espalha dinamicamente em múltiplas direções, a intencionalidade não-derivada se converteria em um polo de recepção que assume a função do escultor de moldar (*sense-making*) os dados irradiados. Como resultado dessa dinâmica de campo, uma estrutura intencional emerge segundo um processo de seleção e diferenciação inerente à própria dinâmica de campo da experiência.

Se considerado *sense-making* como sinônimo de intencionalidade, não parece restar dúvidas de que somente na superfície enativismo e extensionismo são similares e as divergências são mais profundas do que se supõe. Explorando a conjugação entre o pragmatismo e o empirismo de William James, entendo que se, em termos enativistas, a atividade de *sense-making* expressa o sentido fenomenológica de autonomia do organismo, segundo minha compreensão de um extensionismo radical, ela se mostra como um aspecto derivado da dinâmica de campo da experiência.

Considerações finais

Não resta dúvida de que, se considerado o pragmatismo de William James na esteira da sua forma de empirismo, ele mostra ser uma crítica e uma resposta às teorias representacionais da mente na medida em que se deflaciona a noção de representação como condição de sustentação da intencionalidade do pensamento. Por sua vez, segundo a perspectiva 4E em filosofia da mente, como parte da atividade cognitiva, a intencionalidade se estende além do limite da estrutura corporal e incorpora uma dinâmica organismo-ambiente e deveria ser considerada como ‘não-derivada’ no sentido de que se trata de um engajamento prático entre a estrutura corporal do organismo e o ambiente.

Nos termos gerais da proposta deste artigo, partindo da conjugação entre o pragmatismo e a forma de empirismo de William James em paralelo com a noção de mente estendida, chamo de *extensionismo radical* minha abordagem da intencionalidade. Se considerado como sinônimo de intencionalidade, em particular, o que está em questão é um embate entre enativismo e extensionismo em torno de dois sentidos de *sense-making*. Tendo como background filosófico um embate mais amplo, os dois sentidos expressam a distensão entre fenomenologia e empirismo. E minha ideia de um extensionismo radical no contexto 4Es em filosofia da mente se deriva de uma perspectiva filosófica empirista.

Referências

ARAÚJO, A. William James, Jakob von Uexküll et la critique des notions de représentation mentale et d'intentionnalité. **Philosophical Enquiries: Revue des Philosophies Anglophones**, 7, 107-137, 2016.

BACON, M. **Pragmatism** – An Introduction. Cambridge (UK): Polity, 2012.

BATESON, G. **Mind and Nature** – A Necessary Unity. New York: E. P. Dutton, 1979.

BRENTANO, F. **Psychology from an empirical point of view**. Translated by Antos C. Rancurello, D. B. Terrel & Linda L. McAlister. London/New York: Routledge, 2005.

CHEMERO, A. **Radical Embodied Cognitive Science**. Massachusetts: The MIT Press, 2009.

CHEMERO, A. **Radical Embodied Cognitive Science**. Cambridge (MA): A Bradford Books, 2011.

CHEMERO, A. Radical embodied cognitive science. **Review of General Psychology**, 17(2), 145-150, 2013.

CHEMERO, A. & KÄUFER, S. Pragmatism, Phenomenology, and Extended Cognition. In: **Pragmatism and Embodied Cognitive Science** – From Bodily Intersubjectivity to Symbolic Articulation. Edited by Madzia, R. & Jung, M. Berlin/Boston: De Gruyter, 2016.

CHISHOLM, R. **Perceiving: a philosophical study**. Ithaca and London: Cornell University Press, 1957.

CLARK, A. **Supersizing the Mind**. Oxford: Oxford University Press, 2011.

CLARK, A. & CHALMERS, D. The Extended Mind. **Analysis**, No. 1, 1998.

DI FRANCESCO, M., MARRAFFA, M., & PATERNOSTER, A. A self properly embodied: William James and 4E cognition. **The Jamesian Mind**, p. 493-507, Routledge, 2022.

DI PAOLO, E.; ROHDE, M. & DE JAEGHER, H. Horizons for the Enactive Mind: Values, Social Interaction, and Play. In: **Enaction** – Toward a New Paradigm for Cognitive Science. Edited by John Stewart, Olivier Gapenne & Ezequiel A. Di Paolo. Cambridge (MA): The MIT Press, 2010.

GALLAGHER, S. Pragmatic interventions into enactive and extended conceptions of cognition. **Philosophical Issues**, 24, Extended Knowledge, 2014.

HEFT, H. **Ecological Psychology in Context** – James Gibson, Roger Baker, and the Legacy of William James's Radical Empiricism. New York/London: Taylor & Francis Group, 2001.

HEFT, H. William James' psychology, radical empiricism, and field theory: recent developments. **Philosophical Inquiries**, 2, p. 111-130, 2017.

HUFENDIEK, E. William James and John Dewey on Embodied Action-Oriented Emo-

tions. In: **Pragmatism and Embodied Cognitive Science** – From Bodily Intersubjectivity to Symbolic Articulation. Edited by Madzia, R. & Jung, M. Berlin/Boston: De Gruyter, 2016.

HUTTO, D. D. & MYIN, E. **Radicalizing Enactivism**. London: The MIT Press, 2013.

JACKMAN, H. James' Pragmatic Account of Intentionality and Truth. **Transactions of the C.S Peirce Society**, Winter 1998, Vol. XXXIV, No. 1: 155-181.

JAMES, W. Are We Automata? **Mind**, 4, p. 1-22, 1897.

JAMES. The Moral Philosopher and the Moral Life. **International Journal of Ethics**, April, Vol. 1, No. 3, April, pp. 330-354, 1891.

JAMES, W. **The Varieties of Religious Experience**. New York: Longmans, Green, and Co., 1902.

JAMES, W. **Some problems of philosophy** – a beginning of an introduction to philosophy. New York: Longmans, Green, and Co., 1916.

JAMES, W. **The Writings of William James** – A Comprehensive Edition. Edited by John J. McDermott. Chicago: Chicago University Press, 1977.

JAMES, W. **Pragmatism and The Meaning of Truth**. Cambridge (MA): Harvard University Press, 1978.

JAMES, W. **The Principle of Psychology**. Harvard: Harvard University Press, 1983.

JAMES, W. **Writings 1902/1910**. New York: Literary Classics of the United States, 1987.

JAMES, W. **The works of William James**. Cambridge: Cambridge University Press, 1988.

JAMES, W. **Essays in Radical Empiricism**. Lincoln (Nebraska): University of Nebraska Press, 1996.

JAMES, W. **Pragmatism and other writings**. London: Penguin Books, 2000.

JAMES, W. **On a Certain Blindness in Human Beings**. London: Penguin Books, 2009.

JOHNSON, M. Pragmatism, Cognitive Science, and Embodied Mind. In: **Pragmatism and Embodied Cognitive Science** – From Bodily Intersubjectivity to Symbolic Articulation. Edited by Madzia, R. & Jung, M. Berlin/Boston: De Gruyter, 2016.

JOHNSON, M. **Embodied Mind, Meaning, and Reason**. Chicago: The University of Chicago Press, 2017.

KILPINEN, E. Pragmatism as a Philosophy of Action. In: **Pragmatism Perspectives**. Sami Pihlström & Henrik Rydenfelt (Eds.), 163-179, 2009.

KÜHLE, L. William James and the embodied mind. **Contemporary Pragmatism**, 14(1), p. 51-75, 2017.

KRUEGER, J. W. James on experience and the extended mind. **Contemporary Pragmatism**. Editions Rodopi, Vol. 3, No. 1, 165–176, 2006.

LAMBERTH, D. C. **William James and the Metaphysics of Experience**. Cambridge: Cambridge University Press, 1999.

MADZIA, R. & JUNG, M. **Pragmatism and Embodied Cognitive Science – From Bodily Intersubjectivity to Symbolic Articulation**. Berlin/Boston: De Gruyter, 2016.

MALACHOWSKI, A. James's Holism: the human continuum. In: **The Cambridge Companion to Pragmatism**. Edited by Alan Malachowski. Cambridge University Press, 2013.

MENARY, R. Introduction: The Extended Mind in Focus. In: **The Extended Mind**. Richard Menary (Ed.). Cambridge/London: A Bradford Book/The MIT Press, 2010.

NEWEN, A.; DE BRUIN, L. & GALLAGHER, S. (Eds.). **The Oxford Handbook of 4E Cognition**. Oxford (UK): Oxford University Press, 2018.

PEIRCE, C. S. **Selected Writings**. New York: Dover Publications, 1966.

PERRY, R. B. **The thought and character of William James as revealed in unpublished correspondence and notes, together with his published writings**. Volume II Philosophy and Psychology. London: Oxford University Press, 1976.

ROBBINS, P. & MURAT A. A Short Primer on Situated Cognition. In: **The Cambridge Handbook of Situated Cognition**. Cambridge: Cambridge University Press, 2009.

RORTY, R. **Philosophy and the Mirror of Nature**. Princeton, New Jersey: Princeton University Press. Thirtieth Anniversary Edition, Princeton, NJ: Princeton University Press, 1979.

ROWLANDS, M. Enactivism and the Extended Mind. **Topoi**, 28, p. 53–62, 2009. <https://doi.org/10.1007/s11245-008-9046-z>

ROWLANDS, M. **The New Science of the Mind – From Extended Mind to Embodied Phenomenology**. Cambridge: A Bradford Book/The MIT Press, 2010.

ROWLANDS, M. Disclosing the world – Intentionality and 4E Cognition. In: Newen, A.; de Bruin, L. & Gallagher, S. (Eds.). **The Oxford Handbook of 4E Cognition**. Oxford

(UK): Oxford University Press, 2018.

RUPERT, R. D. **Cognitive Systems and the Extended Mind**. Oxford: Oxford University Press, 2009.

RUSSELL, B. **The Analysis of Matter**. London: Routledge, 1992.

RYLE, G. **The Concept of Mind**. Oxford: Oxford University Press, 2009.

SACHS, M. The Field Concept in Contemporary Science. **The Field Concept in Contemporary Science**. Springfield: Charles C. Thomas Publisher, 1973.

THOMPSON, E. & STAPLETON, M. Making Sense of Sense-Making: Reflections on Enactive and Extended Mind Theories. **Topoi**, 28, 23–30, 2009.

THOMPSON, E. Living ways of sense making. **Philosophy Today**, Vol. 55, 2011.

WARD, D. & STAPLETON, M. Es are good: cognition as enacted, embodied, embedded, affective and extended. In: **Consciousness in Interaction: The role of the natural and social context in shaping consciousness**. Fabio Paglieri (Ed.). Amsterdam/Philadelphia: John Benjamins Publishing Company, 2012.

WIENER, P. Comment on Peirce's What Pragmatism is. In: Peirce, C. S. **Selected Writings**. New York: Dover Publications, 1966.

VARELA, F. **Invitation aux Sciences Cognitives**. Paris: Éditions du Séuil, 1988.

VARELA, F.; THOMPSON, E. & ROSCH, E. **The Embodied Mind – Cognitive Science and Human Experience**. Cambridge: The MIT Press, 1993.

ZALTA, E. N. **Intensional Logic and The Metaphysics of Intentionality**. Cambridge (MA): The MIT Press, 1988.





UMA ANÁLISE CRÍTICA DA PROPOSTA NEOMECANICISTA DE WILLIAM BECHTEL PARA A CIÊNCIA COGNITIVA¹♣



Diego Azevedo Leite²

Resumo:

William Bechtel é um filósofo estadunidense que pode ser considerado como um dos autores mais influentes dentro da literatura especializada sobre o neomecanicismo no âmbito da filosofia da ciência contemporânea. Além disso, ele possui publicações influentes voltadas, especificamente, para a área da ciência cognitiva desde, ao menos, a década de 1980. Diante disso, o objetivo deste trabalho é analisar criticamente algumas ideias centrais presentes na proposta teórica desenvolvida nas publicações de Bechtel voltadas à ciência cognitiva. Nesta análise, o foco da atenção estará nos conceitos de “mecanismos cognitivos (ou mentais)”, “explicação mecanicista” e “representações mentais”, os quais são apresentados e elaborados na obra do autor. Por fim, será apresentada uma conclusão com uma avaliação geral da proposta de Bechtel.

Abstract:

William Bechtel is an American philosopher who can be considered one of the most influential authors in the specialized literature on neomechanism within the scope of contemporary philosophy of science. Furthermore, he has influential publications focused specifically on the area of cognitive science since at least the 1980s. Therefore, the objective of this work is to critically analyze some central ideas present in the theoretical proposal developed in Bechtel's publications concerning cognitive science. In this analysis, the focus of attention will be on the concepts of “cognitive (or mental) mechanisms”, “mechanistic explanation” and “mental representations”, which are presented and elaborated in the author's work. Finally, a conclusion will be presented with a general assessment of Bechtel's proposal.

1 *Publicado originalmente como capítulo do livro **Escritos de Filosofia VI: Linguagem e Cognição**, organizado por Marcus José Alves de Souza & Maxwell Morais de Lima Filho. Editora Fi: Cachoeirinha, 2024, p. 239-66. Disponível em: <https://www.editorafi.org/ebook/b86-escritos-filosofia-vi-linguagem-cognicao>

2 Universidade Federal de Alfenas.

Palavras-chave:

William Bechtel; novo mecanicismo; ciência cognitiva teórica; filosofia da mente; representações mentais.

Keywords:

William Bechtel; new mechanism; theoretical cognitive science; philosophy of mind; mental representations.

Introdução

William Bechtel é um filósofo estadunidense que pode ser considerado como um dos autores mais influentes dentro da literatura especializada sobre o neomecanicismo no âmbito da filosofia da ciência contemporânea. Seu livro de 1993, intitulado *Discovering Complexity* (Descobrimos a Complexidade)³, é considerado um dos marcos fundamentais para o desenvolvimento da proposta do novo mecanicismo nas ciências biológicas, as quais constituem o foco de seus trabalhos na área da filosofia da ciência. Suas análises teórico-filosóficas são feitas, principalmente, com base nas áreas da biologia celular, bioquímica, neurociência, neurociência cognitiva e ciência cognitiva (a qual, dentro deste ponto de vista, é também considerada como uma ciência biológica).

Bechtel é um autor que possui publicações influentes voltadas, especificamente, para a área da ciência cognitiva desde, ao menos, a década de 1980. Ao longo do tempo, seu trabalho se desenvolveu em algumas dimensões importantes para o campo. Duas de suas contribuições iniciais, no formato de livro, foram os trabalhos intitulados *Philosophy of Science: An overview for cognitive Science* (Filosofia da Ciência: Uma visão geral para a ciência cognitiva) e *Philosophy of Mind: An overview for cognitive Science* (Filosofia da Mente: Uma visão geral para a ciência cognitiva), que foram publicados em 1988.

Além disso, Bechtel escreveu, junto com uma colega, a obra *Connectionism and the Mind* (Conexionismo e a Mente). Este livro, publicado, pela primeira vez, em 1991, trata de temas ligados à filosofia da ciência cognitiva e neurociência cognitiva, esclarecendo o que é a perspectiva conexionista. Em 1998, Bechtel foi um dos editores do livro *A Companion to Cognitive Science*, no qual também escreveu, como autor principal, um artigo intitulado *The life of cognitive science* (A vida da ciência cognitiva), que trata, com certo nível de detalhe, de um período significativo da história da ciência cognitiva.

3 Publicado em conjunto com Robert C. Richardson.

Em um momento mais avançado de seu percurso intelectual, suas ideias mais amadurecidas foram reunidas em um dos seus trabalhos mais importantes sobre a investigação da cognição humana: *Mental Mechanisms* (Mecanismos Mentais), publicado em 2008. Este livro permanece, ainda nos dias atuais, como uma de suas contribuições mais importantes para a linha de investigação do fenômeno da cognição humana, baseando-se nos campos da ciência cognitiva e neurociência cognitiva. Ademais, essa é uma das obras fundamentais e mais representativas da literatura produzida pelo novo mecanicismo.

Diante disso, o objetivo deste trabalho é analisar algumas ideias centrais presentes na proposta teórica desenvolvida nas publicações de Bechtel voltadas à ciência cognitiva, no contexto do movimento do novo mecanicismo. Tendo em vista este objetivo, a presente análise foi dividida da seguinte forma. Primeiro, será apresentada a proposta de Bechtel a respeito dos conceitos de “mecanismos cognitivos” e “explicação mecanicista” no contexto do novo mecanicismo na ciência cognitiva e neurociência cognitiva. Segundo, será analisada a proposta de Bechtel em relação à forma de se conceber representações mentais. Terceiro, serão apresentadas algumas críticas a estas propostas de Bechtel. Finalmente, será exposta uma conclusão com uma avaliação geral da proposta do autor.

1. Mecanismos neurocognitivos e explicação mecanicista

A gênese do movimento do novo mecanicismo pode ser buscada em, ao menos, três elementos históricos e contextuais fundamentais. Em primeiro lugar, pode-se mencionar o descontentamento, na segunda metade do século XX, no âmbito da filosofia da ciência, com o modelo de explicação científica mais bem aceito no período: o modelo nomológico-dedutivo desenvolvido, principalmente, em obras de Carl G. Hempel (1905-1997) e Ernest Nagel (1901-1985). Em resumo, esta proposta enfatizava o uso de deduções a partir de leis científicas gerais na construção de explicações científicas (cf. WOODWARD & ROSS, 2021).

Contudo, muitos filósofos da ciência e historiadores da ciência, que trabalhavam com o domínio científico da biologia, encontravam dificuldades de aplicar este modelo de explicação científica ao trabalho de investigação empírica desenvolvido nas ciências biológicas (BECHTEL & ABRAHAMSEN, 2005; MACHAMER *et al.*, 2002). Assim, estes autores buscaram desenvolver um modelo explicativo alternativo, que contribuiu para gerar o movimento do novo mecanicismo.

Dessa forma, o novo mecanicismo enquanto modelo de explicação científica pode ser considerado como uma reação ao modelo anterior, possuindo, conseqüentemente, tanto diferenças quanto pontos de contato com o modelo que busca superar. É justamente tendo em vista essa tendência que o livro “Descobrimo a Complexidade”, de Bechtel e Richardson, publicado em 1993, se insere como um fator importante no desenvolvimento destas novas ideias dentro da filosofia da ciência. No livro, os autores buscam formular de forma mais sólida e robusta uma nova estratégia de explicação científica para as ciências biológicas, pautada nas noções de “mecanismo biológico” e de “explicação mecanicista” (Bechtel & Richardson, 2010).

Em segundo lugar, deve ser também mencionado o grande desenvolvimento teórico e metodológico tanto do domínio científico da biologia (sobretudo a biologia molecular) quanto da neurociência, na segunda metade do século XX. De um lado, a biologia contou com a descoberta da estrutura do DNA e de métodos de decodificação do genoma humano. De outro, a neurociência apresentou novas ferramentas metodológicas, tais como a tomografia por emissão de pósitrons (PET), a ressonância magnética funcional (fMRI), a estimulação magnética transcraniana (TMS) e o eletroencefalograma (EEG), por exemplo.

A partir do desenvolvimento destas novas ferramentas, o conhecimento humano a respeito de sistemas biológicos complexos e do sistema nervoso de humanos e outros animais obteve muitos avanços e isso gerou um grande entusiasmo em muitos autores cujo trabalho era investigar e explicar os fenômenos ligados à cognição humana. Um resultado importante destes desdobramentos foi o grande prestígio que a biologia ganhou entre as ciências “genuínas”. Muitos começaram a considerá-la em um nível de cientificidade até mesmo tão grande quanto a física.

Em terceiro lugar, é igualmente importante mencionar que os desenvolvimentos empíricos mencionados acima geraram um grande entusiasmo em muitos filósofos e cientistas influentes simpáticos a visões de mundo fisicalistas⁴. Trabalhos como o livro *Neurophilosophy* (Neurofilosofia), publicado em 1986, de Patricia Churchland, *A Neurocomputational Perspective* (Uma Perspectiva Neurocomputacional), publicado em 1989, de Paul Churchland e *The Computational Brain* (O Cérebro Computacional), publicado em 1992, de Patricia Churchland e Terrence Sejnowski, são exemplos bastante claros disso⁵.

4 Uma discussão mais profunda a respeito do conceito de “fisicalismo” pode ser encontrada em Stoljar (2021).

5 Estas ideias podem ser consideradas como uma expressão ainda mais forte do que já estava sendo sugerido

Estes trabalhos são citados em muitas publicações de defensores proeminentes do neomecanicismo e é muito provável que sua influência sobre eles, incluindo Bechtel, seja forte. Essa influência ocorre no sentido da inspiração fiscalista, em termos ontológicos (apenas entidades concretas, como neurônios e suas atividades, devem ser aceitas como objetos de investigação), epistemológicos (apenas teorias que mencionem objetos concretos devem ser aceitas como científicas) e metodológicos (o trabalho filosófico está em continuidade com o trabalho dos cientistas e as evidências empíricas das ciências são a chave para resolver problemas filosóficos abstratos).

A título de ilustração, em seu artigo de 1994, Patricia Churchland escreve: “Quando eu assumo que a neurociência pode revelar os mecanismos físicos que subjazem funções psicológicas, eu estou assumindo que é de fato o cérebro que realiza essas funções – que capacidades da mente humana são, na verdade, capacidades do cérebro humano” (CHURCHLAND, 1994, p. 23)⁶. Em seguida, ela classifica essa ideia como uma “hipótese altamente provável, baseada em evidência atualmente disponível através da física, química, neurociência e biologia evolucionária” (CHURCHLAND, 1994, p. 23). Ainda que os novos mecanicistas não adotem explicitamente a proposta de Churchland acerca da eliminação dos conceitos psicológicos tradicionais, muitas das ideias da autora são bastante semelhantes àquelas encontradas tanto nos trabalhos de Bechtel quanto nos de outros novos mecanicistas influentes.

A proposta de explicação da cognição humana de Bechtel através do novo mecanicismo, portanto, é, essencialmente, uma reunião dos elementos mencionados acima. Bechtel considera a cognição humana como um grande e complexo mecanismo biológico capaz de produzir fenômenos cognitivos, sendo composto por vários mecanismos neurais menores (BECHTEL, 1994; 2008; 2009a, b, c, d; 2010; 2012; 2016; 2017; 2019).

O sistema neural é a estrutura física responsável pelas funções cognitivas que conhecemos, como, por exemplo, atenção, memória, linguagem, raciocínio, consciência, etc. Dessa forma, uma função específica relacionada à memória é realizada por um mecanismo neural específico que deve ser identificado, investigado, compreendido e utilizado em uma explicação da função cognitiva nas áreas da ciência cognitiva ou neurociência cognitiva. Assim, um mecanismo neurocognitivo é um “sistema biológico” composto de partes que realizam operações particulares, sendo que estas partes possuem uma organização in-

na tradição de pensamento da qual faz parte o trabalho de Place (1956), por exemplo.

6 Todas as traduções do inglês neste trabalho são do próprio autor.

terna específica (BECHTEL, 2008). É essa organização interna que fará com que a função geral do mecanismo seja realizada. Além disso, o mecanismo neurocognitivo está inserido em um contexto específico, que também exercerá uma grande influência no desempenho da função cognitiva investigada.

Igualmente importante na concepção de mecanismo neurocognitivo de Bechtel é o fato de que todos estes mecanismos possuem diferentes níveis de organização, sendo que processos causais autônomos ocorrem nestes diferentes níveis (BECHTEL, 1994, 2009b). Uma explicação mecanicista, na visão de Bechtel, portanto, deve levar em conta estes diferentes níveis de organização e investigação. Na formulação da explicação final, os níveis devem ser relacionados oferecendo um modelo explicativo pluralista da função cognitiva investigada.

Na construção de uma explicação mecanicista, o objetivo central da investigação é, primeiro, caracterizar o fenômeno que deve ser explicado da forma mais precisa possível. Em seguida, é necessário identificar as partes componentes do mecanismo que deve ser explicado, através da decomposição do mecanismo e da separação de suas partes, tanto quanto possível, para o entendimento correto de cada uma das operações realizadas por cada uma destas partes. O próximo passo é investigar como essas partes e operações se relacionam causalmente, isto é, como este mecanismo está internamente organizado em termos de suas estruturas físicas e funções. Finalmente, é preciso investigar como o contexto afeta a performance do mecanismo. É a partir da reunião de todas estas informações em um modelo teórico que a explicação do fenômeno é fornecida (BECHTEL & WRIGHT, 2009).

Além disso, Bechtel (2010) afirma que explicações mecanicistas têm a pretensão de apresentar normas prescritivas e “orientações normativas” para a ciência cognitiva e neurociência cognitiva. Dessa forma, sua teoria fornece critérios para distinguir explicações mecanicistas de explicações não-mecanicistas e para distinguir boas explicações mecanicistas de explicações mecanicistas ruins. Nessa visão, a cognição humana deve ser explicada pelas ciências interessadas através de boas e adequadas explicações mecanicistas. Consequentemente, outras formas de explicação não devem ser consideradas como explicações científicas válidas e não devem, dessa maneira, ser aceitas no âmbito científico.

Na aplicação de sua versão da teoria mecanicista à cognição humana são normalmente escolhidos o sistema neural da percepção visual e o sistema neural da memória como exemplos (cf. BECHTEL, 2008; 2009d). Frequentemente, são fornecidas informações sobre as atividades e processos neurais, ou padrões de conectividade entre grupos

de neurônios, correlacionados com certas funções cognitivas. Esses processos neurais são descritos em vários níveis: molecular; intracelular; celular; de grupos e redes de neurônios; e de sistemas neurais de maior escala.

Essas descrições relacionadas com estruturas e funções neurais são sempre conectadas com descrições de funções cognitivas, as quais são caracterizadas de forma particular no trabalho de Bechtel. Essa caracterização será analisada na sessão seguinte.

2. A questão do conteúdo das representações mentais

Bechtel (2008, p. xi) caracteriza a cognição humana como um mecanismo (ou sistema) biológico complexo que é capaz de “processar informação”. Neste contexto, “informação” é uma noção causal: o efeito carrega informação sobre a causa e a causa afeta o efeito na medida que transmite informação para ele. Além disso, na visão de Bechtel, um mecanismo neurocognitivo possui a particularidade de poder ser estudado a partir de duas instâncias. A primeira diz respeito às alterações estruturais físicas e químicas que acontecem no sistema nervoso (esta é a perspectiva da implementação); e a segunda diz respeito às funções cognitivas tomadas em abstrato enquanto processamento de informação (esta é a perspectiva representacional). Bechtel (2008, p. 31) entende que a noção de “computação mental” foi utilizada por muitos autores influentes na ciência cognitiva na tentativa de melhor caracterizar o processamento de informação; porém, ele é bastante vago em relação a quão promissora é a utilização desta noção.

De qualquer forma, Bechtel aceita a noção geral de processamento de informação para caracterizar “representação mental”. Isso é o que ocorre, por exemplo, no sistema nervoso quando há uma coordenação de atividades motoras a partir do contato com um estímulo externo (informação) que é percebido através de algum órgão da percepção sensível. Podemos notar, portanto, que, ao mesmo tempo em que Bechtel aceita noções básicas da perspectiva do processamento de informação aplicada à cognição humana, sua concepção de representações mentais é bastante alinhada a um ponto de vista neurocientífico em que o centro da atenção é a atividade neural. Provavelmente por conta da influência do conexionismo em seu trabalho, Bechtel (1993) vê como uma vantagem o fato de que representações mentais sejam entendidas com base nas operações cerebrais e não com base em relações entre representações internas simbólicas e estruturadas, conectadas por meio de regras em uma linguagem organizada. Desse modo, ele se afasta de propostas clássicas, na história da ciência cognitiva, em sua caracterização das representações mentais, como a de Fodor (1975).

Além disso, existem necessariamente dois aspectos das representações mentais na visão do autor: o veículo e o conteúdo. O veículo é aquilo que transmite a informação, é o substrato, podendo ser, por exemplo, um mapa, uma palavra, uma atividade específica de uma rede neural, uma onda sonora, uma imagem de fotografia. O conteúdo é a informação que está sendo transmitida por aquele veículo e que possui um significado. Em outras palavras, o conteúdo é aquilo ao qual o veículo se refere. Mecanismos neurocognitivos processam informação, portanto, no sentido de que suas atividades produzem transformações internas de veículos e seus conteúdos, sendo que estes conteúdos, muitas vezes, fazem referência a entidades e propriedades externas a este mecanismo (BECHTEL, 2008, p. ix).

Dessa forma, neste entendimento de representações mentais, pode-se dizer que regiões específicas do cérebro, com seus neurônios, conexões e atividades, representam os músculos que controlam. Igualmente, estímulos externos particulares do ambiente são representados em áreas cerebrais, se os neurônios nessas áreas entram em atividade quando os estímulos são apresentados. Portanto, a transmissão da informação neste caso é feita através da fibra nervosa e o “processamento de informação” é realizado quando os neurônios se comunicam entre si por meio de atividades elétricas e químicas. Em outras palavras, a informação física sobre o ambiente é codificada e decodificada de forma físico-química e o processo informacional é um processo físico-químico. Dentro desta visão, alguns exemplos deste tipo de processamento de informações são células ganglionares na retina de um sapo que servem como detectores de insetos, células no córtex visual primário que servem como detectores de movimento e células do córtex cerebelar que servem como executores de padrões motores.

Estes exemplos também deixam claro que Bechtel prefere uma caracterização de representações mentais e processamento de informações mais alinhada com o trabalho desenvolvido na neurociência. Ele afirma que o trabalho em neurociência sobre essas noções apresenta uma relação menos problemática entre veículo e conteúdo; e ele aponta que “a estratégia neurocientífica para identificar representações no cérebro é relativamente direta”, uma vez que ela “busca determinar qual estímulo sensorial ou resposta motora está causalmente ligada à atividade de determinados neurônios”; como resultado, “as representações da neurociência são mais claramente fundamentadas nonexo causal que relaciona os organismos com seus ambientes do que aquelas apresentadas na ciência cognitiva” (BECHTEL, 2008, p. 186).

Outra particularidade na proposta de Bechtel é a incorporação de ideias vindas da teoria de sistemas dinâmicos (BECHTEL, 1993; 1998; 2001; 2002; 2008). Isso significa que, na sua visão, a cognição humana deve ser caracterizada como um sistema não-linear complexo. Neste viés, Bechtel caracteriza mecanismos cognitivos como sistemas de controle, e é com base nessa estrutura teórica que se deve compreender representações mentais e questões de conteúdo. Pode-se dizer, de acordo com essa visão, que representações são aspectos de sistemas de controle; elas desempenham um papel crucial em sistemas que controlam outros sistemas, uma vez que “eles podem alcançar o controle apropriado apenas representando informações relevantes” (BECHTEL, 2008, p. 161).

No caso de mecanismos biológicos (organismos), o cérebro é o responsável pelo controle: alguns mecanismos no cérebro estão envolvidos na regulação ou controle de outras partes do cérebro ou órgãos dentro do organismo. Desta forma, o organismo coordena seu comportamento de acordo com as características do ambiente. As informações sobre um determinado processo são fornecidas ao sistema de controle para regular o comportamento; por exemplo, a informação do sistema sensorial é fornecida ao sistema motor para regular sua atividade. Então, quando houver a utilização de informação de um outro sistema pelo sistema de controle haverá, nessa perspectiva, representação. Contudo, o próprio Bechtel (2008, p. xi) reconhece que, embora essa abordagem funcione bem para muitos fenômenos estudados pela neurociência, ela é muito problemática para o que normalmente é considerado fenômeno cognitivo típico, “como resolução de problemas e planejamento”.

Em resumo, podemos perceber, portanto, que Bechtel incorpora a noção de “processamento de informação” para caracterizar aspectos da cognição humana. No entanto, muito pouco é especificado em relação a qual exatamente seria a caracterização do tipo de processamento realizado pela cognição humana. Seria alguma forma específica de computação, como, por exemplo, computação digital, analógica, ou de outro tipo? Onde exatamente ocorreria essa computação, nos neurônios individuais, nos pequenos agrupamentos de neurônios, nas grandes redes neurais? Existem outras células do sistema nervoso envolvidas? Ou essa computação ocorreria apenas no nível das representações e do conteúdo? Como essa forma de computação lidaria com questões complexas e tradicionais, no campo da filosofia da mente, sobre consciência, racionalidade, moralidade e liberdade da vontade humana? Não há discussões profundas sobre essas questões nos seus trabalhos mais centrais e populares.

Bechtel busca reunir em sua proposta mecanicista elementos centrais que fizeram parte do desenvolvimento intelectual do seu trabalho: a noção de processamento de informação mental em redes neurais, de um lado, e ideias relacionadas à teoria de sistemas dinâmicos, de outro. Com isso, podemos constatar que o projeto de Bechtel busca expandir a aplicação da teoria mecanicista nas ciências biológicas tradicionais para as ciências que lidam com a cognição, isto é, a ciência cognitiva e a neurociência cognitiva. Trata-se, assim, de uma tentativa de aplicação de uma estrutura teórica e conceitual formulada com base em uma ciência, a biologia, a outra ciência, a ciência cognitiva.

Resta, então, analisar se essa tentativa de expansão é correta e promissora ou se estamos diante de um caminho investigativo problemático ou equivocado.

3. Os problemas centrais da proposta de Bechtel

Em primeiro lugar, podemos mencionar as dificuldades relacionadas ao “reducionismo” presente na proposta do autor. Bechtel enfatiza, em muitas passagens de seu trabalho, que sua abordagem busca manter alguma autonomia das explicações psicológicas em relação às explicações puramente neurais; ao mesmo tempo, no entanto, ele se compromete com uma posição reducionista a este respeito. O próprio Bechtel caracteriza sua explicação mecanicista como “reducionista” (BECHTEL, 2007). De acordo com o autor, na medida em que a análise mecanicista enfatiza as contribuições feitas pelas partes do mecanismo para a sua operação, ela é, em um sentido importante, “reducionista” (BECHTEL, 2008, p. 21). Portanto, ao mesmo tempo em que a sua teoria mecanicista busca preservar, em algum sentido, alguma autonomia explicativa em relação à ciência cognitiva (dizendo que a explicação final de um fenômeno cognitivo deve levar em conta os múltiplos níveis causais de organização, descrição e análise de um mecanismo neurocognitivo particular), a abordagem está também comprometida com algum tipo de reducionismo – na medida em que é a interação e a organização causal das partes componentes de um mecanismo que explicam o fenômeno produzido por aquele sistema biológico. Vários autores contemporâneos têm explorado até as últimas consequências o que muitos consideram como uma inconsistência substancial da teoria neomecanicista a respeito deste ponto (cf. LEITE, 2019).

Além disso, como foi visto, a abordagem de Bechtel (na medida em que busca uma integração entre disciplinas como neurociência, ciência cognitiva e filosofia da ciência, tendo uma visão de mundo fisicalista como base) se assemelha muito a certas propostas

reducionistas de Churchland (1986; 1994). A autora defende que a estratégia correta para o entendimento das capacidades psicológicas é “essencialmente reducionista”, isto é, explicar as propriedades psicológicas de nível macro em termos das propriedades de redes neurais de nível micro (CHURCHLAND, 1994, p. 23). O raciocínio fundamental por trás dessa estratégia, segundo ela, é o seguinte:

se você quiser entender como algo funciona, você precisa entender não apenas o seu perfil comportamental, mas também os seus componentes básicos e como eles estão organizados de forma a constituir um sistema (CHURCHLAND, 1994, p. 24).

Diante disso, pode-se questionar se, dentro desta proposta mecanicista, restaria, de fato, alguma autonomia para explicações puramente psicológicas na ciência cognitiva, uma vez que a explicação mecanicista pressupõe que alguma estrutura física (neural) é a responsável por qualquer função psicológica. Assim, a explicação exige que esta estrutura física seja apresentada. No entanto, inúmeras explicações psicológicas em diversas áreas da ciência cognitiva, como psicologia do desenvolvimento, psicologia social, psicologia da educação, psicologia computacional, psicologia da percepção e psicologia da tomada de decisão, não seguem essa diretriz mecanicista. Ainda assim, as explicações psicológicas produzidas nessas áreas possuem grande valor científico teórico, empírico e prático. Isso, no entanto, não é considerado de forma adequada na abordagem de Bechtel, que é centrada em pesquisas na área da neurociência.

É exatamente pelo fato de ter uma abordagem centrada na neurociência e em algum tipo de reducionismo que Bechtel ataca também o famoso *argumento da múltipla realização* (BECHTEL & MUNDALÉ, 1999). Em sua visão, não há capacidades ou funções cognitivas específicas que sejam realizadas por diferentes mecanismos neurais específicos. Ou seja, cada função psicológica específica deve ser localizada em algum mecanismo neural específico responsável pela sua realização. O problema é que não há ainda correspondência exata entre função cognitiva e mecanismo neural para inúmeros fenômenos cognitivos complexos conhecidos, sendo possível argumentar, por conseguinte, que algumas capacidades psicológicas específicas são realizadas por múltiplos mecanismos neurais e várias outras podem ter essa mesma característica (cf. AIZAWA, 2009). O argumento da múltipla realização de capacidades cognitivas tem sido defendido até os dias atuais (cf. AIZAWA, 2022).

Ao defender um certo tipo de reducionismo tanto do ponto de vista dos objetos de investigação quanto da explicação desses objetos, Bechtel também propõe o que denominou como *teoria heurística da identidade* (BECHTEL & MCCAULEY, 1999). Na sua visão, afirmações sobre identidades na ciência têm um papel fundamental na medida em que funcionam como hipóteses que podem ser estabelecidas no curso das investigações empíricas. Dessa forma, Bechtel está dizendo que é plausível pensar que um fenômeno psicológico específico ligado a um raciocínio é idêntico a um fenômeno ligado a uma atividade neural específica. Esta identidade, segundo ele, pode ser estabelecida como uma hipótese de investigação, não como uma determinação científica.

O problema com esse raciocínio é que as identidades clássicas estabelecidas na história da ciência são de uma natureza bastante diferente das identidades problemáticas estabelecidas entre fenômenos psicológicos e fenômenos físicos, justamente porque no caso da psicologia há uma discussão histórica profunda sobre a verdadeira natureza de fenômenos mentais. É com base nisso que Eronen (2013) argumenta que identidades, de fato, têm um papel importante na prática científica e na construção de explicações científicas, mas elas não são utilizadas para explicar correlações entre fenômenos físicos e fenômenos mentais. Quando uma identidade é estabelecida na ciência, como, por exemplo, a identidade entre água = H₂O, isso é feito porque há uma correspondência entre as propriedades das duas partes dessa relação. O mesmo não acontece, porém, no caso de fenômenos mentais, os quais possuem propriedades bastante distintas das propriedades físicas, como por exemplo: as capacidades de representações mentais conscientes de se referirem a algo externo a elas mesmas ou de se relacionarem logicamente; e a capacidade de sensações conscientes de se referirem a experiências subjetivas muito particulares, como a experiência de experimentar o sabor de um chocolate pela primeira vez na vida. Nestes casos, falar de uma identidade entre tais fenômenos e atividades neurais é extremamente problemático, o que torna a proposta de Bechtel insustentável.

Em segundo lugar, há também graves problemas em relação ao entendimento de Bechtel acerca da natureza das representações mentais humanas e do conteúdo destas representações. Em relação a este ponto, como foi visto anteriormente, o trabalho do autor se baseia fortemente na abordagem do processamento da informação e na investigação da atividade físico-química de redes neurais dinâmicas. É por conta disso que encontramos informações tão detalhadas em seu trabalho a respeito das atividades e processos neurais tanto no nível molecular e celular quanto no de redes e sistemas de neurônios.

Contudo, essa ênfase nos processos neurais faz com que a abordagem deste autor seja muito centrada na parte neural do problema, enquanto a parte dos problemas complexos ligados ao fenômeno cognitivo é negligenciada. No trabalho de Bechtel, não há discussões tão profundas sobre a histórica dificuldade de se explicar o fenômeno do “significado” de uma palavra ou um símbolo, nem sobre como diferenças na interpretação do significado de um mesmo símbolo podem gerar enormes diferenças no comportamento das pessoas. Este “conteúdo” das representações mentais humanas é reduzido a interações físico-químicas entre redes e sistemas neurais. Dessa forma, muitos aspectos da complexidade da cognição humana e do comportamento humano são deixados de lado na análise de Bechtel sobre representações mentais (cf. QUILTY-DUNN *et al.*, 2022).

Além disso, ainda que seja aceita a ideia de que o sistema nervoso é *necessário* para a manifestação de funções cognitivas na natureza, pode-se questionar se o sistema nervoso é *suficiente* para a manifestação de todos os fenômenos cognitivos que o ser humano é capaz de produzir. Isso, evidentemente, não é questionado por Bechtel porque seu projeto, de antemão, assume uma forma de fisicalismo. A própria caracterização da cognição humana como um “mecanismo” ou “sistema” composto de partes já pressupõe a ideia fisicalista de um objeto no espaço que pode ser dividido em partes. Mas o pressuposto fisicalista de que todos os fenômenos cognitivos possuem uma base física, neural, é extremamente controverso ainda nos dias atuais.

De forma contrária a essa visão, pode-se argumentar que, em uma representação mental original, não há uma separação radical em partes, por exemplo, dos elementos da memória, da atenção, do raciocínio, da consciência e da emoção. Em um processo complexo de tomada decisão (como a escolha do próximo presidente de uma república), por exemplo, é, frequentemente, muito difícil separar o raciocínio das emoções, as representações conscientes das não-conscientes, o interesse pessoal da preocupação com o impacto social.

Nesse caso, a explicação relevante das ações é feita através de considerações: 1) sobre a criação e alteração do conteúdo de percepções, crenças, sensações, emoções, máximas, vontades, desejos, intenções etc.; 2) sobre as suas relações internas; e 3) sobre as suas relações externas com o contexto físico, social, histórico e cultural. Portanto, a utilização de uma estrutura teórico-conceitual para investigar a cognição humana baseada em mecanismos biológicos, sistemas dinâmicos complexos e processamento de informação parece problemática quando estamos analisando uma entidade plenamente consciente, autoconsciente, que raciocina, age através da liberdade, é capaz de pensamento moral e atua em um contexto social dinâmico e complexo.

Quando estamos analisando o tema da cognição humana é preciso ter em mente que as representações mentais têm um caráter bastante complexo. Na verdade, o fenômeno cognitivo humano aparenta possuir particularidades que não são encontradas em sistemas mecânicos naturais e artificiais, ainda que complexos. O ser humano é capaz, por exemplo, de produzir representações mentais originais construindo sistemas de pensamento inovadores. O ser humano possui desejos, objetivos e metas e fica frustrado quando não as alcança, mas também se sente alegre quando os desejos são realizados. Além disso, o ser humano busca o desenvolvimento intelectual ao longo de toda a sua vida e é afetado pelas suas emoções e sentimentos. A cognição humana é capaz de relacionar ideias de forma lógica e pensar sobre direito, ética e moralidade, apresentando normas e prescrições para a sua própria conduta, com base no que julga ser bom, justo e correto.

De fato, muitos sistemas computacionais atuais são capazes de simular habilidades cognitivas humanas particulares e até mesmo superá-las. No entanto, a configuração cognitiva humana difere em muitos aspectos de um sistema computacional artificial. Em processos de tomada de decisão complexos, há inúmeros elementos que interagem para que um resultado comportamental particular seja obtido: sistemas de crenças, prazer e desprazer, emoções complexas, desejos e motivações, autoconsciência e autoconceito, a relação entre a pessoa e o ambiente e entre a pessoa e a sociedade.

Particularmente, crenças e sistemas de crenças podem ser compreendidos como representações abstratas simbólicas de conceitos ou ideias que podem ser comunicadas através de palavras e da linguagem natural, que opera por meio de regras aplicadas às dimensões lexical, sintática e semântica (cf. KRÜGER, 2018). Sistemas de crenças aceitos por uma determinada pessoa desempenham frequentemente um papel central no seu comportamento individual. As crenças estão, muitas vezes, logicamente relacionadas e fazem sentido para a pessoa que as possui. Sistemas de crenças afetam um indivíduo em sua forma de pensar, em seus processos mentais, seus sentimentos e suas emoções, suas motivações, sua aprendizagem, sua tomada de decisão e suas formas de agir. Através da autoconsciência, o ser humano adquire consciência sobre si mesmo e sobre as razões que fundamentam o seu agir, podendo assim questioná-las e aprimorar a si mesmo.

Essas crenças podem dizer respeito também à forma como a pessoa se observa no mundo ou à forma como ela interage com o ambiente, com grupos de pessoas, com a sociedade e com a cultura (cf. BRUNER, 1990, 1996; STURM, 2008). Sistemas de crenças podem adquirir uma dimensão histórica, atravessando séculos e sendo transmitidos de uma geração para outra. Além disso, o ser humano interage em uma sociedade complexa

através de suas ideias e possui uma concepção de si mesmo e do seu valor nesta sociedade. Por exemplo, um atleta de alta performance pode se ver mais motivado a buscar um resultado se acreditar que este resultado irá orgulhar a sua família ou o seu país. Nestes momentos, os obstáculos em relação ao cansaço mental ou físico ou à dor física podem ser superados tendo em vista um resultado que é importante e faz sentido para o atleta. Estes obstáculos podem ser encarados como sacrifícios necessários na busca de um bem maior.

Este tipo de configuração cognitiva, contudo, não ocorre em sistemas artificiais. Uma máquina não possui vínculo afetivo familiar, não possui vínculo afetivo com sua nação ou com a comunidade da qual faz parte, assim como com a cultura desta comunidade, na qual uma pessoa é educada desde o momento do seu nascimento. É dentro desta cultura que as linguagens simbólicas e os símbolos ganham os seus significados, de acordo com as relações interpessoais entre indivíduos e entre eles e o meio ambiente onde vivem e se comportam. Por exemplo, o significado da imagem de uma bandeira nacional para um atleta de alta performance em uma competição, enquanto símbolo que representa o seu país, não é o mesmo que o processamento da imagem desta bandeira, enquanto símbolo arbitrário, realizado por qualquer sistema cognitivo artificial.

Máquinas não superam obstáculos na vida, pois não sofrem, não sentem dor; elas não precisam constantemente buscar a sua própria sobrevivência, elas não têm consciência de que irão morrer um dia. Igualmente, elas não se preocupam com o seu desenvolvimento pessoal durante o seu período de vida e não se preocupam em refletir sobre a sua própria experiência de viver, nem em dar algum sentido a sua própria existência. Máquinas tampouco se preocupam em deixar um legado para gerações futuras ou ter o seu trabalho socialmente reconhecido pelo seu alto nível de qualidade – que foi alcançado através de muito esforço e sacrifícios durante o período de vida. Máquinas, portanto, não se importam se terão reconhecimento e prestígio social ou se serão desprezadas socialmente. Isso não fará nenhuma diferença para elas, na medida em que não precisam lidar com questões de autoestima. Estes aspectos da experiência de viver, porém, claramente, fazem parte da experiência humana cotidiana, afetando com frequência inúmeros processos cognitivos. É tendo em vista justamente esses tipos de reflexões que se pode constatar que a configuração cognitiva humana é ainda muito distinta da configuração presente em um sistema cognitivo artificial que meramente processa informações.

Tendo em vista as novas tecnologias atualmente disponíveis, algumas pessoas podem pensar que as máquinas finalmente superaram a inteligência humana. Ou, mais dramaticamente, que as máquinas triunfaram sobre o homem. Porém, antes de tudo, os

responsáveis pela criação dessas máquinas inteligentes são os próprios seres humanos. Além disso, quando os sistemas inteligentes e robóticos mais sofisticados atualmente são comparados com a inteligência e cognição humana, é possível perceber que a diferença ainda é enorme. Sistemas cognitivos artificiais e robóticos não sofrem nem sentem prazer. Eles não nascem em contexto de violência doméstica. Eles não sentem a alegria de alcançar um objetivo de longo prazo. Eles também não procuram orgulhar os pais e não buscam reconhecimento social, elogios, honrarias ou prestígio. Eles não refletem sobre questões morais nem têm seus próprios princípios morais.

Existem diferenças substanciais na forma como os humanos e os sistemas de computador usam o raciocínio e a linguagem. A cognição humana não é uma máquina estatística que procura encontrar padrões em uma grande quantidade de dados sem sentido e gerar a resposta mais provável. A cognição humana parece ter um sofisticado sistema interno de representações baseadas em princípios lógicos que permitem aos humanos gerar complexas e longas cadeias de pensamentos originais.

Ademais, a mente humana tem a capacidade de descrever e prever fatos: ou seja, dizer o que foi, o que é e o que será. Mas também tem a capacidade de explicar as coisas: isto é, de dizer o que pode ou não acontecer e por quê. A mente humana tem, então, a capacidade de gerar conjecturas e raciocinar em função de eventos ou fatos relacionados causalmente. Além disso, a cognição humana também tem a capacidade de raciocínio moral. O ser humano tem a capacidade de entender o que deve ou não ser feito em função de seus princípios éticos. Isso significa que, a partir do enorme número de comportamentos possíveis para um ser humano, ele é capaz de distinguir entre o que é certo e o que é errado fazer em uma determinada situação. Por fim, o ser humano tem a capacidade de refletir criticamente sobre seu próprio sistema de crenças e princípios morais, aprimorando-o ao longo de toda a vida. Ao fazer isso, cada um pode construir sua maneira significativa e original de pensar sobre o mundo, sobre a existência e sobre a moralidade.

Os sistemas computacionais, ao contrário, até agora, não formam crenças por conta própria, não têm a capacidade de avaliá-las e aperfeiçoá-las por si mesmos. Se olharmos para o problema de um ponto de vista muito concreto e objetivo, observamos que mesmo os sistemas de computador mais avançados, os robôs mais modernos e as arquiteturas neurais e cognitivas artificiais mais desenvolvidas ainda estão muito longe de se comportarem como seres humanos em relação à linguagem e ações que envolvem consciência e racionalidade informal. Os seres humanos são capazes de jogar xadrez, fazer pizza, fazer café, conversar sobre política, criar uma nova música no violão e jogar tênis no mesmo

dia. Nenhum sistema artificial computacional é atualmente capaz dessa generalidade na cognição. Portanto, ainda existem muitas diferenças substanciais entre os sistemas artificiais mecânicos que meramente processam informação e a cognição humana.

Finalmente, é importante mencionar a questão da divergência entre as versões da teoria mecanicista da cognição humana. Como foi visto, Bechtel é um autor que utiliza o termo “explicação mecanicista reducionista” e defende alguma forma de redução do vocabulário da psicologia ao vocabulário dos mecanismos neurais. Outros autores influentes no novo mecanicismo tentam deixar claro que suas propostas não são reducionistas, evitando a todo custo até mesmo o uso deste termo. Bechtel também caracteriza representações mentais e lida com a questão do conteúdo através da estrutura teórica da teoria dos sistemas dinâmicos.

Não obstante, entre os autores mais influentes do novo mecanicismo, ele é o único que adota essa perspectiva enquanto pilar central da sua proposta. Outro ponto de divergência se relaciona com o fato de Bechtel utilizar a noção de identidade heurística em sua reflexão teórica, o que não está explicitamente presente em nenhuma outra versão influente da teoria mecanicista. E existem também inúmeras divergências terminológicas e conceituais entre Bechtel e outros novos mecanicistas. Portanto, não é possível considerar a proposta de Bechtel como uma proposta que unifica ou que integra em um grau significativo o movimento intelectual do novo mecanicismo na ciência cognitiva e áreas afins. Este movimento intelectual, assim como as várias versões da teoria mecanicista da cognição humana, têm um caráter bastante difuso (cf. LEITE, 2021).

Conclusão

De fato, o movimento do novo mecanicismo na ciência cognitiva conseguiu reunir muitos defensores influentes que, como Bechtel, possuem um foco de atenção bastante voltado para a investigação da atividade e dos processos em células e redes neurais. Certamente, essa é uma dimensão bastante forte na ciência cognitiva atual e nas reflexões contemporâneas sobre a natureza da cognição humana. Dentro deste contexto, a proposta de Bechtel faz algumas contribuições úteis para esclarecer o que exatamente são mecanismos neurocognitivos humanos, explicações mecanicistas e como exatamente esses conceitos se relacionam com noções clássicas da ciência cognitiva, tais como representações mentais e processamento de informações. É, seguramente, muito importante que o trabalho teórico de esclarecimento de conceitos fundamentais e de diferentes noções de explicação científica seja feito na ciência cognitiva.

Por outro lado, a proposta teórica de Bechtel, fundamentada em um modelo antropológico fisicalista muito presente na cultura intelectual predominante na ciência cognitiva nos tempos atuais, apresenta dificuldades bastante substanciais. Uma destas dificuldades diz respeito ao fato de que suas contribuições são, tipicamente, sugestões muito gerais sobre como se pode tentar compreender a extremamente complexa relação entre o cérebro humano e os fenômenos cognitivos humanos. Não há ainda uma explicação propriamente dita, em termos mecanicistas, para fenômenos complexos como o conteúdo das representações mentais, os fenômenos da consciência e autoconsciência humana, os fenômenos que envolvem racionalidade, moralidade e liberdade da vontade.

A estrutura teórica mecanicista apresenta inúmeras limitações quando aplicada a atividades cognitivas de alto nível de complexidade. Exemplos de fenômenos cognitivos humanos mais complexos não são sequer mencionados nos trabalhos. Ao invés disso, as discussões são centradas em exemplos retirados da neurociência e da biologia. De fato, toda a abordagem é muito centrada na área da neurociência. Com isso, muitas questões psicológicas abstratas são simplesmente negligenciadas.

Ademais, o movimento neomecanicista é bastante difuso. Existe, claramente, uma falta de unidade de pensamento entre os principais defensores deste ponto de vista teórico, até mesmo quando consideramos as questões mais fundamentais. Embora Bechtel seja um dos principais autores dentro do movimento neomecanicista, sua versão da teoria mecanicista da cognição humana apresenta pontos de divergência substanciais em relação a formulações de seus colegas mecanicistas, de modo que a proposta neomecanicista não apresenta uma estrutura teórica minimamente unificada. A análise mais profunda e detalhada das propostas dos defensores do movimento precisa ser feita, dessa forma, individualmente. E a análise individual feita neste trabalho dos pontos centrais da proposta de Bechtel mostra que sua teoria possui uma série de limitações. Consequentemente, afirmações de que essa é a melhor teoria que temos atualmente não se sustentam.

Referências

AIZAWA, K. Neuroscience and multiple realization: a reply to Bechtel and Mundale. *Synthese*, v. 167, n. 3, p. 493-510, 2009.

AIZAWA, K. The multiple realization of human color vision revisited. *Frontiers in Psychology*, 2022. DOI: 10.3389/fpsyg.2022.985267.

BECHTEL, W. & ABRAHAMSEN, A. Explanation: A mechanist alternative. **Studies in History and Philosophy of Biological and Biomedical Sciences**, v. 36, p. 421-41, 2005.

BECHTEL, W. & MCCAULEY, R. N. Heuristic identity theory (or back to the future). In: M. Hahn & S. C. Stoness (Eds.). **Proceedings of the 21st Annual Meeting of the Cognitive Science Society**. Mahwah, NJ: Lawrence Erlbaum Associates, 1999. p. 67-72.

BECHTEL, W. & MUNDALE, J. Multiple realizability revisited. **Philosophy of Science**, v. 66, p. 175-207, 1999.

BECHTEL, W. & RICHARDSON, R. **Discovering complexity: Decomposition and localization as strategies in scientific research**. Cambridge, MA: MIT Press, 2010 (Publicado originalmente em 1993).

BECHTEL, W. & WRIGHT, C. D. What is psychological explanation? In: Symons, John & Calvo, Paco (Eds.). **The Routledge companion to philosophy of psychology**. New York: Routledge/Taylor & Francis Group, 2009, p. 113-30.

BECHTEL, W. The case for connectionism. **Philosophical Studies**, v. 71, n. 2, p. 119-54, 1993.

BECHTEL, W. Levels of description and explanation in cognitive science. **Minds and Machines**, v. 4, p. 1-25, 1994.

BECHTEL, W. Representations and cognitive explanations: assessing the dynamicist's challenge in cognitive science. **Cognitive Science**, v. 22, n. 3, p. 295-318, 1998.

BECHTEL, W. The compatibility of complex systems and reduction: a case analysis of memory research. **Minds and Machines**, v. 11, p. 483-502, 2001.

BECHTEL, W. Decomposing the mind-brain: a long-term pursuit. **Brain and Mind**, v. 3, p. 229-42, 2002.

BECHTEL, W. Reducing psychology while maintaining its autonomy via mechanistic explanations. In: Schouten, Maurice & Looren de Jong, Huib (Eds.). **The matter of the mind: Philosophical essays on psychology, neuroscience, and reduction**. Malden, MA: Blackwell Publishing, 2007, p. 172-198.

BECHTEL, W. **Mental mechanisms: Philosophical perspectives on cognitive neuroscience**. New York: Routledge, 2008.

BECHTEL, W. Constructing a philosophy of science of cognitive science. **Topics in Cognitive Science**, v. 1, n. 3, p. 548-69, 2009a.

BECHTEL, W. Looking down, around and up: mechanistic explanations in psychology. **Philosophical Psychology**, v. 22, n. 5, p. 543-64, 2009b.

BECHTEL, W. Mechanism, modularity, and situated cognition. In: Robbins, Philip & Aydede, Murat (Eds.). **The Cambridge handbook of situated cognition**. Cambridge: Cambridge University Press, 2009c, p. 155-70.

BECHTEL, W. Molecules, systems, and behavior: Another view of memory consolidation. In: Bickle, J. (Ed.). **The Oxford handbook of philosophy and neuroscience**. New York: Oxford University Press, 2009d, p. 13-40.

BECHTEL, W. How can philosophy be a true cognitive science discipline? **Topics in Cognitive Science**, v. 2, p. 357-66, 2010.

BECHTEL, W. Identity, reduction, and conserved mechanisms: perspectives from circadian rhythm research. In: S. Gozzano & C. S. Hill (Eds.). **New perspectives on type identity: The mental and the physical**. Cambridge: Cambridge University Press, 2012, p. 43-65.

BECHTEL, W. Investigating neural representations: the tale of place cells. **Synthese**, v. 193, p. 1287-321, 2016.

BECHTEL, W. Explicating top-down causation using networks and dynamics. **Philosophy of Science**, v. 84, p. 253-74, 2017.

BECHTEL, W. Resituating cognitive mechanisms within heterarchical networks controlling physiology and behavior. **Theory and Psychology**, v. 29, n. 5, p. 620-39, 2019.

BECHTEL, W.; Abrahamsen, A. & Graham, G. The life of cognitive science. In: Bechtel W. & Graham, G. (Eds.). **A companion to cognitive science**. Oxford: Basil Blackwell, 1998, p. 1-104.

BRUNER, J. **Acts of meaning**. Cambridge, MA: Harvard University Press, 1990.

BRUNER, J. **The culture of education**. Cambridge, MA: Harvard University Press, 1996.

CHURCHLAND, P. S. **Neurophilosophy**. Cambridge, MA: MIT Press, 1986.

CHURCHLAND, P. S. Can Neurobiology Teach Us Anything about Consciousness? **Proceedings and Addresses of the American Philosophical Association**, v. 67, p. 23-40, 1994.

ERONEN, M. Hypothetical identities: explanatory problems for the explanatory argument. **Philosophical Psychology**, 2013. DOI:10.1080/09515089.2012.736076.

FODOR, J. **The language of thought**. New York: Crowell, 1975.

KRÜGER, H. **Psicologia social das crianças**. Curitiba: Editora CRV, 2018.

LEITE, D. A. Neo-mechanistic explanatory integration for cognitive science: The problem of reduction remains. *Sofia*, v. 8, n. 1, p. 124-45, 2019 [Dossiê filosofia da mente e da linguagem]. Disponível em: <https://periodicos.ufes.br/sofia/issue/view/1078>

LEITE, D. A. **The twenty-first century mechanistic theory of human cognition: A critical analysis**. Cham: Springer, 2021.

MACHAMER, P., DARDEN, L. & CRAVER, C. Thinking about mechanisms. *Philosophy of Science*, v. 67, p. 1-25, 2000.

PLACE, U. T. Is consciousness a brain process? *British Journal of Psychology*, v. 47, p. 44-50, 1956.

QUILTY-DUNN, J., POROT, N. & MANDELBAUM, E. The best game in town: the re-emergence of the language of thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences*, 2022. DOI: 10.1017/S0140525X22002849.

STOLJAR, D. Physicalism. *The Stanford encyclopedia of philosophy*, 2021. <https://plato.stanford.edu/archives/sum2022/entries/physicalism/>

STURM, T. Why did Kant reject physiological explanations in his anthropology? *Studies in History and Philosophy of Science*, v. 39, p. 495-505, 2008.

WOODWARD, J. & ROSS, L. Scientific Explanation. *The Stanford encyclopedia of philosophy*, 2021. <https://plato.stanford.edu/archives/sum2021/entries/scientific-explanation/>





DAS

ANÁLISES INVESTIGATIVAS DOS SERES VIVOS E A QUESTÃO DO PRINCÍPIO ANÍMICO EM ARISTÓTELES



Rodrigo Romão de Carvalho

Resumo:

Na investigação composicional dos organismos, Aristóteles, sem deixar de considerar o papel dos elementos materiais e os movimentos espontâneos a eles intrinsecamente associados, empenha-se em conceder uma primazia explanatória, de tipo teleológica, a determinados princípios de natureza formal e final. Nesse sentido, pretendo desenvolver uma análise na qual aponte às razões pelas quais Aristóteles procura sustentar esta primazia explanatória, a partir de certas considerações críticas que ele direciona aos filósofos naturalistas do século V a.C., notadamente Empédocles e Demócrito. Com isso, procurarei estabelecer uma relação entre (i) os fatores explicativos de caráter *formal-final* no estudo dos seres vivos em as *Partes dos Animais*, e (ii) a noção de alma, tal como expressa no *De Anima*. A partir desta investigação, pretenderei ver em que medida e por quais razões caberia ou não à ciência da natureza examinar, além das

Abstract:

In the compositional investigation of organisms, Aristotle, taking also into consideration the role of the material elements and the spontaneous movements inherently associated with them, endeavors to give an explanatory primacy, of teleological type, to certain principles of formal and final nature. In this sense, I intend to develop an analysis of the reasons why Aristotle supports this explanatory primacy from certain critical considerations which he addresses to the naturalistic philosophers of the 5th century BC, notably Empedocles and Democritus. Thus, I will try to establish afterwards a relation between (i) the explanatory factors of *formal-final* character in the study of the living beings in the *Parts of the Animals*, and (ii) the notion of the soul, as expressed in *De Anima*. Finally, I will deal with Aristotle's problem of knowing whether it would be proper for the science of nature to examine, besides the psychic faculties common to all

faculdades anímicas comuns a todo vivente, a faculdade intelectual exclusiva do ser humano.

living beings, the intellectual faculty exclusive of the human being.

Palavras-chave:

Causalidade; investigação da natureza; constituição orgânica; faculdades da alma; Aristóteles.

Keywords:

Causality; nature research; organic constitution; faculties of the soul; Aristotle.

1. Introdução

De acordo com a concepção aristotélica de natureza, deve-se priorizar as explicações em termos de causas *formais-finais* em relação às explicações em termos de causas *materiais-eficientes* (*Partes dos Animais*, I.1, 639b11-30), o que se ajustaria, portanto, a um modelo teleológico de natureza, em detrimento de certos modelos *necessitaristas*. Estes modelos estariam ligados àqueles sustentados pelos filósofos naturalistas do século V a.C., ou pelos, como Aristóteles costuma denominá-los, *physiologoi* (φυσιολόγοι), dentre os quais, nos escritos aristotélicos, se destacam os nomes de Empédocles (cerca de 490-435 a.C.) e de Demócrito (cerca de 460-370 a.C.).

As concepções filosóficas destes filósofos representam algumas das posições adversárias, contra as quais Aristóteles argumenta em diversos textos. Segundo Aristóteles, Empédocles pretendia, equivocadamente, recorrer tão somente a um conjunto de causas materiais e eficientes, independentemente associadas entre si por uma relação de concomitância, para explicar, por exemplo, a constituição da coluna vertebral no processo de geração de um novo indivíduo:

Empédocles não se pronunciou corretamente ao afirmar que muitos itens pertencem aos animais porque assim sucedeu concomitantemente no vir a ser; como, por exemplo, ter a espinha de tal e tal qualidade porque, ao se retorcer [sc. o corpo], sucedeu-lhe concomitantemente quebrar-se – ele não reconheceu, primeiramente, que é preciso que o esperma constituidor esteja já disposto no começo com uma capacidade de tal e tal tipo e, em seguida, que o produtor se apresente como anterior não apenas por definição, mas também no tempo: pois é um homem que gera um homem, de modo que é porque aquele homem

é de tal e tal qualidade que o vir a ser sucede assim de tal modo para este outro¹ (*Partes dos Animais*, I.1, 640a19-24).

Para Empédocles, seria, então, suficiente explicar a formação da espinha dorsal unicamente pelos movimentos absolutamente necessários da matéria que ocorrem espontaneamente. É pela razão de a coluna (ἀκάνθα) ser constituída por certo tipo de material, acrescentado ao fato de o embrião, casualmente, contorcer-se muitas vezes, que ela apresentaria certas características, ou seja, que ela se apresentaria como um conjunto de vértebras articuladas. Portanto, a função exercida pela espinha no animal como um todo seria devido a uma mera consequência dos movimentos espontâneos das disposições dos elementos materiais.

Contudo, Aristóteles considera que, apesar de necessário, certo conjunto de causas materiais e eficientes não é suficiente para engendrar as partes do organismo vivo; antes, é preciso lançar mão das causas formais e finais, as quais conduzem o conjunto de causas materiais e eficientes, de modo a estabelecer entre elas uma articulação de interdependência, a fim de promover as propriedades e as disposições composicionais capazes de possibilitar, ao animal, o exercício efetivo de suas funções vitais.

2. Processos espontâneos e teleológicos

Tanto Aristóteles como Empédocles entendem que a constituição orgânica depende da conjunção de diversas séries causais, através de interações entre causas materiais e eficientes, cuja devida ordem de encadeamento permitem originar os organismos a serem examinados e explicados. Porém, as questões que se devem colocar com relação, por um lado, à posição de Aristóteles, e, por outro, à de Empédocles, dizem respeito ao *como* as séries causais viriam a ser apropriadamente encadeadas, e *se* ocorreria ou não algum desvio do curso que a matéria elementar seguiria espontaneamente, por meio da conjunção destas séries causais (ANGIONI, 1999, p. 59).

A posição que Aristóteles atribui a Empédocles, mas que também poderia ser atribuída a Demócrito, estabelece que o apropriado encadeamento das séries de causas materiais e eficientes, que promove a formação das partes orgânicas, ocorreria por espontaneidade, devido, exclusivamente, a uma interação fortuita entre os elementos a

¹ Com relação às citações das obras de Aristóteles traduzidas para o português, utilizei a edição “Bekker I”, as traduções para o inglês das edições bilíngues da “Loeb Classical Library”, as traduções para o espanhol das edições “Gredos”, e a tradução para o português de L. Angioni do Livro I das *Partes dos Animais*.

partir de suas disposições essenciais (*Física*, II.8, 198b23-31). Aristóteles não aceitará esta posição, visto que, para ele, a devida concatenação das séries de causas materiais e eficientes, através da qual o organismo vem a ser constituído, não poderia efetuar-se tão somente por interações casuais da matéria elementar. Antes, estas séries causais se ordenariam de um modo apropriado por intermédio de causas formais e finais, que atuariam sob o pressuposto de um princípio anterior, ou de certo fator teleológico. Desta maneira, os componentes elementares desviam o curso espontâneo de suas disposições essenciais, e, em função deste desvio, adquirem propriedades acidentais, as quais permitem aos seres vivos a realização de suas funções características².

O princípio anterior, que intercede sobre a série causal das interações elementares, é justamente a forma do vivente. Esta forma, própria à espécie de cada ser vivo, provém do esperma constituidor, contendo, em si, desde o início do processo de geração, o potencial (δύναμις) para dar origem ao novo indivíduo (LENNOX, 2001, p. 132). Deste modo, o desenvolvimento de um animal provido de coluna vertebral depende de um organismo do mesmo tipo, com a capacidade de produzir o sêmen específico, no qual há o potencial de originar, nas condições favoráveis, outro animal do mesmo tipo (FURTH, 1988, p. 74). Sendo assim, a coluna vertebral, ou espinha dorsal, não vem a ser constituída por um processo meramente casual, pois o arranjo composicional do embrião é determinado pelo esperma, que, ao solidificar o resíduo menstrual, no ato da concepção, também imprime os movimentos responsáveis pela conseqüente formação das articulações estruturais, incluindo a espinha dorsal, que compõe certos organismos (*Geração dos Animais*, II.1, 734b7-12, II.4, 739b21-23; COHEN, 1996, p. 156; BALME, 1987, p. 282).

No Livro I da *Metafísica*, Aristóteles declara que grande parte dos filósofos naturalistas do século V a.C. atribui, exclusivamente, a fatores materiais, e às propriedades essenciais a eles inerentes, os princípios pelos quais todos os entes naturais, incluindo os organismos vivos, vêm a se constituírem:

Dos que primeiro filosofaram, a maioria pensara que os únicos princípios de todas as coisas são de natureza material: é do que estão constituídas todas as coisas que são, a partir do qual primeiramente se geram, e no qual ultimamente

2 Em seu artigo *Teleology and necessity*, Balme afirma que Aristóteles “vê a alternativa à teleologia não como a uma ordem universal mecanicamente determinada por um nexos de causas físicas e efeitos, no qual cada efeito é a si mesmo determinado, e causa de um previsível efeito ulterior, mas como a um caos do qual nada favorável à explicação científica poderia emergir. Os elementos agem de acordo com seus próprios movimentos naturais, mas as ações são ilimitadas”. É, no entanto, pela determinação formal, enquanto princípio anterior responsável pela constituição dos entes naturais, que estas ações dos elementos vêm a ser limitadas, e ordenadas em um estado preciso (BALME, 1987, p. 283).

se descompõe – subsistindo uma essência, modificada, porém, em suas afecções – eis o que afirmam ser elemento e princípio dos entes e, por isso, julgaram não ser verdade que algo vem a ser e se destrói, dado que essa natureza sempre se preservaria (*Metafísica*, I.3, 983b6-11).

Assim, a verdadeira natureza de todas as coisas corresponderia ao(s) elemento(s), a partir dos quais eles são primeiramente gerados, e, posteriormente, vêm a se decompor, de modo que não haveria, em última instância, geração e corrupção entre os seres que compõem a realidade natural. Conforme Aristóteles, Anaxágoras de Clazômenas (cerca de 500-428 a.C), por exemplo, defendia que os primeiros princípios, as homeomerias (τὰ ὁμοιομερής)³, eram infinitos, e que os entes naturais eram gerados e destruídos através, unicamente, de causas materiais, por mera agregação e desagregação (*Metafísica*, I.3, 984a11-19).

Então, para Aristóteles, a maioria dos filósofos naturalistas do século V a.C. defendia a tese de que os animais e as plantas desenvolvem as suas partes constituintes por acidente ou por mera espontaneidade, sob as quais subjaz a verdadeira natureza das coisas, isto é, o(s) elemento(s). Sendo assim, não haveria alguma causa intrínseca atuando no desenvolvimento dessas partes, que explicaria porque elas se formam de tais e tais modos (SAUVÉ, 1992, p. 822). Por conseguinte, a constituição dos seres vivos não seria nada mais além do que estados ou afecções passageiras de certo arranjo ou composição da matéria elementar, a qual sempre se preserva.

Nesse viés, SauvÉ irá dizer que a tese dos *physiologoi* adversária da de Aristóteles seria um tipo de concepção *eliminativista*, visto que “propõe eliminar da categoria ontológica da substância todas as outras entidades, que não os elementos” (SAUVÉ, 1992, p. 825). A autora explica que, apesar de os *physiologoi* reconhecerem a existência de animais e plantas, não contemplariam uma existência *sui generis*, não podendo ser concedido a eles, portanto, o título de substâncias propriamente ditas. Aristóteles não apenas irá negar esta posição, como irá inverter as credenciais ontológicas: estipulando o primado do

3 Na sua tradução da *Física* de Aristóteles, Echandía aponta em nota que as homeomerias “seriam para Anaxágoras, segundo Aristóteles, as partes semelhantes de que estão feitas as coisas, semelhantes a cada coisa e semelhantes entre si. Enquanto que para Empédocles um corpo por divisão se dissolve nos quatro elementos, para Anaxágoras, por mais que seja dividido, sempre se encontrará partes semelhantes, ou seja, os elementos são partes semelhantes (*De Caelo*, III.3, 302a28). No entanto, embora consagrada pela tradição doxográfica posterior (na que se usa ὁμοιομέρεια, termo utilizado pela primeira vez em Epicuro), a palavra ὁμοιομερής não aparece nos fragmentos de Anaxágoras, e é provável que tenha sido uma invenção de Aristóteles (Ross, *Metaph.* I, p. 132). A ideia, não a palavra, se encontra em Platão, *Prot.* 329d-e. A palavra usada por Anaxágoras é, antes, σπέρματα, para designar as “sementes” das coisas, inicialmente confundidas no caos primordial (DK 59 B 4)” (ECHANDÍA, 2002, p. 98, nota 48).

princípio formal em relação ao princípio material, e concedendo um papel relevante ao processo de reprodução. Pelo fato de outorgar aos seres vivos um aspecto permanente e essencial, ele elevará os organismos a um primeiro plano, de modo a relegar os elementos a um segundo plano.

Embora Aristóteles considere os animais como substâncias (οὐσίαι) paradigmáticas, de maneira alguma deixa de atribuir aos elementos um papel significativo à investigação dos seres vivos. Nas explicações naturais, eles se apresentam como condições necessárias e causas auxiliares (ANGIONI, 1999, p. 70). Com efeito, em uma passagem do *De Anima*, Aristóteles afirma:

A alguns parece que a natureza do fogo é, em absoluto, a causa da nutrição e do crescimento, posto que, pelo que se observa, é o único dentre os corpos ou elementos que se nutre e cresce; pelo que, alguém poderia julgar que também é ele que realiza tal função nas plantas e animais. Entretanto, é em certo sentido causa auxiliar, mas não causa em absoluto, que é a alma; pois o crescimento do fogo não tem limite, na medida em que houver combustível, ao passo que todos os seres naturalmente constituídos têm um limite e uma proporção quanto ao seu tamanho e crescimento (*De Anima*, II.4, 416a10-18).

Segundo Aristóteles, certos autores atribuíam exclusivamente ao elemento fogo o fator responsável pela nutrição (τροφή) e o crescimento (αύξησις) dos animais. Sendo assim, estas características vitais seriam apenas explicadas em termos de causas materiais e eficientes, a partir de processos simplesmente espontâneos, associadas a determinadas propriedades essenciais, ou formais, do fogo.

Pelo que se pode depreender do trecho acima citado, Aristóteles não nega que certas propriedades do elemento fogo desempenham um papel relevante nos processos e atividades orgânicos. Este componente material apresenta-se como condição necessária e como causa auxiliar (συναίτιον) à função nutritiva e ao crescimento do animal. Porém, recorrer tão somente às propriedades do elemento fogo - como, por exemplo, a capacidade de nutrir mediante a ação do calor, o qual permite a produção da cocção (πέψις),⁴ e

⁴ De acordo com Sánchez-Escariche e Miguel, o termo πέψις “significa tanto cocção como digestão. Designa toda mudança das matérias alimentícias dentro do corpo. [...] É um vocábulo que deriva da tradição médica” (SÁNCHEZ-ESCARICHE, Y MIGUEL, 2000, p. 90, nota 29). Em uma passagem de *Meteorológicos*, Aristóteles diz o seguinte sobre o processo de cocção: “A cocção é a consumação pelo calor natural e próprio, a partir de potenciais passivos opostos [i.e. úmido e seco]: estas são a matéria apropriada de cada coisa. De fato, quando chega à cocção, é realizada e gerada. E o princípio da realização se dá pelo calor apropriado, mesmo se levada a termo, também, mediante o concurso de algo externo, tal como, por exemplo, o alimento é digerido, também, graças aos banhos e outras coisas assim: mas o princípio é o calor presente nele” (*Mete.* IV.2 379b17- 25). Então, parece que a assimilação do alimento pelo organismo deve-se à produção de certa proporção entre características úmidas e secas pela ação do calor natural, o qual proviria, em última ins-

os movimentos delas decorrentes - não seria suficiente para explicar a apropriada porção de calor requerida à elaboração dos alimentos, por meio dos quais o animal cresce e aumenta o seu tamanho de acordo com certa proporção e limite. Esta proporção e limite são regulados pela alma (*ψυχή*) do ser vivo, através de um processo de caráter teleológico, os quais vêm a ser explicados em termos de causa formais e finais. No entanto, para Aristóteles, o todo orgânico não constitui uma mera soma de suas partes, pois a proporção e limite de seu tamanho, ou seja, a sua configuração (*σχῆμα*) não é determinada por uma simples conjunção espontânea de séries causais, ocasionada por interações entre os movimentos dos componentes materiais. A devida configuração, ou arranjo disposicional de suas partes, isto é, o modo como as partes são concatenadas para poder realizar certa função, ou um conjunto articulado de funções, é promovida pela forma do organismo, de modo que o todo orgânico vem a ser algo distinto daquilo de que é constituído⁵.

Por isso, segundo Aristóteles, Demócrito não se pronunciou corretamente ao identificar a forma com as configurações externas dos seres vivos, apreensíveis pelos sentidos. A este respeito, é dito seguinte no Livro I das *Partes dos Animais*:

Se cada um dos animais (ou cada uma de suas partes) fosse pela figura e pela cor, Demócrito teria se pronunciado de maneira correta: pois parece que ele assim concebeu. Em todo caso, ao menos, ele afirma que a qualquer um é evidente de que qualidade é o homem em sua forma, como se o homem fosse conhecido pela figura e pela cor. No entanto, também o homem morto possui a mesma forma de configuração, mas, entretanto, não é homem. Além do mais, é impossível haver uma mão disposta de qualquer maneira que houver (por exemplo, uma de bronze ou uma de madeira), a não ser homonimamente, tal como o médico desenhado. Pois ela não seria capaz de efetuar a sua função própria [...]. E semelhantemente a estes casos, tampouco nenhuma das partes do morto seria ainda de tal e tal qualidade, quero dizer, por exemplo, olho, mão (*Partes dos Animais*, I.1, 640b29-641a5).

tância, do elemento fogo. Sobre a ideia de nutrição em Aristóteles, ver: *Partes dos Animais*, II.3, 650a3 e ss.

5 Em uma passagem da *Geração e Corrupção*, Aristóteles escreve o seguinte: “Dado que há entes que são em potência e entes que são em ato, é possível que as coisas combinadas sejam em um sentido e, em outro sentido, não sejam, resultando o produto de sua combinação diverso delas em ato, mas podendo cada ingrediente ser em potência o que era antes de se combinar, e não ser destruído” (*Geração e Corrupção*, I.10, 327b23-26). O que Aristóteles pretende dizer nesta passagem é que a combinação que resulta na constituição dos organismos é, em ato, algo distinto dos elementos materiais da qual é composta, mas esses elementos não deixariam de preservar, em potência, as suas características essenciais. Deste modo, as coisas combinadas enquanto tais são, e não são, na medida em que as suas partes constituintes deixam de ser o que são em ato, e passam a ser em potência em relação ao todo do qual são partes.

De acordo com Aristóteles, uma mão – ou qualquer que seja a parte do organismo – só é o que é em função do todo orgânico do qual faz parte. Esta tese tem por base o princípio da homonímia (*Metereológicos*, IV.12, 389b31-390a2; *De Anima*, II.1, 412b11-15, 19-21; *Geração dos Animais*, I.19, 726b22-24, II.1 734b25-27, II.1 735a8-9, II.5 741a10-13; *Metafísica*, VII.10, 1035b23-25): a mão separada do corpo, ou a mão de um homem morto, só se diz mão por homonímia, tal como uma mão pintada ou esculpida, pois ela não desempenha a função que a caracteriza enquanto tal e, pela qual, definimos o que ela é.

Da mesma maneira, um homem morto pode apresentar um mesmo formato externo de um homem vivo, enquanto se mantém de algum modo preservado, ou enquanto não aparenta estar se decompondo. Mas, um homem morto não é, de fato, um homem, a não ser por homonímia. O que distingue o homem é justamente a sua capacidade de executar determinadas funções que lhe são próprias, e que o habilita a desempenhar certas atividades características, como, por exemplo, a nutrição, a respiração, o crescimento, a sensação, a locomoção etc.

Assim, conforme Aristóteles, não se deve pensar em um organismo vivo em termos de configuração externa, como Demócrito equivocadamente supunha, ao pretender que o homem poderia ser conhecido “pela figura e pela cor” (τῷ σχήματι καὶ τῷ χρώματι). Mas se deve pensar em um organismo vivo em termos de função ou, mais precisamente, em termos de um conjunto articulado de funções anímicas, em vista das quais ele vem a ser como tal e, pelo qual, apreendemos a sua definição, o seu *logos* (λόγος).

3. O princípio anímico

No primeiro capítulo do Livro II do *De Anima*, Aristóteles define a alma como “a efetividade primeira de um corpo natural que em potência possui vida” (ἐντελέχεια ἢ πρώτη σώματος φυσικοῦ δυνάμει ζῶν ἔχοντος) (*De Anima*, II.1, 412a27-28), ou ainda, após se referir às plantas, as quais representam as formas mais simples, ou mais básicas, de vida, como “a efetividade primeira de um corpo natural orgânico” (ἐντελέχεια ἢ πρώτη σώματος φυσικοῦ ὀργανικοῦ) (*De Anima*, II.1, 412b5-6), acrescentando, logo em seguida, que não cabe a pergunta a respeito de saber se “a alma e o corpo são uma única realidade, como não cabe fazer tal pergunta acerca da cera e da figura e, em geral, acerca da matéria de cada coisa, e aquilo de que é matéria” (*de An.* II.1 412b6-8). Deste modo, a alma não seria nada mais do que a “efetividade primeira” (ἐντελέχεια ἢ πρώτη)⁶ das funções vitais,

6 A expressão “efetividade primeira” diz respeito à disposição (ἔξις) ao desempenho das funções vitais, pois,

isto é, a capacidade que um corpo natural tem de desempenhar um conjunto articulado de funções, ou operações próprias, que caracterizam os seres vivos (ANGIONI, 1999, p. 78).

No Livro I das *Partes dos Animais*, Aristóteles prescreve que “compete ao estudioso da natureza afirmar e conhecer a respeito da alma” (*Partes dos Animais*, I.1, 641a21-22)⁷, pois sendo a alma o princípio do animal (*De Anima*, I.1, 402a6-7), e, portanto, causa através da qual ele vem a ser o que é, ela explica porque a matéria do organismo apresenta certas propriedades características, que são requisitadas pelas exigências da vida. Assim, em consonância com o *hilemorfismo* (teoria da matéria e forma), tal como formulado no Livro II da *Física*, a ciência relativa aos seres vivos deve levar em conta tanto a forma quanto a matéria em suas investigações, visto que a alma, correspondendo à forma do animal, é o fator responsável por aquilo a partir de que ele vem a se constituir. Através da composição material, a própria forma ou alma do vivente se realiza enquanto essência e acabamento, na medida em que esta composição possibilita ao organismo o exercício das funções vitais pelas quais o ser vivo é definido.

Com efeito, Aristóteles irá dizer, no Livro I das *Partes dos Animais*, que a alma é “como movente e como acabamento” (ὡς ἡ κινουσα καὶ ὡς τὸ τέλος) (*PA* I.1 641a27). De acordo com isto, no Livro II do *De Anima*, Aristóteles argumenta que na alma, enquanto princípio e causa do corpo vivente, convergem os modos de causalidade formal, final e eficiente⁸:

A alma é causa conforme três dos modos delimitados: pois ela é causa enquanto princípio do movimento, bem como aquilo em vista de que, e, também, como essência dos corpos animados. Ora, que a alma é causa como essência é evidente: pois, para todos os entes, o responsável pelo ser é a essência, e o ser para os que vivem é o viver, e deste, é a alma que é causa e princípio. Além disso,

Aristóteles afirma que a efetividade anímica poderia ser dita “de duas maneiras, uma como conhecimento, enquanto que outra, como o estar contemplando”, sendo que a alma é efetividade, primeiramente, “como conhecimento: pois tanto o sono, como a vigília, ocorrem quando há uma alma, e a vigília é análoga, por seu lado, ao estar contemplando, ao passo que o sono é análogo ao possuir conhecimento, mas não exercê-lo” (*De Anima*, II.1, 412a22-26). Em outras palavras, uma pessoa, por exemplo, que não é cega, durante o sono, ou quando está com os olhos fechados, havendo, evidentemente, luz, não está vendo; porém, ela não deixa de ter a capacidade ou disposição de ver. É justamente a esta capacidade, ou disposição a realizar determinada função vital, que Aristóteles se refere por “efetividade primeira”. Como Hamlyn (1993, p.85) notou, Alexandre de Afrodísias (198–209 d.C), comentador antigo de Aristóteles, declarou que “a efetividade primeira é a ἔξις, a qual é pressuposta pela ἐνέργεια”.

⁷ Ver, também: *De Anima*, I.1, 403a28; *Metafísica*, VI.1, 1026a5- 6.

⁸ Esta ideia encontra-se, também, na *Física*: “Muitas vezes, estas três [*sc.* causas] convergem para uma só coisa: o ‘o que é’ e aquilo em vista de que são uma só, e lhes é especificamente idêntico aquilo de que procede primeiramente o movimento, pois é um homem que gera um homem” (*Física*, II.7, 198a34- 27).

a efetividade é determinação daquilo que é em potência. E é manifesto que a alma é causa também como aquilo em vista de quê [...]. Nos animais, é a alma que conforme a natureza é um acabamento, pois todos os corpos naturais são instrumentos da alma, e assim com o são as partes dos animais [...]. Mas, seguramente a alma é, também, aquilo de onde inicialmente se dá o movimento (*De Anima*, II.4, 415b8-21).

Do ponto de vista do ser, ou da essência (οὐσία), o princípio anímico, ou a alma, atua como causa formal ao promover a realização do que faz com que o ser vivo seja exatamente o que ele é: um ser animado, sendo, por isso, capaz de exercer as suas atividades próprias, que correspondem à reprodução, à nutrição, ao crescimento, à sensação, à locomoção etc. E, dado que as atividades ou funções vitais caracterizam os seres vivos enquanto tais, essas devem constar no enunciado definitório do vivente.

Contudo, do ponto de vista do vir a ser, a alma atua como causa final, pois é em vista do ser animado que ocorrem os processos dos quais resultam o organismo vivo. Ademais, o princípio anímico atua como causa eficiente, (i) na medida em que a forma contida no sêmen do progenitor, no ato da concepção, engendra, em condições normais⁹, a sequência dos movimentos necessários à formação de um novo indivíduo especificamente idêntico, e (ii) na medida em que a forma do animal determina um conjunto articulado de funções, permitindo aos organismos reproduzirem-se, nutrirem-se, locomoverem-se etc. (ANGIONI, 1999, p. 79-80; BALME, 2001, p. 89-90).

Por conseguinte, a natureza material do animal é devido à alma, não sendo a inversa verdadeira (*Partes dos Animais*, I.1, 641a30-31), pois é a partir do princípio anímico que se delimitam as propriedades relevantes e se desenvolve o devido arranjo composicional que capacita o organismo executar as funções requisitadas pelas necessidades vitais. Deste modo, na constituição orgânica, (i) os movimentos necessários e (ii) as propriedades essenciais dos elementos materiais são, respectivamente, reconduzidos – ao ser desviado o curso que seguiriam espontaneamente –, e apropriadas pela forma do animal, vindo a ser alteradas.

Estas propriedades são alteradas não porque, através do princípio anímico, elas destoariam de um comportamento normal, mas sim porque resultam, simplesmente, de um processo de interação com outras propriedades. Assim, a alma não seria algum tipo de “substância imaterial”, a qual intercederia sobre a matéria elementar, de modo a alterar, completamente, as suas propriedades essenciais, ou atribuisse algo além do que

⁹ Há casos atípicos nos quais os movimentos que procedem do sêmen são enfraquecidos, resultando na ocorrência dos monstros, isto é, na ocorrência de organismos mal formados (*Geração dos Animais*, IV.3, 769b11-18).

poderia resultar (i) de uma composição na qual se estabelece uma mistura (μίξις)¹⁰, e da influência exercida pelo calor inato (*Partes dos Animais*, II.3, 649b28-33), que auxilia nas funções da alma (*Partes dos Animais*, II.7, 652b10-13).

Tal calor é primeiramente transmitido ao embrião através do πνεῦμα, ou sopro vital, contido no sêmen (*Geração e Corrupção*, II.3, 736b34-37), mas cuja principal fonte, após a fecundação, procede do coração (*Parva Naturalia*, 469b10-12), o qual é o primeiro órgão a se formar no feto, antes mesmo da formação das outras partes orgânicas (*Partes dos Animais*, III.4, 666a25-27). Portanto, a alma é causa pela qual vêm a ser ajustadas às devidas disposições dos elementos materiais,¹¹ os quais, “deixados à própria sorte”, sem a intervenção de certa causalidade reguladora, seriam incapazes de promover adequadamente as complexas redes de propriedades e estruturas que os seres vivos apresentam.

4. Natureza e alma

Não há dúvidas que, para Aristóteles, a ciência da natureza deve tratar da alma, enquanto fator responsável pela determinação daquelas funções vitais comuns a todos os animais, as quais já foram mencionadas anteriormente: reprodução, nutrição, sensação, locomoção etc. Porém, comporta dificuldade saber se caberia à ciência natural ocupar-se, também, da faculdade intelectual do ser humano; dificuldade esta formulada por Aristóteles, no primeiro capítulo do Livro I das *Partes dos Animais*, da seguinte maneira:

Compete à ciência da natureza afirmar a respeito de toda alma, ou a respeito de alguma? Pois, se for a respeito de toda alma, não restará nenhuma filosofia além da ciência da natureza, pois a inteligência é dos inteligíveis. Por conseguinte, a ciência da natureza seria conhecimento de tudo; pois, compete a uma mesma ciência contemplar a respeito da inteligência e a respeito do inteligível (*Partes dos Animais*, I.1, 641a34-38).

Antes de entrarmos no cerne da questão, convém fazermos um breve relato sobre o modo pelo qual Aristóteles concebe, no *De Anima*, a distinção entre as partes ou faculdades da alma.

¹⁰ De acordo com Aristóteles, “a mistura é uma união de coisas combinadas que têm sofrido alteração” (*Geração e Corrupção*, I.10, 328b22). A mistura acontece quando há certo grau de equilíbrio entre os elementos de uma combinação, de modo que “cada um transforma sua própria natureza, sem converter-se no outro, mas em algo intermediário e comum” (*Geração e Corrupção*, I.10, 328a28-31).

¹¹ Como Balme (2001, p. 90) escreve: “Ele [Aristóteles] não considera que uma tendência diretiva tem que ser sobreposta sobre interações físicas e químicas normais, mas que interações normais incluem uma tendência diretiva. Por “causa”, ele quer dizer um fator explicativo, não necessariamente um item separado.

Conforme Aristóteles, a alma pode ser analisada em potencialidades ou faculdades (δυνάμεις) (*de An.* II.3 414a29-32), as quais perfazem uma série ordenadamente determinada, sendo que, cada faculdade, situada num plano posterior, pressupõe outra(s), situada(s) em um plano anterior (*De Anima*, II.3, 414b31-415a3). A mais fundamental diz respeito à alma nutritiva (τὸ θρεπτικόν) e à reprodutiva (τὸ γεννητικόν), visto que existe em todos os organismos vivos: plantas e animais (*De Anima*, II.3, 414a33).

Em seguida, vem a alma sensitiva (τὸ αἰσθητικόν), da qual estão excluídas as plantas. Na alma sensitiva, apresenta-se uma equivalente série ordenada, na qual certas faculdades sensitivas pressupõem outras, assentadas num plano mais fundamental. A faculdade sensitiva mais básica corresponde ao tato¹², a qual pressupõe as demais, a saber, as relativas ao olfato, à audição e à visão, de modo que se houver estas últimas três, sempre, haverá a primeira, e, por vezes, acontece de o tato ocorrer sem as demais, “pois muitos dos animais não possuem nem vista, nem audição, nem sensação do cheiro” (*De Anima*, II.3, 415a5-6).

No entanto, a alma sensitiva não se restringe às funções perceptivas, mas, como consequência disto, compreende, também, o prazer ou a dor e, por isso, a capacidade desiderativa (τὸ ὀρεκτικόν), presente em todos os animais (*De Anima*, II.3, 414b1-2). Além do mais, há duas capacidades que derivam da faculdade sensitiva, mas que não se encontram em todos os animais: trata-se da capacidade (i) imaginativa e (ii) motriz (*De Anima*, II.3, 414b16-17).

Por fim, há a faculdade discursiva (τὸ διανοητικόν) ou intelectual (τὸ νοητικόν) da alma, associada à razão, a qual pertence exclusivamente ao ser humano (*de An.* II.3 414b18). Tal faculdade é considerada como especificamente distinta da sensitiva, entretanto, na medida em que a sensação (αἴσθησις) atua de uma maneira geral em relação aos sensíveis próprios ligados aos órgãos sensoriais, como a audição, a visão etc., ou seja, na medida em que a sensação atua enquanto percepção dos sensíveis comuns: movimento, repouso, figura, tamanho, número e unidade (*De Anima*, III.1, 425a14-20), se estabelece um vínculo entre os sentidos e a inteligência (*De Anima*, III.8).

Como Balme (2001) observou, para Aristóteles, “os animais podem ser habilitados e inteligentes (φρόνιμος) (*Metafísica*, I.1, 980a27-980b25), mas não podem contemplar (θεωρεῖν), isto é, apreender universais”. Os inteligíveis, acrescenta Balme, são

12 De acordo com Aristóteles, o gosto ou paladar é uma espécie de tato, pois, tal como este, o gosto é percebido por um contato direto entre o objeto tangível e o corpo, e não através de um meio transmissor, como acontece no caso do olfato, da audição e da visão (*De Anima*, II.10, 422a8-11, III.12 434b13 e ss.; *Partes dos Animais*, II.10, 657a1, II.17 660a23).

considerados “unicamente como objetos que não se movem ou sofrem mudanças, ou seja, formas abstraídas ou conceitos matemáticos” (BALME, 2001, P. 9). Entretanto, toda ciência, ou técnica, nasce da experiência, a qual surge aos homens através da memória. A memória (μνήμη), por sua vez, se desenvolve a partir da imaginação (φαντασία) (*Parva Naturalia*, 450a22-24),¹³ e a capacidade imaginativa procede da faculdade sensitiva¹⁴. Apesar de alguns animais, além do homem, possuírem memória, a experiência adquirida por ela não é suficiente para gerar conhecimentos. Neste sentido, Aristóteles diz o seguinte no Livro I da *Metafísica*:

Os outros animais vivem com base em impressões e recordações, mas compartilham pouco da experiência; o gênero dos homens, por sua vez, vive também com técnica e raciocínios. É da memória que a experiência surge aos homens: diversas recordações de um mesmo fato perfazem a capacidade de uma experiência. A experiência parece muito similar à ciência e à técnica, mas na realidade é através da experiência que os seres humanos obtêm ciência e técnica (*Metafísica*, I.1, 980b25-981a4).

Então, retornemos à questão de saber se caberia à ciência natural ocupar-se, também, da faculdade intelectual do ser humano, quando Aristóteles discute se competiria “à ciência natural afirmar a respeito de toda alma, ou a respeito de alguma”, de modo que, “se for a respeito de toda alma, não restará nenhuma filosofia além da ciência da natureza”.

Parece que o filósofo pretende dizer que a faculdade sensitiva, até o ponto em que é possível alcançar um conjunto de recordações que possibilitam ao homem adquirir experiências (ἐμπειρίαι), através das quais se obtêm a ciência, pertence ao domínio do estudo natural; porém, a faculdade intelectual, a qual está associada à atividade científica e à ciência como tal, não, pois, aquilo pelo que a alma pensa (διανοεῖται) e discerne (ὑπολαμβάνει) se estabelece a partir das formas abstraídas da matéria (*De Anima*, III.4, 429a22-29).

As formas, consideradas em si mesmas, não experimentam quaisquer tipos de mudança, mas o objeto de estudo da ciência da natureza diz respeito aos seres providos de matéria e forma, caracterizados, justamente, por apresentarem um princípio de mudança. Penso que é por esta razão que Aristóteles, no Livro I das *Partes dos Animais*, diz

13 O sentido de φαντασία se aproxima ao de φανεῖσθαι (aparecer), designando tanto a aparição de um objeto quanto ao processo mental daquele que o percebe (MARTÍNEZ, 2003, p. 227, nota 71).

14 Segundo Aristóteles, o termo φαντασία (imaginação) deriva do termo φάος (luz). O vínculo etimológico entre φαντασία e φάος se deve ao fato de que a apreensão da imagem perceptiva, através da visão, só é possível por intermédio da luz (*De Anima*, III.3, 429a3-4).

“que não se deve afirmar a respeito de toda a alma: pois não é toda alma que é natureza, mas sim alguma ou até mesmo várias partes dela” (*Partes dos Animais*, I.1, 641b8-10).

5. Considerações Finais

Quando Aristóteles se refere a almas, ou a partes da alma, entendo que não se deveria pensar em tais referências a partir de um ponto de vista ontológico, como se houvesse substancialmente diversas almas, ou que houvesse distintas regiões, nas quais as partes da alma estariam localizadas, visto que se trata de diferentes atribuições de uma mesma e única alma. Assim, quando Aristóteles diz que não é toda alma que é natureza, reportando-se, pelo que se pode depreender do contexto, à faculdade intelectual do homem, ele não estaria defendendo a ideia de que o intelecto pertence a um domínio extra-natural, de modo que manteria uma existência separada do corpo como se, durante a vida da pessoa, permanecesse de algum modo associado a ele.

Como Marco Zingano constatou no seu *Ensaio* sobre os capítulos quatro e cinco do Livro III do *De Anima*: “O intelecto é sempre uma parte da alma (aquela pela qual a alma pensa e concebe), e a alma é a forma de um corpo, não podendo existir separadamente (ZINGANO, 1998, p. 176)”.

De fato, Aristóteles sustenta a ideia de que a alma, apesar de não ser um corpo, é algo de um corpo, e somente se realiza nele (*De anima*, II.2, 414a19-21). No entanto, acontece que a alma pode ser analisada sob várias perspectivas, e a faculdade intelectual sob duas: (i) a atividade de conhecer de um indivíduo, e (ii) o conhecimento como resultado dessa atividade (ZINGANO, 1998, p. 177).

No que diz respeito ao item (i), a atividade de conhecer corresponde a certa função da alma, a qual não ocorre senão por meio do organismo humano, mas se trata de abstrações, ideias, pensamentos, que são elementos noéticos e não materiais. E quanto ao item (ii), o conhecimento como resultado desta função anímica, isto é, a ciência, não é natureza, pois se trata de uma instância puramente epistemológica, não substancial. Neste sentido, Lennox declara o seguinte:

Os objetos da razão teórica são abstrações. Tais objetos, sendo imateriais, não mudam [...]. Eles não são, portanto, objetos naturais. Se estudar a razão é estudar seus objetos, e esses objetos são abstrações, o estudo da razão teórica seria excluído por esse motivo (LENNOX, 1999, p. 5).

Sendo assim, seria tão somente a alma do ser humano, analisada sob o seu aspecto intelectual, que não competiria à ciência da natureza investigar, apesar de a capacidade intelectual da alma não deixar de estar vinculada ao corpo humano, o qual corresponde

a determinado composto de matéria e forma natural. Isto porque o objeto de estudo da ciência da natureza se limitaria, de um modo estrito, às substâncias naturais compostas de matéria e forma, de maneira que os objetos de caráter abstrato da razão teórica não se ajustaria, apropriadamente, ao domínio de investigação da ciência natural.

Referências

Obras de Aristóteles

ARISTÓTELES. **Aristotelis Opera**. BEKKER, I. (Ed.). Vols. I-II. Berlim, Academia Regia Borussica, 1831.

ARISTÓTELES. **On the heavens**. Trad. by W. K. C. Guthrie. London, Loeb Classical Library, 1939.

ARISTÓTELES. **The metaphysics**, Books I-IX. Trad. Hugh Tredennick. London, The Loeb Classical Library, 1947.

ARISTÓTELES. **Meteorologica**. Trad. by H. D. P. Lee. London, The Loeb Classical Library, 1952.

ARISTÓTELES. **Generation of animals**. Trad. A. L. Peck. London, The Loeb Classical Library, 1953.

ARISTÓTELES. **On sophistical refutations; On coming-to-be and passing away; On the cosmos**. Trad. E. S. Foster, and D. J. Furley. London, The Loeb Classical Library, 1955.

ARISTÓTELES. **On the soul; Parva naturalia, On breath**. Trad. W. S. Hett. London, The Loeb Classical Library, 1957.

ARISTÓTELES. **Physics**. Trad. by P. M. Wicksteed, and F. M. Conford. London, The Loeb Classical Library, 1957.

ARISTÓTELES. **Parts of animals; Movement of animals; Progression of animals**. Trad. A. L. Peck. London, The Loeb Classical Library, 1961.

ARISTÓTELES. **Reproducción de los animales**. Trad. E. Sánchez. Madrid, Biblioteca Clásica Gredos, 1994.

ARISTÓTELES. **Acerca del cielo; meteorológicos**. Trad. y notas de M. Candel. Madrid, Biblioteca Clásica Gredos, 1996.

ARISTÓTELES. **Acerca de la generación y la corrupción; tratados breves de historia natural.** Trad. E. La Croce, e A. B. Pajares. Madri, Biblioteca Clásica Gredos, 1998.

ARISTÓTELES. As partes dos animais, Livro I. Trad. e comentários de L. Angioni. **Cadernos de História e Filosofia da Ciência**, 9 (3), número especial, 1999.

ARISTÓTELES. **Partes de los animales; Marcha de los animales; Movimiento de los animales.** Trad. E. J. Sánchez-Escariche y A. A. Miguel. Madrid, Biblioteca Clásica Gredos, 2000.

ARISTÓTELES. **Física.** Trad. G. R. de Echandía. Madrid, Biblioteca Clásica Gredos, 2002.

ARISTÓTELES. **Acerca del alma.** Trad. T. C. Martínez. Madrid, Biblioteca Clásica Gredos, 2003.

ARISTÓTELES. **Metafísica.** Trad. T. C. Martínez. Madrid, Biblioteca Clásica Gredos, 2006.

ARISTÓTELES. **Física I-II.** Trad. e comentários de L. Angioni. Campinas, Ed. Unicamp, 2009.

Literatura secundária

ANGIONI, L. As partes dos animais, Livro I. Tradução e comentários de L. Angioni. **Cadernos de História e Filosofia da Ciência**, 9 (3) (n. especial), 1999.

BALME, D. M. Teleology and necessity. In: A. GOTTHELF & LENNOX, J. (Eds.). **Philosophical issues in Aristotle's biology.** Cambridge, Cambridge University Press, 1987, p. 275-285.

BALME, D. M. **Aristotle de partibus animalium I and de generatione animalium I:** with passages from ii. 1-3. Translated with notes by D. Balme. Oxford, Clarendon Press, 2001 ("Clarendon Aristotle Series").

COHEN, S. M. **Aristotle on nature and incomplete substance.** New York, Cambridge University Press, 1996.

ECHANDÍA, G. R. de. **Aristóteles – Física.** Madri, Biblioteca Clásica Gredos, 1998.

FURTH, M. **Substance, form and psyche:** an Aristotelian metaphysics. Cambridge, Cambridge University Press, 1998.

HAMLIN, D. W. **Aristotle – de Anima:** Books II and III (with passages from Book I). Oxford, Clarendon Press, 1993.

LENNOX, J. G. The place of mankind in Aristotle's zoology. **Philosophical Topics**, Vol. 27, Nº 1, 1999.

LENNOX, J. G. **Aristotle: On the parts of animals**, Books I-IV. Translated with an Introduction and Commentary by J. Lennox. Oxford, Clarendon Press, 2001.

MARTÍNEZ, T. C. **Acerca del alma**. Madrid, Biblioteca Clásica Gredos, 2003.

ROSS, D. **Aristotle's Metaphysics**. A revised text with introduction and commentary. 2 vols. Oxford, Clarendon Press, 1924.

SAUVÉ, S. M. Aristotle, teleology and reduction. **Philosophical Review** 101, n. 4, p. 791-825, 1992.

SÁNCHEZ-ESCARICHE, E. J. & MIGUEL, A. A. **Partes de los animales; Marcha de los animales; Movimiento de los animales**. Madrid, Biblioteca Clásica Gredos, 2000.

SÁNCHEZ, E. **Reproducción de los animales**. Introducciones, traducciones y notas de E. J. Sánchez-Escariche. Madrid, Biblioteca Clásica Gredos, 1994.

ZINGANO, M. A. **Razão e sensação em Aristóteles: um ensaio sobre de Anima III 4-5**. Porto Alegre, L&PM editores, 1998.





O ABSURDO DA MATÉRIA E A PALAVRA EXTRAVIADA: A CRÍTICA DE BERKELEY AOS ASPECTOS ESTRUTURANTES DA TEORIA DO CONHECIMENTO



Luama Socio¹

Resumo:

Abordar a filosofia de Berkeley justifica a impressão de que seja especialmente esclarecedor voltar a questões filosóficas através de uma perspectiva mais elementar ou mais radical. Se pensamos geralmente que a abstração tem sua positiva existência entre os fatos, porque é que ela não se esclarece como objeto à luz da percepção? É certo que os tempos, a partir do século XX, trouxeram fecundas elaborações para os ângulos mais inquietantes das relações entre linguagem, sujeito e objeto, dos dualismos conceituais formados pelas oposições entre idealismo e realismo, subjetividade e objetividade, totalidade e relativismo ou fragmentação etc. No entanto, a perspectiva de Berkeley continua desafiadoramente fecunda à medida que une em-

Abstract:

Approaching Berkeley's philosophy justifies the impression that it is especially illuminating to return to philosophical questions through a more elementary or more radical perspective. If we generally think that abstraction has its positive existence among facts, why is it not clarified as an object in the light of perception? It is true that times, from the 20th century onwards, have brought fruitful elaborations to the most disturbing angles of the relations between language, subject and object, of the conceptual dualisms formed by the oppositions between idealism and realism, subjectivity and objectivity, totality and relativism or fragmentation, etc. However, Berkeley's perspective remains defiantly fruitful as it unites empirical with spirit, persisting at the conflict

¹ Doutora em Filosofia pela Universidade de São Paulo (USP). Graduada e mestra em Letras pela Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP/IBILCE). Docente do Curso de Graduação em Letras da Universidade Estadual do Tocantins (Unitins). Lattes: <http://lattes.cnpq.br/9753517040257731>. ORCID: <https://orcid.org/0000-0002-1837-2166>. E-mail: luamasocio@gmail.com

pírico com espírito persistindo no ponto de conflito do limite entre materialismo e metafísica ao trabalhar a investigação do funcionamento e da natureza da percepção dos sentidos, das ideias e da linguagem.

Palavras-chaves:

Berkeley; filosofia da ciência; linguagem; abstração.

point of the boundary between materialism and metaphysics when working on the investigation of the functioning and nature of the perception of senses, ideas and language.

Keywords:

Berkeley; philosophy of Science; language; abstraction.

A matéria existe a partir da mente

Berkeley propõe o desafio da demonstração de objetos exteriores indicando a contradição existente entre admitir que o que se vê, ouve ou sente é percebido sempre por uma mente e, ao mesmo tempo, afirmar que tais coisas são ocasionadas por objetos exteriores a ela:

(...) dar-vos-ei plena razão (...), ainda que só possais dar-me como razão acreditar na sua existência ou indicar algum uso para o que se supõe existir. (...) Ao esforçarmo-nos no máximo para conceber a existência de corpos externos, contemplamos sempre e somente as nossas próprias ideias (BERKELEY, 1984, p. 17).

A questão da inexistência da matéria como um ser exterior inerte, independente do percipiente, dentro de um contexto filosófico rigidamente empirista, é o cerne da contenda que Berkeley propõe à filosofia de Locke.

Para este último, o percurso inicial do processamento das ideias é o seguinte: as ideias provêm dos sentidos; os sentidos percebem qualidades, que são elas mesmas as ideias em seus princípios; as qualidades estão nas coisas; portanto as qualidades estão, primeiramente, fora da mente e são levadas das coisas à mente pelos sentidos.

Assim, os sentidos configuram o próprio “saber”, contíguos às ideias, que, por sua vez, configuram mais precisamente o conhecimento. Os sentidos discriminam as qualidades antes da intervenção da razão e antes da palavra.

Em contraposição, Berkeley apontará para a limitação desse conceito de “ideia”, qual seja, o de significar o único e restrito conteúdo do conhecimento, sempre retirado da matéria exterior ao ser. Por conseguinte, aspectos da percepção não sujeitos à representação, tais como o espírito e as “relações” concebidas pela razão ou intuição, devem ser considerados objetos do conhecimento:

Comprendemos nossa existência por intuição ou reflexão; a dos outros espíritos pela razão. Pode dizer-se termos algum conhecimento ou noção do nosso próprio espírito e seres ativos mas em estrito sentido não temos ideias deles. (...) A mim parece-me que *ideias, espíritos e relações* são na sua respectiva espécie o objeto do conhecimento humano e o sujeito do discurso; e que o termo *ideia* seria impropriamente aplicado a cada coisa conhecida ou de que temos noção (BERKELEY, 1984, p. 30).

A crítica à noção de “ideia” como fundamento único do conhecimento relaciona-se com a crítica à noção de independência da matéria com relação à mente. Para Berkeley (1984, p. 16), a própria admissão, por parte de Locke, de que há ideias geradas pela mente, à parte do funcionamento dos sentidos, evidencia “a desnecessidade de corpos exteriores à produção de nossas ideias, desde que nos concedem que elas sejam produzidas às vezes e podem talvez sê-lo sempre na mesma ordem presente, sem o seu auxílio”.

As qualidades percebidas pela mente são expressões dessa mesma mente, na exata medida em que são ideias produzidas por ela. Os objetos não possuem qualidades em si. Isso é evidenciado pelo fato de os seres poderem perceber qualidades diferentes em objetos supostamente iguais, ou seja, da mesma espécie: “porque sem alteração do objeto o doce pode volver-se em amargo” (BERKELEY, 1984, p. 15). Daí infere-se que o espírito é variado, dentro de sua totalidade. Nessa concepção, a própria percepção da ideia só é possível porque há identidade entre o ser e aquilo que ele percebe. “A ideia só pode assemelhar-se a outra ideia; portanto, nem elas nem os seus arquétipos podem existir em uma substância incapaz de perceber” (BERKELEY, 1984, p. 14).

Esse ponto apresenta a clássica alternativa à refutação recorrente que se faz à filosofia de Berkeley, imputando-a ingênuo subjetivismo sobre o tema da polaridade entre objetividade e subjetividade aos modos de um Thomas Nagel (2004, p. 38-40):

Nenhuma concepção objetiva do mundo mental poderá incluí-lo por completo. (...) O mundo simplesmente *não é* o mundo que se revela a um único ponto de vista altamente abstrato que pode ser perseguido por todos os seres racionais. (...) O mundo não é o meu mundo, ou o nosso mundo – nem mesmo o mundo mental é assim. (...) Ainda assim, mesmo que o conhecimento objetivo

seja apenas parcial, vale a pena tentar ampliá-lo, por uma razão simples. A busca de um entendimento objetivo da realidade é a única maneira de expandir nosso conhecimento do que existe para além da aparência que ela tem para nós.

Esse ponto fulcral, que grosso modo integra o fundamento dos limites entre posições materialistas e metafísicas que, na modernidade, incidirá mais particularmente sobre o problema da determinação dos sujeitos do conhecimento e todas as consequências éticas relacionadas ao desenvolvimento das ciências e seus contextos, parecia especialmente desafiador e de relevante importância para Berkeley.

Em sua filosofia, o espírito causa a ideia, mas a maior parte de toda ideia não depende da vontade individualizada do ser humano. Há ideias compartilhadas entre os espíritos individuais, causadas por um espírito em comum e ao mesmo tempo “outro”: o espírito soberano, Deus. “Ninguém ignora que as chamadas obras da natureza, isto é, a maior parte das nossas sensações e ideias não são produzidas pela vontade humana nem dependentes dela. Há pois algum outro Espírito que as causa, visto não poderem subsistir por si” (BERKELEY, 1984, p. 42).

Sob essa perspectiva, o homem vive, efetivamente, num contexto mental, e o espírito individual, para conhecer a realidade, poderá discernir entre ideias realizadas pelo “Espírito Soberano” e as realizadas pela vontade individual. As primeiras parecem mais reais que as segundas. No entanto, ambas só “podem existir apenas no espírito que percebe” (BERKELEY, 1984, p. 19).

Esse viés explicativo expressa a contradição inerente à concepção convencional de matéria no contexto empirista, qual seja, a contradição de conciliar a própria validade de um conhecimento que, por sua vez, aparece contíguo à óbvia modificação do objeto de acordo com a mudança do ponto de vista existente em qualquer possibilidade de experiência.

Aqui Berkeley opera a inversão do sentido do conhecimento. Ao considerar o ser como o próprio movimento da objetividade, este movimento perceptivo expressará a norma, a natureza, a correta expressão do espírito que aparece como objetividade na elaboração do conhecimento. A mobilidade do ser humano, bem como sua constituição específica, condiciona o ponto de vista de forma inexorável e espontânea dentro do paradigma da realidade. O que poderá impedir o que se chama de clara objetividade é sempre a confusão causada pela linguagem, pela palavra extraviada de um contexto dado:

Suponha-se, por exemplo, que, olhando para a Lua, eu diga que ela está a 50 ou 60 raios terrestres distantes de mim. De que Lua se está falando? É claro que não pode ser a Lua visível, ou algo similar à Lua visível, isto é, a Lua que eu vejo, que é apenas uma figura plana redonda e luminosa, com cerca de 30 pontos visíveis de diâmetro. Pois, caso eu fosse transportado do lugar em que estou diretamente em direção à Lua, é claro que o objeto varia à medida que avanço, e, no momento em que eu tiver percorrido 50 ou 60 raios terrestres, não estarei de modo algum próximo a uma figura plana pequena, redonda e luminosa, e não perceberei nada semelhante a ela, pois esse objeto há muito desapareceu e, se eu quiser recuperá-lo, só posso fazê-lo retornando à Terra da qual parti (BERKELEY, 2010, p. 61).

É possível interpretar esse aspecto da filosofia de Berkeley, qual seja, a questão da palavra extraviada, como um elemento fácil de associar à ideia de ciência como um campo de criatividade linguística ilimitada, pois, à medida que se elimina a dualidade sujeito-objeto, elimina-se também a abstração espaço-tempo, a qual, como toda abstração, não tem existência em si independente do percipiente, mas vincula-se justamente à configuração humana em seu existir tornado linguagem. Espaço e tempo são configurações da ação e do ser, dados por movimento e linguagem, porque só existem aqueles por consequência destes. A ciência doravante terá que se preocupar em não se deixar cair nos ilusionismos espaço-temporais e Berkeley se dedicará a elucidar alguns desses ilusionismos em seus trabalhos de investigação sobre o movimento e a visão. O espaço está relacionado a uma clara corporalidade e ao movimento, na medida em que este seja compreendido como relação de posição entre corpos:

Quando provoço um movimento do meu corpo, se não há resistência digo que há *espaço*; se há resistência digo que há *corpo* (...). Assim, quando falo de espaço puro ou vazio não deve supor-se que a palavra *espaço* representa uma ideia distinta de ou concebível sem corpos e movimento (BERKELEY, 1984, p. 36).

Em conjunção a isso, Berkeley não vê uma estabilidade do que se chama matéria, mas vê, em seu lugar, um constante fluir da realidade mental ou espiritual. O ser aparece através do próprio pensamento, do próprio conhecimento. O tempo é sua manifestação, seu produto e realização:

Não sendo o tempo coisa alguma, abstraído da sucessão de nossas ideias, segue-se que a duração de um espírito finito deve medir-se pelo número de ideias ou ações sucessivas no mesmo espírito. Consequentemente a alma pensa sempre. E na verdade não parece fácil tarefa tentar separar nos pensamentos ou abstrair a *existência* do espírito, da sua *cognição* (BERKELEY, 1984, p. 32).

As consequências dessa abordagem rigorosamente empirista associada a uma compreensão da maleabilidade criativa da mente apontam desde então para o paradoxo dos valores de neutralidade, imparcialidade e autonomia das organizações operacionais técnicas dos desenvolvimentos científicos a partir da modernidade. Uma possível consequência da visão berkeleyana de que não há distância entre sujeito e objeto seria difícil de suportar num contexto científico e filosófico que se desenvolveu através da concepção materialista de “domínio da natureza”, segundo a qual o mundo é concebido como algo diferente do sujeito que o apreende, e também como algo passivo, modelável, acessível ao controle e apreensível pela prática científica². Ironicamente, a ideia de ciência também se desenvolveu através da ideia de linguagem, iluminada sob a filosofia de Wittgenstein e variações, tal como a de Kuhn, ao longo do século XX, expressando sua natureza inerentemente subjetiva.

A palavra extraviada aloja-se no senso comum

O primeiro diálogo de “Hilas e Filonous”³ parte do princípio de que o pensamento pode ser caracterizado de acordo com a classificação em duas categorias, a saber: há o pensamento do vulgo, senso comum; na sua maior parte, enredado nas abstrações originando fantasias sem referenciais, que ocupam o lugar da verdade confundindo a mente, e há o pensamento do filósofo. Filonous vai mostrar que o pensamento (palavras) do vulgo parece útil, mas que, para o filósofo, esse pensamento, na verdade, é supérfluo, ao passo que o pensamento (palavras) do filósofo parece supérfluo ao vulgo, mas, na realidade, é a revelação de uma prática específica que conduz à verdade (a filosofia). O método do filósofo se expressa da seguinte maneira: usando de um estratagema paradoxal em seu discurso, Filonous aponta para o “nada da matéria” através da palavra, duvidando da mesma palavra.

Se, por um lado, assim se delineia a noção de uma forma específica para o discurso filosófico, indicando a palavra como instrumento de uma prática, ao mesmo tempo, por outro, Berkeley esvazia a importância da palavra quando utilizada pelo senso comum, na maior parte das vezes, palavra abstrata, o que equivale à palavra sem sentido.

² “O materialismo científico propõe que podemos controlar um determinado objeto porque o apreendemos tal como ele é, e como um objeto *do mundo* que subsiste independentemente das percepções e práticas humanas?” (LACEY, 2008, p. 40).

³ BERKELEY, *Três diálogos entre Hilas e Filonous em oposição aos cétricos e ateus*.

Resgatando a gênese das ideias, Filonous afirma a percepção como faculdade da mente e os sentidos, se dando no corpo, como partes integrantes dessa mesma mente: “(o prazer) não pode existir sem mente; quer dizer, não pode existir numa substância impercepcionante, isto é, num corpo” (BERKELEY, 1965, p. 12). Portanto, o conceito de percepção separado do conceito de mente seria uma ilusão abstracionista, facilmente propagada pelo vulgo.

A diferença entre substância e qualidade, contida em dois conceitos separados supostos pelo senso comum, é dissolvida quando Filonous aponta para o fato de que essas pretensas duas coisas são realmente percebidas numa unidade através da experiência. Para Filonous, essa é a comprovação de que ambos os conceitos ocupam o mesmo plano de um raciocínio, sendo absurdo desnivelá-los hierarquicamente. A substância não contém a qualidade porque ela e a qualidade são a mesma coisa. Assim, o que é captado pelos sentidos é de fato parte da mente, manifestação da mente. Filonous defende a unidade de mente e matéria a partir da crítica feita ao argumento do *substratum material*, a relação imaginária entre substância e qualidade, presente no raciocínio de Hílas, a quem “se não concebem qualidades sem que exista também o seu suporte” (BERKELEY, 1965, p. 12).

A abordagem de Filonous leva a uma crítica do imaginário espacial prévio que estaria, segundo o senso comum, ligado a uma pretensão lógica do discurso. Sem prejuízo à lógica, de acordo com Filonous, é possível tomar as palavras só pelo que elas são literalmente, e isto ajuda a dissolver a figura ilusória de uma matéria independente da mente:

(...) não imponho o significado aos vossos termos (...). Só rogo que por meio deles me exprimais qualquer coisa que se perceba. Ao que me dizíeis, a matéria suporta os acidentes, ou acha-se por debaixo dos acidentes. De que maneira? Assim como as pernas vos suportam o corpo? (BERKELEY, 1965, p. 42).

Sobre essa questão, a seguinte afirmação de Bergson (1979, p. 60) é ilustrativa: “o que o idealismo de Berkeley significa é que a matéria é coextensiva à nossa representação; que ela não tem interior, nada sob si; que ela não esconde nada, não contém nada”.

Em determinada passagem do *Tratado sobre os princípios do conhecimento humano*, Berkeley desafia o leitor a converter, na organização da frase, o sentido da palavra “matéria” para o sentido da palavra “nada”, sem que de fato depreenda-se dessa palavra alguma ideia que possa ser completada pelo seu contrário:

É possível, se a alguém agradar, usar o termo “matéria” no mesmo sentido em que os homens usam o termo “nada”, o que neste estilo torna esses dois termos convertíveis. (...) Replicar-se-á talvez que da definição se inclui a distinção suficiente de “nada”: a ideia abstrata positiva de quiddidade, entidade ou existência. Os que pensam ter a faculdade de formar ideias abstratas gerais falam como se tivessem essa ideia, para mim a mais incompreensível (BERKELEY, 1984, p. 28).

Assim, a forma do pensamento do filósofo já tem uma qualidade reconhecível sob o ponto de vista do vulgo, que, no entanto, não a discerne, a despeito deste reconhecimento, mais elucidativa do que o pensamento comum. Isto acontece porque, aos olhos do vulgo, a palavra do filósofo ainda reveste apenas uma espécie de irônico estereótipo: o discurso do filósofo é aquele que contém “conceitos fantásticos”, “mistério”, “enigma”, ainda que, aos olhos do filósofo, o amor ao enigma seria a própria caracterização do vulgo.

Da relação entre filosofia e linguagem

O enigma filosófico a ser elucidado, proposto por Hílas a Filonous, é que este explique a afirmação de que “a matéria não tem existência”. A explicação começa a ser dada a partir de uma resposta à acusação de ceticismo feita por Hílas a Filonous. Este quer, por introdução persuasiva, mostrar que a afirmação contrária à acusação não será tão absurda quanto esta. O artifício de Filonous é ter a pretensão de revestir a forma de linguagem do senso comum, mas praticar a filosofia. E mostra a Hílas, primeiro, que classificar o sujeito que nega a existência da matéria de “cético” é inexato, pois o cético é alguém que duvida de tudo, ao passo que quem nega alguma coisa está na verdade tendo a intenção de afirmar algo: a própria negação.

Hílas rebate explicando o conteúdo da acusação recorrendo a um “querer dizer” oculto mas ao mesmo tempo evidente, como que por debaixo da sentença. Um “querer dizer” que poderia ter sido entendido por Filonous se este o quisesse. Ou seja, Hílas recorre a uma obscuridade, a um expediente psicológico no nível imaginativo, emocional, a uma ausência que, por esse ângulo, deveria ter sido pressentida numa comunicação honesta entre os participantes do diálogo. Ele explica que o que quis dizer foi que Filonous “nega a realidade e a verdade das coisas”. Tais palavras se organizam na mentalidade de Hílas de forma que “matéria” seja equivalente, em seu significado, a “realidade” e “verdade”.

Introduz-se assim o desmascaramento dessa “profundidade obscura” dos significados propostos por Hílas como sinônimos, cujo núcleo é constituído pela palavra “maté-

ria”. Para Filonous, coisas não são coisas: são princípios e teoremas das ciências. As ciências localizam-se em algo diferente do que se chama vulgarmente “matéria”.

Filonous tenta a seguinte explicação, apelando ao entendimento do senso comum, através de um exemplo da mais simples apreensão da matéria pelos sentidos significando um acesso a algo “não-material”: por exemplo, o significado de uma palavra escrita, distanciando-se do sentido da visão presentificado no ato de ler a palavra. Aqui há o exemplo de como o sentido da visão não está sendo utilizado para dar acesso a nenhuma percepção “material”, com significado em si, mas que está acessando algo que não se encontra realmente diante dos olhos. E nem por isso se estará lidando com uma necessária gradação entre realidade e não realidade.

Essa maneira de abordar a questão da diferença de nível perceptivo entre significante e significado localiza-se historicamente, segundo Foucault (1967), justamente na transição filosófica de uma concepção ternária do signo (forma, conteúdo e similitude), predominante desde o “estoicismo” até a Renascença, para uma concepção binária (significante e significado) do signo, que irá predominar a partir do século XVII:

(...) a partir do século XVII perguntar-se-á como é que um signo pode estar ligado ao que significa. (...) A profunda interdependência da linguagem e do mundo acha-se desfeita. O primado da escrita é suspenso. Desaparece então essa camada uniforme em que se entrecruzavam indefinidamente o *visto* e o *lido*, o visível e o enunciável. As coisas e as palavras vão separar-se. O olho será destinado a ver, e a ver apenas; o ouvido, apenas a ouvir. O discurso terá então por objetivo dizer o que é, mas já não será coisa alguma do que diz (FOUCAULT, 1967, p. 68).

Façamos um breve paralelo dessa transição da concepção ternária para a concepção binária no tema filosófico do signo, apontada por Foucault (1967), com a transição de uma mentalidade aristotélica anteriormente predominante, em que a fundamentação do ser-mundo está na tríade espírito-forma-matéria, para a concepção binária espírito-matéria, que extingue a questão da forma e irá desenvolver toda uma certa linha filosófica posterior que substitui a forma pela representação, sem que esta tenha o mesmo estatuto de realidade daquela.

Talvez possamos enxergar o reflexo da correspondência do paralelo mencionado quando curiosamente vemos em Berkeley duas visões a respeito do tema da linguagem, que parecem, ao mesmo tempo, a) atribuir uma importância radical ao conceito de linguagem num caso, o filosófico, em que persiste a certeza do sentido correto, mesmo num

signo desprovido da instância de similitude e b) atribuir uma importância relativa à linguagem, noutro caso, o senso comum.

O primeiro caso se dá da seguinte forma: Berkeley vê na conexão entre as ideias, durante o processo da percepção e pensamentos que se desenrolam pela mente, a expressão do próprio “plano divino” em que se constitui a realidade. E ao invés de ver relações de causa e efeito entre as ideias, Berkeley vê relações de “sinal” e “significado”:

(...) A conexão das ideias não implica a relação de causa e efeito mas somente a de um *sinal* da coisa *significada*. O fogo que vejo não é causa da dor sentida se me aproximar mas o sinal para me acautelar dele. O ruído que ouço não é efeito de movimento ou colisão de corpos externos mas o sinal disso (BERKELEY, 1984, p. 26).

Essa concepção do “sinal”, ao invés da causa, parece, em primeiro lugar, ser necessária à coesão do discurso filosófico, que pretende demolir a concepção de matéria para fazer ver em seu lugar um espírito que se manifesta em infinita variedade de formas diretamente acessíveis aos sentidos. E, em segundo lugar, mostra como o sinal, sendo algo que vem depois da coisa significada, não poderá logicamente ser causa, mas, sim, consequência. Há uma anterioridade lógica localizada no espírito ou ser percipiente que se reflete na anterioridade da significação em relação ao esquema da representação:

O que se tem dito da existência absoluta de coisas impensáveis sem alguma relação com o seu ser-percebidas parece perfeitamente ininteligível. O seu *esse* é *percipi*; nem é possível terem existência fora dos espíritos ou coisas pensantes que os percebem (BERKELEY, 1984, p. 13).

Aqui arriscamo-nos a imaginar o mundo de Berkeley (1984, p. 26) como que puramente expressivo, uma espécie de linguagem de Deus, no qual “o esforço do filósofo natural” deve ser

procurar entender esses sinais instituídos pelo Autor da Natureza (...) e não o pretender explicar coisas por causas corpóreas, doutrina que parece ter afastado a mente humana do seu princípio ativo.

Essa questão foi apontada por Bergson (1979) da seguinte maneira: a mensagem mais eficiente e duradoura que o filósofo pode comunicar reside na imagem e não na série de argumentos, explicações, teses, ou análises de que sua obra pode consistir. Em Berkeley, haveria duas imagens unindo as concepções de substância espiritual, ideia e Deus.

A primeira é a de que “Berkeley percebe a matéria como uma fina película transparente situada entre o homem e Deus”, e a segunda é a de que “a matéria seria uma língua em que Deus nos fala” (BERGSON, 1979, p. 62).

Por sua vez, a implicação da ideia de “consequência” no conceito de sinal se imiscui na concepção de linguagem como estrutura utilitária para fins de comunicação. É nesse caso, quando o conceito de linguagem se afunila à modalidade da linguagem verbal, que aparece toda a possibilidade de inexatidão, irrealidade e senso comum inerentes à natureza dessa mesma linguagem. Vê-se aqui o entrelaçamento entre os temas da própria constituição do conhecimento dentro de uma noção de ser e da linguagem.

Embora a pessoa que concorde com a doutrina de Berkeley possa saber que a única causa de todos os fenômenos é o espírito, nem por isso exigirá que as expressões linguísticas expressem a exatidão dessa verdade.

Com tais princípios não devemos continuar a dizer que o fogo aquece e a água refresca, mas que o espírito aquece e assim por diante. Não seria justamente ridicularizado quem assim falasse? Respondo que sim; nessas coisas devemos *pensar com os doutos, falar como o vulgo* (BERKELEY, 1984, p. 23).

A questão do contexto é uma questão de linguagem

Então, a linguagem verbal sozinha, não sendo suficiente para expressar ideias precisas, dependerá do contexto para ser empregada e compreendida corretamente: “um leitor sério e simples buscará o sentido pelo fim e teor da conexão do discurso” (BERKELEY, 1984, p. 23). Relacionando a linguagem verbal com a questão da abstração, Berkeley sugere que a palavra é na verdade uma espécie de instrumento que, por conta de sua múltipla utilidade, se apresenta sob uma aparência universal, mas que seu significado sempre será uma ideia particular. Sob sua filosofia não se deverá fazer obstáculos das palavras. É possível dissociar palavra e ideia e acessar esta última a despeito da primeira. O conhecimento verdadeiro, para Berkeley, transcende as palavras quando não se deixa turvar por elas:

Quem sabe que só tem ideias particulares não se esforçará em vão por achar e conceber a ideia abstrata ligada a qualquer nome; e quem sabe que os nomes nem sempre significam ideias poupa-se ao trabalho de perseguir ideias onde nada há. Bom seria que todos se esforçassem por obter visão clara das ideias consideradas, separando-as da vestidura e acúmulo da palavra que muito contribuem para cegar o juízo e dividir a atenção (BERKELEY, 1984, p. 11).

Aparentemente, dir-se-ia que, dessa visão instrumentalista da linguagem, poder-se-á inferir o desenvolvimento do conceito de “paradigma”, “incomensurabilidade”, entre outros, da futura filosofia da ciência. No entanto, no nível do discurso filosófico concebido como então, Berkeley (1984, p. 23) dispensa as especialidades lexicais: “Quem aceita por demonstração o sistema de Copérnico nem por isso deixa de dizer ‘o sol nasce’, ‘o sol põe-se’ ou ‘chega ao meridiano’; e decerto seria muito ridículo afetar um estilo contrário”.

A linguagem fundará a noção de contexto na filosofia da ciência do século XX, em que Kuhn (2003), por exemplo, diz que a linguagem na ciência não deve ser apenas “clara” ou “precisa”, ela deve ser específica e se diferenciar, principalmente no tocante à classe gramatical dos “nomes”, de outras. O léxico especializado é o fundamento dos limites entre as ramificações das especialidades científicas. Aqui o contexto não se dissocia de um léxico específico. A combinação entre contexto e léxico forma o que Kuhn (2003, p. 116) chama de “incomensurabilidade”: “componente essencial de qualquer concepção histórica, desenvolvimentista ou evolucionária do conhecimento científico”.

A sutil diferença entre as noções de contexto, conforme sejam filiadas principalmente à linguagem ou em sentido inverso principalmente à percepção, delineia a diferença de rumos entre as consequências da contenda Berkeley-Locke. Uma certa linha representando Locke delineia uma realidade que parte da instância abstrata interna do ser para a concretude exterior da matéria e, nesse movimento do conhecimento, as palavras são os títulos de uma referencialidade construída e, portanto, contígua à incomensurabilidade e aí é necessária a possibilidade de tradução, que implica a abstração. Já uma linha berkeleyana ligará o ser diretamente à “matéria”, e a palavra do conhecimento será sempre uma clara expressão da experiência num dado contexto concretizado. Nesse caso, não existirá a questão da tradução porque cada experiência será única, singular e objetiva.

Uma tradução no modelo de Kuhn (2003), por exemplo, implica o deslocamento de um mesmo problema entre especialidades científicas diferentes. Assim, algo entendido como objeto material, na acepção convencional, e, ao mesmo tempo abstrato, se configurando em projetos e discursos científicos através de uma aura de permanência, atravessando contextos diferentes, não faria sentido numa filosofia da ciência nos moldes de Berkeley. Se há algo que se aproxime de uma tal aura de permanência em Berkeley, será com certeza o espírito percipiente, que se deslocará pelo mundo talvez como um Gulliver em suas viagens.

Referências

BERKELEY, George. **Tratado sobre os princípios do conhecimento humano**. Trad. Antônio Sérgio. São Paulo: Abril Cultural, 1984 (Os pensadores).

BERKELEY, Jorge. **Três diálogos entre Hilas e Filonous em oposição aos céticos e ateus**. Trad. de Antônio Sérgio. Coimbra: Atlântida, 1965.

BERKELEY, George. **Tratados sobre a visão**. Trad. José Oscar de Almeida Marques. Edição em inglês e português. Campinas: Editora da Unicamp, 2010.

BERGSON, Henri. **Cartas, conferências e outros escritos**. Trad. Franklin Leopoldo Silva, Nathanael Caxeiro. São Paulo: Abril Cultural, 1979.

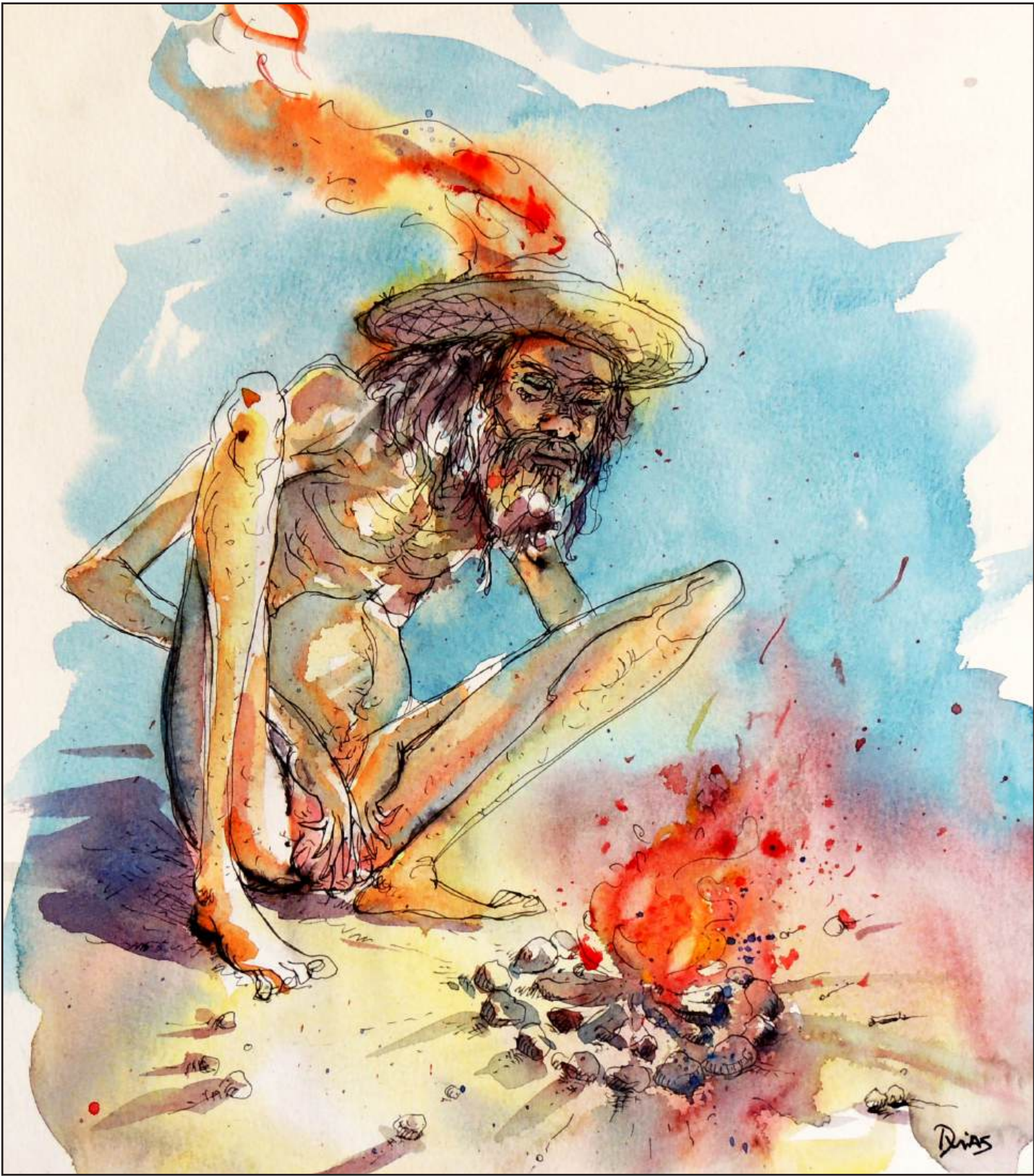
FOUCAULT, Michel. **As palavras e as coisas**. São Paulo: Martins Fontes, 1967.

KUHN, Thomas. **O caminho desde a estrutura**. Trad.: Cesar Mortari. São Paulo: Unesp, 2003.

LACEY, Hugh. **Valores e atividade científica 1**. Editora 34: São Paulo, 2008.

NAGEL, Thomas. **Visão a partir de lugar nenhum**. Trad. Silvana Vieira. São Paulo: Martins Fontes, 2004.





RESENHA DO LIVRO *NEUROCOGNITIVE MECHANISMS: EXPLAINING BIOLOGICAL COGNITION* (OXFORD UNIVERSITY PRESS, 2020), DE GUALTIERO PICCININI



Diego Azevedo Leite¹

Publicado em 2020, o livro do filósofo Ítalo-americano, Gualtiero Piccinini, intitulado *Neurocognitive mechanisms: Explaining biological cognition* (que pode ser traduzido² por “Mecanismos neurocognitivos: Explicando a cognição biológica”)³, pode ser considerado como um dos trabalhos mais representativos da extensa literatura especializada publicada recentemente sobre o neomecanicismo na filosofia da mente e da ciência cognitiva. O projeto apresentado no livro de Piccinini é bastante ambicioso em escopo, contando com discussões sobre questões metafísicas complexas e bastante controversas, sobre história da ciência cognitiva e sobre explicações científicas na neurociência cognitiva. Seu estilo de escrita é, no geral, claro e direto, e a bibliografia utilizada, embora esteja longe de ser exaustiva, dada a complexidade do tema tratado, é relevante para os propósitos do trabalho, sendo que o autor mostra ter domínio adequado sobre ela.

Neste livro, o objetivo central de Piccinini é defender sua teoria da cognição biológica (incluindo a cognição humana), que enfatiza mecanismos neurais e processos computacionais a partir de um contexto filosófico mais abrangente, que conta com, pelo menos, dois grandes pilares. O primeiro pilar é o fisicalismo (ou materialismo)⁴, o qual

1 Universidade Federal de Alfenas.

2 Todas as traduções do inglês neste trabalho são minhas.

3 Ainda não há tradução deste livro para o português.

4 Piccinini e Craver (2011, p. 284), por exemplo, afirmam o seguinte: “Nós endossamos o reducionismo no sentido de que toda coisa concreta é feita de componentes físicos e as atividades organizadas dos componentes de um sistema explicam as atividades do todo.” Outro exemplo é o que segue: “Embora nós, junto com qualquer outro participante neste debate, assumamos que sistemas funcionais são realizados por mecanismos, alguns dualistas discordam; eles mantêm que um sistema funcional pode ser um sistema implementado de forma não-física, não-mecanística. Nós desconsideramos essa possibilidade com base nas razões usuais do fechamento causal do físico e da falta de uma teoria adequada a respeito da interação entre

permanece sendo continuamente reelaborado durante este primeiro quarto do século XXI (cf. STOLJAR, 2021)⁵. O segundo pilar é o movimento neomecanicista na filosofia da ciência, o qual foi desenvolvido a partir, sobretudo, de reflexões sobre as ciências biológicas (cf. BECHTEL & RICHARDSON, 2010).

Em seus trabalhos publicados nas últimas décadas, voltados para as áreas da filosofia da neurociência, filosofia da ciência cognitiva e filosofia da mente, Piccinini busca defender uma abordagem neomecanicista e, ao mesmo tempo, computacionalista da cognição humana (e da cognição biológica, no geral), a qual ele próprio caracteriza como “revolucionária” (BOONE & PICCININI, 2016)⁶. Seu ponto de vista é centrado em pesquisas científicas e reflexões filosóficas a respeito das atividades e processos dos cérebros de organismos humanos e de outros animais, assim como em pesquisas científicas e reflexões filosóficas a respeito de processos computacionais nestes “mecanismos biológicos”. Esta visão neomecanicista, defendida também por inúmeros outros autores atualmente, ganhou grande relevância e popularidade no presente, devido também ao uso, em suas construções teóricas, de uma linguagem que constantemente se apropria dos avanços científicos e metodológicos de pesquisas empíricas na área da neurociência, bem como dos avanços tecnológicos nas áreas da ciência da computação e inteligência artificial.

Partindo desse contexto, na visão de Piccinini apresentada em seu livro, o sistema nervoso humano é um mecanismo funcional que realiza computações e processa informações por meio da ativação de neurônios. Capacidades cognitivas são explicadas, dessa forma, por mecanismos neurocognitivos de múltiplos níveis (por exemplo: nível celular e molecular; nível de pequenos conjuntos de neurônios; nível de grandes sistemas e redes neurais; nível do sistema nervoso como um todo) que realizam “computações neurais” sobre “representações neurais”. Para Piccinini, a “computação neural” (ou seja, computações definidas nos elementos funcionalmente relevantes da atividade neural) não deve ser caracterizada como puramente digital, como ela era entendida classicamente (NEWELL & SIMON, 1961). Mas ela também não deve ser compreendida de uma forma puramente analógica, como era entendida pela abordagem conexionista de redes neurais (SMOLENSKY, 1988). De acordo com o autor, a computação neural é “*sui generis*”; isto é, nem totalmente digital, nem totalmente analógica (p. 1, 144, 155, 207, 219, 311).

propriedades físicas e não-físicas” (PICCININI & CRAVER, 2011, p. 289).

5 A verdade ou falsidade de posições fisicalistas, no entanto, continua sendo tema fortemente debatido nas discussões recentes.

6 Apesar de se referirem a uma “revolução” da neurociência cognitiva, os autores do trabalho não esclarecem o que eles entendem por este complexo conceito.

Estrutura e conteúdo do livro

Com a finalidade de defender sua teoria neuro-computacional da cognição biológica, o autor organiza seu trabalho da seguinte forma. O livro, além da introdução, conta com uma divisão do seu conteúdo em 14 capítulos, nos quais são discutidas questões relevantes acerca de temas bastante complexos e controversos, tais como: níveis ontológicos de investigação; realização múltipla; natureza e caracterização de mecanismos biológicos; história da teoria computacional da cognição; natureza e caracterização de processos computacionais; representações neurais; relação entre computação e consciência; dentre outros. É com base nesta estrutura e neste conteúdo que Piccinini alcança suas conclusões centrais, que aparecem, em um breve resumo, no capítulo final da obra.

Análise do livro

Há uma série de pontos obscuros, argumentos problemáticos, lacunas e outras dificuldades e limitações na literatura especializada sobre o neomecanicismo enquanto teoria aplicada à cognição humana e à ciência cognitiva (cf. LEITE, 2021). O livro de Piccinini não é uma exceção. Eu vou me limitar, porém, a mencionar dois problemas com esta visão, os quais eu considero bastante substanciais.

O problema com a noção de computação neurocognitiva

O que é, afinal, computação cognitiva? Qual é o tipo de processo computacional realizado pelo cérebro humano ou pelos processos cognitivos tomados em abstrato que explicam as capacidades humanas de exercer a liberdade da vontade, de utilizar sua consciência e autoconsciência, de estabelecer normas morais e de utilizar seu raciocínio informal em diversas dimensões da vida? É mesmo possível explicar o fenômeno da cognição humana completamente em termos de computação neurocognitiva? Essas questões são, atualmente, extremamente controversas. Além disso, historicamente, a ideia de computação aplicada à cognição humana sofreu críticas substanciais (e.g., BRUNER, 1990, 1996; DREYFUS, 1972; FODOR, 2000, 2005, 2008; PUTNAM, 1988; SEARLE, 1980). Ironicamente, muitas dessas críticas foram elaboradas por autores centrais no desenvolvimento do próprio movimento cognitivista.

Na visão de Piccinini, um sistema computacional físico é um mecanismo funcional que tem a função teleológica de manipular variáveis (ou veículos), as quais são independentes do meio concreto de implementação, de acordo com uma regra definida sobre estas variáveis. O autor desenvolve, assim, uma concepção de processos computacionais em termos sintáticos, formais e biologicamente construídos. Um dos problemas mais graves com essa abordagem é que ela não é sensível à dimensão semântica do conteúdo cognitivo. A dificuldade que surge a partir disso, na visão de muitos outros autores, é que propriedades semânticas desempenham um papel essencial, na ciência cognitiva, para a explicação de inúmeros comportamentos humanos que envolvem cognição (cf. e.g., RESCORLA, 2012)⁷.

A dimensão semântica (ou do significado) do conteúdo de representações mentais está relacionada às propriedades de elas serem originais (ou não-derivadas), verdadeiras ou falsas, precisas ou imprecisas, e se referirem a algo que não seja elas mesmas (isto é, a propriedade de ser sobre algo diferente de si mesmo). É dessa forma que o ser humano pode ter representações sobre, por exemplo, objetos que se encontram no seu ambiente físico (como uma bola de futebol), mesmo quando não está diretamente exposto ao objeto.

O problema fundamental com teorias computacionais não-semânticas é que elas explicam propriedades apenas não-semânticas das representações mentais, enquanto as propriedades semânticas permanecem sem uma explicação. A teoria de Piccinini pode explicar qual é o mecanismo neuro-computacional responsável pelo processamento de informação a respeito da forma circular e do movimento de uma bola de futebol no campo visual. Contudo, a teoria não diz nada sobre o significado que aquele objeto pode assumir na dinâmica cognitiva de uma pessoa e da relação entre este significado e alguns dos comportamentos realizados por ela.

⁷ Outra linha de discussão se refere à questão da *individuação* de sistemas computacionais físicos. Muitos autores pensam que propriedades semânticas são essenciais para uma teoria sobre quais sistemas físicos são realmente capazes de executar computações (cf. SHAGRIR, 2020). Ou seja, é através do uso de propriedades semânticas que se pode dizer, por exemplo, que um cérebro realiza computações de uma forma, que um computador digital realiza computações de outra e que uma pedra ou um certo tipo de relógio não realiza computações. Dito de outra forma, é apenas através de propriedades semânticas que se pode diferenciar um cérebro humano de outros sistemas computacionais físicos, como um PC desktop, um notebook, um tablet ou um smartphone.

O problema da unificação

A obra de Piccinini busca desenvolver uma estrutura teórica para construir uma ciência unificada da cognição (cf. PICCININI & CRAVER, 2011). Essa unificação, na sua visão, é alcançada ao se mostrar como análises funcionais de capacidades cognitivas devem ser integradas com explicações mecanicistas multiníveis de sistemas neurais. Boone e Piccinini (2016, p. 1530), igualmente, escrevem sobre a integração mecanicista da psicologia e neurociência, formando assim a revolucionária neurociência cognitiva, “uma ciência da cognição completamente integrada”.

Este é um tópico, de fato, bastante significativo nas ciências, uma vez que a posição adotada para buscar unidade em uma determinada área científica pode fornecer instrução e justificação para a formulação de hipóteses, projetos e objetivos específicos (CAT, 2017, sect. 6). Além disso, essa posição, normalmente, carrega a autoridade para apresentar normas a serem seguidas e dizer o que é ou não legítimo no campo de estudo. Além de influenciar nos tipos de questões a serem respondidas e quais áreas específicas devem ser mais exploradas, as posições adotadas para unificação podem ter a autoridade para dizer o que é realmente científico, influenciando assim a construção de programas e projetos científicos e filosóficos.

No entanto, unificar a ciência cognitiva por meio de uma teoria abrangente já é por si só uma tarefa extremamente árdua. Muitos autores ao longo da história apresentaram propostas de unificação e integração interna, mas o que se observa ainda hoje é uma fragmentação e diversidade teórica acentuadas. O projeto de Piccinini não parece fornecer uma solução para a questão da enorme fragmentação e diversidade encontradas na ciência cognitiva de forma substancial⁸. Isso ocorre porque a proposta de Piccinini apresenta teses extremamente controversas até mesmo entre os próprios defensores do neomecanicismo na ciência cognitiva.

Além disso, unificar ou integrar externamente a ciência cognitiva e a neurociência é outro projeto que acrescenta mais um grau de dificuldade ao cenário. Piccinini defende que explicações científicas na neurociência cognitiva (que seria a ciência da cognição uni-

⁸ É muito importante enfatizar também que a cooperação entre cientistas de diferentes disciplinas não implica necessariamente a aceitação ou o desenvolvimento de uma estrutura teórica unificada ou integrada de forma coerente e sistemática. Um trabalho científico multidisciplinar integrado é uma coisa, e o desenvolvimento de uma teoria integrada ou unificada com coerência, bem articulada, coesão interna e externa e sistematicidade é outra. Simplesmente incentivar a cooperação de cientistas trabalhando em campos diferentes não é o mesmo que gerar integração teórica coerente em uma área científica.

ficada, na sua visão) deveriam substituir explicações na psicologia ou ciência cognitiva, como explicações legítimas da cognição humana. De acordo com esta visão, as explicações psicológicas não são autônomas no que diz respeito às explicações mecanicistas neuro-científicas. Explicações psicológicas válidas, nessa visão, devem ser construídas através da investigação da função ou capacidade psicológica específica e da estrutura neural que é responsável pela sua ocorrência. Uma explicação que não indique esta estrutura estaria incompleta e seria inapropriada.

Contudo, os exemplos de fenômenos psicológicos aos quais Piccinini aplica sua teoria são muito limitados e são retirados, principalmente, de pesquisas particulares da neurociência. Não há discussões aprofundadas a respeito de temas tradicionalmente centrais na discussão filosófica que atravessam a ciência psicológica. Alguns exemplos são a questão da liberdade da vontade humana, discussões sobre moralidade, discussões sobre consciência e autoconsciência e discussões sobre a racionalidade humana. Portanto, existem ainda fenômenos complexos relacionados à cognição humana que parecem não admitir um tratamento em termos neuro-computacionais mecanicistas.

Além disso, pode-se argumentar que explicações psicológicas bem-sucedidas, ao contrário, têm uma autonomia robusta em relação a explicações neuro-computacionais mecanicistas. Considere, por exemplo, o caso de uma pessoa que precisa tomar uma decisão sobre em quem votar para presidente do seu país e busca refletir seriamente sobre o seu voto (cf. LEITE, 2021, cap. 6; LEITE, 2022). Considere também que existem três grandes escolhas possíveis, entre os presidentes A, B e C. Baseando-se em reflexões simples sobre seu conjunto de crenças, emoções, motivações, valores e princípios e refletindo conscientemente talvez por alguns minutos ou algumas horas, uma pessoa consegue facilmente descartar o candidato C. Porém, a escolha entre os candidatos A e B não é tão simples. Essa escolha envolve a reflexão sobre argumentos apresentados por inúmeras pessoas, sejam familiares, jornalistas, analistas políticos, ou mesmo pelos próprios candidatos. Essa escolha racional com base na análise de argumentos informais pode demorar meses para ser realizada.

Além disso, ela pode envolver não apenas a reflexão sobre o próprio conjunto de crenças, emoções, motivações, intenções, valores e princípios e os argumentos apresentados em favor de cada candidato, mas também uma reflexão sobre o que a pessoa desejaria para o seu próprio futuro, pensando no seu próprio desenvolvimento humano e no da sociedade em que vive. Quer dizer, a capacidade de autoconsciência, a influência social no sistema de crenças e a capacidade de pensar no bem da própria sociedade também poderão desempenhar um papel essencial na decisão final.

Quando refletimos, no entanto, sobre os processos computacionais de sistemas computacionais físicos e abstratos atuais, percebemos que este conjunto de capacidades não está presente neles. Primeiro, não há neles um conjunto de crenças desenvolvidas a partir da própria reflexão e da capacidade de representações mentais originais que se relacionam de forma lógica e podem ser verdadeiras ou falsas, precisas ou imprecisas. Segundo, não há a interação dinâmica e complexa entre crenças, emoções, motivações, valores, princípios, consciência e autoconsciência. Terceiro, não há influência social e cultural no desenvolvimento cognitivo de sistemas computacionais em todas as dimensões em que há no desenvolvimento cognitivo humano. Em resumo, a cognição humana e os sistemas computacionais apresentam propriedades bastante diferentes, e isso também pode ser observado nos diferentes comportamentos realizados por esses diferentes tipos de entidades.

O ponto de vista defendido por Piccinini, portanto, é apenas um em meio a uma grande diversidade de opiniões. Conseqüentemente, ele não pode ser utilizado como ponto de integração ou unificação nem mesmo para os próprios proponentes do neomecanicismo na ciência cognitiva atual, menos ainda para a ciência cognitiva como um todo.

Conclusão

Por conta das limitações que foram expostas, o livro não apresenta força argumentativa suficiente para sustentar as suas conclusões mais importantes. Entretanto, ele apresenta muitas considerações teóricas que podem contribuir para a diversidade de propostas e para o amplo debate. Sendo assim, ele pode ser indicado para os interessados em discussões contemporâneas a respeito da cognição humana, especialmente nas áreas de filosofia da mente, de filosofia da ciência cognitiva e de ciência cognitiva teórica.

Referências

BECHTEL, W. & RICHARDSON, R. **Discovering complexity**: Decomposition and localization as strategies in scientific research. Cambridge, MA: MIT Press, 2010 (Publicado originalmente em 1993).

BOONE, W. & PICCININI, G. Cognitive neuroscience revolution. **Synthese**, v. 193, p. 1509-34, 2016. DOI: <https://doi.org/10.1007/s11229-015-0783-4>.

- BRUNER, J. **Acts of meaning**. Cambridge, MA: Harvard University Press, 1990.
- BRUNER, J. **The culture of education**. Cambridge, MA: Harvard University Press, 1996.
- CAT, J. The unity of science. **The Stanford encyclopedia of philosophy**, 2017. Disponível em: <https://plato.stanford.edu/archives/fall2017/entries/scientific-unity/>.
- DREYFUS, H. L. **What computers can't do: A critique of artificial reason**. New York: Harper & Row, Publishers, 1972.
- FODOR, J. **LOT 2: The language of thought revisited**. Oxford University Press, 2008.
- FODOR, J. Reply to Steven Pinker "So how does the mind work?" **Mind & Language**, v. 20, n. 1, p. 25-32, 2005. DOI: <https://doi.org/10.1111/j.0268-1064.2005.00274.x>
- FODOR, J. **The mind doesn't work that way: The scope and limits of computational psychology**. Cambridge, MA: MIT Press, 2000.
- LEITE, D. A. **The twenty-first century mechanistic theory of human cognition: A critical analysis**. Cham: Springer, 2021.
- LEITE, D. A. The neo-mechanistic model of human cognitive computation and its major challenges. In: Habib, M. K. (Ed.). **Cognitive robotics and adaptative behaviors**. London: IntechOpen, 2022, p. 1-20. DOI: 10.5772/intechopen.104995
- NEWELL, A. & SIMON, H. A. Computer simulation of human thinking. **Science**, v. 134, n. 3495, p. 2011-7, 1961.
- PICCININI, G. & CRAVER, C. Integrating psychology and neuroscience: Functional analyses as mechanism sketches. **Synthese**, v. 183, n. 3, p. 283-311, 2011.
- PUTNAM, H. **Representation and reality**. Cambridge, MA: MIT Press, 1988.
- RESCORLA, M. Are computational transitions sensitive to semantics? **Australasian Journal of Philosophy**, v. 90, n. 4, 703-21, 2012. DOI: <https://doi.org/10.1080/00048402.2011.615333>
- SEARLE, J. R. Minds, brains, and programs. **The Behavioral and Brain Sciences**, v. 3, p. 417-57, 1980.
- SHAGRIR, O. In defense of the semantic view of computation. **Synthese**, v. 197, p. 4083-108, 2020. <https://doi.org/10.1007/s11229-018-01921-z>
- SMOLENSKY, P. On the proper treatment of connectionism. **Behavioral and Brain Sciences**, v. 11, p. 1-74, 1988.

STOLJAR, D. Physicalism. **The Stanford encyclopedia of philosophy**, 2021. Disponível em: <https://plato.stanford.edu/archives/sum2022/entries/physicalism/>





PALESTRA 1 DO LIVRO *EMERGENT EVOLUTION* (THE GIFFORD LECTURES): EMERGÊNCIA



C. Lloyd Morgan

Tradutora: Cristiane Xerez Barroso

I. Emergentes e Resultantes. II. O esquema piramidal. III. Involução e Dependência. IV. Em direção ao Espaço-Tempo. V. Deidade.

§ I. Emergentes e Resultantes

Nós vivemos em um mundo em que parece haver uma sequência ordenada de eventos. É função da ciência, e de uma filosofia que se mantém em contato com a ciência, descrever o curso dos eventos neste ou naquele momento de sua ocorrência e descobrir o plano segundo o qual eles prosseguem. Evolução, no sentido amplo da palavra, é o nome que damos ao abrangente plano de sequência de todos os eventos naturais.

Mas a sequência ordenada, vista historicamente, parece apresentar, de tempos em tempos, algo genuinamente novo. No que aqui denomino de evolução emergente, a ênfase é colocada na chegada do novo. Exemplos relevantes são oferecidos no advento da vida, no advento da mente e no advento do pensamento reflexivo. Mas, no mundo físico, a emergência não é menos exemplificada pelo advento de cada novo tipo de átomo e de cada novo tipo de molécula. Está além da capacidade do homem enumerar os casos de emergência. Mas se nada de novo surgir – se houver apenas um reagrupamento de eventos pré-existentes *e nada mais* –, então não há evolução emergente.

A afirmação naturalista é que, com base nas evidências, não apenas átomos e moléculas, mas organismos e mentes, são suscetíveis de tratamento por métodos científicos

fundamentalmente do mesmo tipo; que todos pertencem a um mesmo tecido de eventos; e tudo isso exemplifica um plano fundamental. Em outras palavras, a posição é que, em uma filosofia baseada no procedimento sancionado pelo progresso da pesquisa e do pensamento científicos, o advento de qualquer tipo de novidade deve ser lealmente aceito onde quer que seja encontrado, sem invocar qualquer Poder extranatural (Força, Enteléquia, *Élan* ou Deus) através da Atividade eficiente da qual os fatos observados podem ser explicados. Surge então a questão de saber se tal interpretação científica ou naturalista é suficiente, ou se alguma explicação supranaturalista adicional é admissível no tribunal da filosofia, não como substituta, mas como complemento do resultado da investigação científica. Eu afirmarei que isso é admissível e que não há nada na evolução emergente, que se pretende estritamente naturalista, que impeça um reconhecimento de Deus. Isto implica (1) que uma filosofia construtiva é mais do que ciência, e (2) que tal reconhecimento deve aqui ser fundado apenas em considerações filosóficas.

O conceito de emergência foi tratado (para não retroceder ainda mais) por J. S. Mill, em seu *Logic* (Bk. III. cap. vi. §2), sob a discussão de “leis heteropáticas” na causalidade. A palavra “emergente”, em contraste com “resultante”, foi sugerida por G. H. Lewes em seu *Problems of Life and Mind* (Vol. II. Prob. V. ch. iii. p. 412). Ambos apresentam exemplos da química e da fisiologia; ambos tratam de propriedades; ambos distinguem aquelas propriedades (a) que são apenas aditivas e subtrativas e previsíveis, daquelas (b) que são novas e imprevisíveis; ambos insistem na afirmação de que estas últimas, tanto quanto as primeiras, se enquadram na rubrica da causalidade uniforme. Um exemplo simples e familiar deve ser suficiente. Quando o carbono com certas propriedades se combina com o enxofre com outras propriedades, forma-se não uma mera mistura, mas um novo composto, algumas das propriedades são bastante diferentes daquelas de qualquer um dos componentes. Já o peso do composto é uma resultante aditiva, a soma dos pesos dos componentes; e isso poderia ser previsto antes que qualquer molécula de dissulfeto de carbono fosse formada. Nós poderíamos dizer antecipadamente que, se o carbono e o enxofre se combinarem em quaisquer proporções verificáveis, o resultado será tal e tal peso. Mas diversas outras propriedades são emergentes constitutivas que (afirma-se) não poderiam ser previstas antecipadamente em qualquer caso de tal combinação. É claro que quando se aprende o que emerge *neste* caso particular, pode-se prever o que surgirá *naquele* caso sob circunstâncias semelhantes. Nós aprendemos algo sobre o plano natural da evolução emergente.

Tal emergência do novo é agora amplamente aceita no que diz respeito à vida e à mente. É uma doutrina incansavelmente defendida pelo Professor Bergson. Wundt

pressionou a sua aceitação sob o seu “princípio das resultantes criativas” (isto é, o que distinguimos como emergentes) que, diz ele, “tenta afirmar o fato de que, em todas as combinações psíquicas, o produto não é uma mera soma dos elementos separados... mas representa uma nova criação”. (I.P. p. 164). Browning, em *Abt Vogler*, enfatizou-a poeticamente em referência à nossa apreciação de um acorde musical.

E não sei se, exceto nisso, tal dádiva será permitida ao homem

Que de três sons ele enquadra, não um quarto som, mas uma estrela¹.

Ao chamar “estrela”, ele dá ênfase poética ao caráter emergente da “harmonia”, que é algo mais do que a resultante aditiva dos tons constituintes – algo genuinamente novo. Se for dado em, ou para, nossa audição, tudo o que podemos dizer é: “Considere e incline a cabeça”. Essa, em certo sentido, deveria ser a nossa atitude leal para com todos os emergentes. Como diz o professor Alexander, devemos aceitá-los, todos e cada um, “com piedade natural”.

O professor M'Dougall distingue analiticamente o que se pode denominar de notas constituintes no acorde de *reverência*. Existe um elemento de terna emoção ou amor, de medo, adequadamente definido, de admiração; há uma atitude de consideração ascendente em relação a alguém que está em um nível superior; e assim por diante. Estas e outras semelhantes são as notas aditivas que são resumidas em reverência. Mas não há também algo mais; algo que confere ao resultado aditivo seu caráter distintivo de reverência; algo do qual podemos dizer: “Considere e incline a cabeça”? Se assim for, aquilo que dá à combinação destas diversas notas o seu caráter de acorde é, na nossa interpretação, uma qualidade emergente.

Browning, note-se, não nega a soma das notas constituintes do acorde; ele afirma que há mais no acorde do que pode ser interpretado apenas como o resultado da soma. Os caracteres aditivos, como resultantes, podem ser – aceitarei a hipótese de que sempre são – coexistentes com caracteres constitutivos, como emergentes. Muitas vezes, pode haver resultantes sem emergência; mas não há emergências que não envolvam também efei-

¹ “And I know not if, save in this, such gift be allowed to man
That out of three sounds he frame, not a fourth sound, but a star”.

tos resultantes. Resultantes dão continuidade quantitativa que fundamenta novos passos constitutivos na emergência. E o passo emergente, embora possa parecer mais ou menos saltitante, é melhor considerado como uma mudança qualitativa de direção, ou um ponto de virada crítico, no curso dos eventos. Nesse sentido, não há a ruptura descontínua de uma lacuna ou de um hiato. Pode-se dizer, então, que através das resultantes há continuidade no progresso; através da emergência há progresso na continuidade.

Lewes diz que a natureza dos caracteres emergentes só pode ser aprendida pela experiência da sua ocorrência; portanto, eles são imprevisíveis antes do evento. Mas, pode-se afirmar, que isto é verdade para todos os caracteres, sejam eles resultantes ou emergentes. Somente como resultado da experiência, eles podem ser preditos. Isso, em certo sentido, é assim. O ponto de ênfase, no entanto, é este. Suponhamos que haja três níveis sucessivos de eventos naturais, A, B e C. Que haja em B um *tipo de relação* que não está presente em A; e em C, um tipo de relação ainda não presente em B ou em A. Se então alguém vivesse e ganhasse experiência no nível B, não poderia prever os caracteres emergentes do nível C, porque as relações, das quais eles são a expressão, ainda não existem. Nem se vivêssemos no nível A, nós poderíamos prever o caráter emergente dos eventos-*b*, porque, *ex hypothesi, tais eventos não* existem ainda. O que, afirma-se, não se pode prever então é a expressão emergente de algum novo tipo de relacionamento entre eventos preexistentes. Não se poderia prever o caráter emergente dos eventos vitais apenas a partir do conhecimento mais completo possível dos eventos físico-químicos, se a vida for um acorde emergente e não meramente uma soma, por mais complexa que fosse, das notas-*a* constituintes. Essa é a hipótese aceita na evolução emergente.

Não se nega nem se ignora a evidência de que alguns caracteres aditivos ou resultantes são, por assim dizer, discretamente incrementais. Nem se pode negar que, somente através da experiência, se pode aprender a ordem incremental. Não parece improvável que os denominados elementos difiram pela adição sucessiva de um elétron. Até oito deles podem ser representados como formando um elétron planetário interno, ou um conjunto de elétrons, girando em torno de um núcleo solar. Outras adições são em uma esfera orbital mais ampla, novamente até oito. Além disso, temos um terceiro e ainda mais amplo curso orbital dos elétrons adicionados; e assim por diante. Mas parece também que existem certos caracteres constitutivos ou qualitativos que distinguem instâncias de +1, +2, +3, . . . incrementos em órbitas sucessivas. Eles possuem certas características em comum e formam grupos familiares. Pode-se dizer que, em cada grupo familiar, não existe somente uma resultante incremental, mas também um tipo específico de relacionamento integral no qual os caracteres constitutivos de cada membro do grupo são uma expressão emergente? Se assim for, temos aqui um exemplo do que se entende por evolução emergente.

Em um campo diferente da investigação científica, muito tem sido feito ultimamente para tornar provável a continuidade resultante entre o não-vivo e o vivo. Nenhum evolucionista provavelmente subestimar­á o seu valor. Mas ainda podemos perguntar se não existe, em algum estágio deste processo, um novo caráter emergente de vida, cuja superveniência deve ser aceita com piedade natural e descrita em termos adequados de integração vital ou não. Parece haver algo genuinamente novo em algum estágio da continuidade resultante.

E se continuarmos a história, com o discurso presidencial do Dr. E. J. Allen (Brit. Assoc. Sec. D. 1922), sobre *The Progression of Life in the Sea*, como nosso guia, embora a ênfase esteja talvez na continuidade resultante, ele pergunta repetidamente se não há emergência também.

Há mais uma questão preliminar sobre a qual algumas palavras devem ser ditas. É bastante certo que a interpretação da natureza que apresento será, em alguns setores, caracterizada como mecânica e viciada por uma aceitação acrítica do que, às vezes, é denominado de “o dogma mecanicista”. O estranho aqui é que toda a doutrina da emergência é um protesto contínuo contra a interpretação mecânica, e a própria antítese daquela que é mecanicista. Não interpreta a vida em termos de física e química. Não interpreta a mente em termos de padrões de receptores e rotas neuronais. Aqueles que supõem que isso acontece entendem totalmente mal o seu significado.

Deve-se, contudo, caracterizar de alguma forma o que deve ser considerado aqui como a nota-chave do mecanismo. Eu deveria caracterizá-la assim: a característica essencial de uma interpretação mecânica – ou, se preferir, mecanicista – é que ela é feita apenas em termos de efeitos resultantes, calculáveis por soma algébrica. Ela ignora algo mais que deve ser aceito como emergente. Ela considera um composto químico apenas como uma mistura mecânica mais complexa, sem qualquer novo tipo de relação entre os seus constituintes. Ela considera a vida como um reagrupamento de eventos físico-químicos sem qualquer novo tipo de relação expressa em uma integração que parece, segundo as evidências, marcar um novo ponto de partida na passagem dos eventos naturais. Contra tal interpretação mecânica – tal dogma mecanicista – a evolução emergente levanta-se em protesto. A essência da sua afirmação é que tal interpretação é bastante inadequada. Resultantes existem; mas também há emergência. Sob tratamento naturalista, porém, a emergência, em todos os seus graus ascendentes, é lealmente aceita, segundo as evidências, com piedade natural. O fato de não poder ser interpretada mecanicamente apenas em termos de resultantes é, precisamente, aquilo pelo qual é nosso objetivo lutar com

ênfase reiterada. Mas o fato de só poder ser explicada invocando alguma força química, algum *élan* vital, alguma entelúquia, em certo sentido extranatural, parece-nos uma metafísica questionável. Pode ser que tenhamos apenas de aceitar os fatos recentemente apresentados – todos os fatos tal como os encontramos – na atitude francamente agnóstica própria da ciência. Ou pode ser que, no reconhecimento de Deus, se encontre uma explicação filosófica última, complementar à interpretação científica. Essa será a posição que eu tentarei manter.

§ II. O esquema piramidal

A tentativa mais resoluta de fornecer uma interpretação filosófica da natureza como um todo, com ênfase adequada no conceito de emergência, é a do Professor S. Alexander em *Space, Time, and Deity*. Para chegar aos fundamentos da natureza tal como ela é agora, ele convida-nos a pensar sobre tudo o que surgiu no decurso do progresso evolutivo – tudo o que pode ser excluído antes da aniquilação. Isso nos dá, como um resto inexpugnável, um plano básico de eventos basais finais (movimentos puros) com nada além dos termos espaço-temporais (pontos-istantes) em relações fluentes de ordem semelhante. Ele denomina isso de espaço-tempo, ubíquo, onipresente e inseparavelmente hifenizado. Desta primeira emergiu a “matéria”, com as suas qualidades primárias e, em uma fase posterior, as suas qualidades secundárias. Aqui sobrevêm novas relações, além daquelas que são apenas espaço-temporais. Até agora, portanto, supervenientes aos eventos espaço-temporais, temos também eventos físicos e químicos em graus progressivamente ascendentes. Mais tarde, na sequência evolutiva, surge a vida – uma nova “qualidade” de certos sistemas materiais ou físico-químicos com relações vitais supervenientes que até então não existiam. Aqui, novamente, há graus progressivamente ascendentes. Então, dentro desta matriz orgânica, ou de alguma parte altamente diferenciada dela, já “qualificada”, como ele diz, pela vida, emerge a qualidade superior da consciência ou mente. Aqui, mais uma vez, há graus progressivamente ascendentes. À medida que a evolução mental segue o seu curso, emergem, no estágio reflexivo da mente, as “qualidades terciárias” – ideais de verdade, de beleza e do eticamente correto – tendo relações de “valor”. E além disso, no ápice da pirâmide evolutiva ou próximo a ele, da qual o espaço-tempo é a base, a qualidade da deidade – a mais elevada de todas – emerge em nós, os mais recentes produtos da evolução até o momento.

Este esboço em miniatura faz pouca justiça a uma imagem construída com detalhes elaborados em uma grande tela. O tratamento pretende formular todo o plano natural de evolução. Do espaço-tempo onipresente emergem, na devida ordem histórica, o inorgânico, o orgânico e o mental, em todos os seus graus ascendentes, até que a qualidade da deidade seja alcançada em alguns homens.

Eu posso dar uma expressão diagramática – quanto mais simples e grosseira, melhor – a essa pirâmide de evolução emergente? Na sua base [Figura 1], o espaço-tempo (S.T.) se estende por tudo o que existe. No seu ápice, mas dentro dele, não menos que o espaço-tempo, está a deidade (D), uma qualidade emergente que caracteriza apenas certas pessoas no estágio mais elevado e mais recente da evolução ao longo de uma linha central de avanço. O estreitamento que dá a forma piramidal expressa o fato de que a gama de ocorrência de eventos materiais, como tais, é mais extensa do que a de eventos que também são vitais, mas não é, na opinião do Sr. Alexander, coextensiva com a gama de espaço-tempo. A seta vertical acima de N representa o que o Sr. Alexander denomina de *nisus*. Ele fala disso como o *nisus* em direção à deidade.

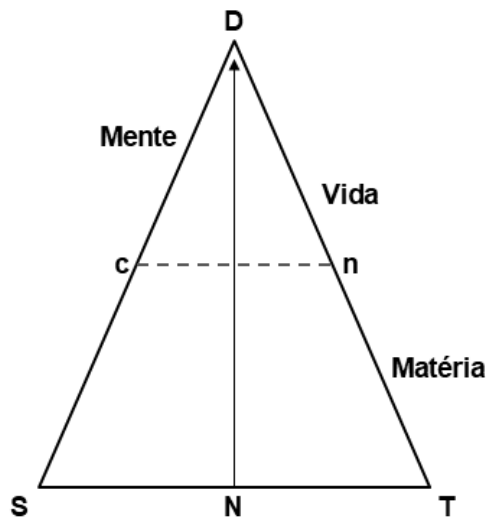


Figura 1

Tal diagrama – pelo qual o Sr. Alexander não é responsável – é, por assim dizer, uma expressão sinóptica, ou gráfico composto, de uma vasta multidão de pirâmides individuais – pirâmides de átomos perto da base, moléculas um pouco mais acima, ainda mais altas, “coisas” (por exemplo, cristais), mais alto ainda, plantas (nas quais a mente ainda não emerge), depois animais (com consciência) e, perto do topo, nossos eus humanos. Classifique como quiser; mas deixemos que cada entidade individual tenha o seu lugar apropriado na pirâmide sinóptica. A intenção é abranger todas as entidades naturais a partir dos átomos – ou, para o Sr. Alexander, de “pontos-instantes” em diante.

Não devemos supor que isso signifique que um átomo se desenvolve em uma molécula, esta em um plastídulo (ou como quer que seja denominado no nível da vida), e assim por diante. Cada entidade superior na série ascendente é um “complexo” emergente de muitas entidades de graus inferiores, dentro do qual um novo tipo de relacionamento dá uma unidade integral. Nós poderíamos dizer que cada *com-plexo* superior assume o papel de *com-plexo* em virtude de sua unidade integral; e que, quanto mais elevado o status de qualquer entidade ao longo da linha de avanço, mais ambos os ramos da palavra composta, e o conceito que ela nomeia, recebem a ênfase indicada pelo itálico.

Uma vez que é bastante certo afirmar que falar de uma qualidade de vida emergente tem cheiro de vitalismo, nós deveríamos aqui dizer entre parênteses, com a devida ênfase, que se o vitalismo conota algo da natureza da Enteléquia ou do *Élan* – qualquer inserção na evolução físico-química de uma influência alienígena que deve ser invocada para explicar os fenômenos da vida – então, longe de isto estar implícito, é explicitamente rejeitada sob o conceito de evolução emergente. Ela começa, deixe-nos dizer, com elétrons e similares; ela vê no átomo um complexo superior; vê na molécula um complexo ainda mais elevado; vê em um cristal de quartzo, ao longo de sua linha de avanço, uma entidade ainda mais complexa; e vê em um organismo, ao longo da sua linha de avanço, uma entidade com um tipo diferente de complexidade referida como integração vital. Se ela fala de vitalismo, por que não também de cristalismo, de molecularismo, de atomismo? Talvez não seja melhor, a este respeito, lançar ao mar todos esses *-ismos* e aliviar o navio de tais estorvos; ou, de qualquer forma, apenas manter “vitalismo” para designar uma doutrina que invoca (como a evolução emergente não invoca) o conceito suplementar de Enteléquia ou *Élan* de alguma ordem díspar de ser?

Aqui, descartando todos esses *-ismos*, nós procuramos indicar linhas de avanço puramente naturalistas, aceitando tais novos tipos de relacionamento como supervenientes, com piedade natural. Mas certamente não devemos supor que o progresso ao longo das linhas de avanço implique que não haja, em detalhe, nenhum retrocesso – nenhuma conversão de entidades superiores em outras de condição inferior – nenhuma degradação ou descida dentro da pirâmide. A desintegração ou involução, não menos que a integração com a evolução emergente, tem de ser levada em conta na história dos sistemas naturais.

Mais uma questão preliminar pode ser colocada em termos do diagrama – cujo benefício (tal como é) reside nas questões que provoca. Se reconhecermos alguma Atividade da qual todos os eventos piramidais são a manifestação, o diagrama sugere que, neste, naquele ou em outro nível – este da matéria, aquele da vida, o outro da mente, e talvez

acima de tudo quando a autoconsciência racional é emergente – há uma inserção especial *Ab Extra*? Ele sugere que a superveniência emergente deve ser explicada pela intervenção divina (ou outra)? Isto é exatamente o que o diagrama pretende, para o bem ou para o mal, impedir. Do ponto de vista estritamente emergente, qualquer noção de um denominado “influxo alienígena na natureza” é proibida. E se reconhecermos a Atividade Divina, da qual para a minha filosofia construtiva a evolução emergente é a expressão, ela deve ser concebida como *onipresente e manifestada em cada uma das numerosas entidades dentro da pirâmide*. Deus, se existe, está em todos, sem distinção de entidades.

E se não houver inserção Divina em pontos esporádicos – digamos, no nível da vida, da mente em seu início, ou da consciência reflexiva – não há, certamente, para nós, nenhum outro tipo de inserção. Todas as qualidades emergem *dentro da pirâmide*. A vida e a mente em nenhum sentido agem nela, ou em qualquer parte dela, de fora – de alguma ordem díspar de ser.

Na medida em que expressa, ainda que inadequadamente, o esquema filosófico do Sr. Alexander, as principais dificuldades sugeridas pelo diagrama surgem em conexão com a base e o ápice da pirâmide, e no que diz respeito ao conceito de *nisus* que introduzi no diagrama, porque ele é, penso eu, para ele uma característica fundamental. Como ele lida com essas dificuldades pode ser aprendido em seu livro, em suas declarações subsequentes em *Mind* (Vol. XXX. N.S. p. 409) e em sua recente palestra sobre *Spinoza and Time*.

Uma dificuldade adicional centra-se na relação da mente para a vida e, portanto, na ordem descendente para a matéria. Pois a mente, por um lado, e a matéria, por outro, parecem ser, em algum sentido especial, heterogêneas na própria natureza do seu ser. Como então, certamente se perguntará, um pode “emergir” do outro?

Ainda surge outra dificuldade quando lembramos que o diagrama pretende ser a expressão sinóptica de um vasto número de pirâmides individuais. Tomemos dois deles – um em que a mente é emergente; outro, digamos, um cristal de quartzo, em que o ápice não se eleva acima do nível da matéria. Como pode o primeiro, em certo sentido, conhecer (perceber) o último? Como o Sr. Alexander poderia dizer: como pode o mental, como qualidade de um, apreender o não-mental pelo qual o outro é “qualificado” de acordo com seu status evolutivo inferior? Este problema cognitivo é central para qualquer filosofia. Isso atrairá grande parte da nossa atenção em tudo o que se segue.

§ III. Involução e Dependência

Até o momento, eu usei as palavras “superior” e “inferior”, assumindo como certo que o seu significado seria entendido de uma forma geral. Com base neste entendimento, nós poderíamos concordar que os eventos naturais ao nível da mente são superiores aos do nível da vida, e estes são superiores aos eventos ao nível da matéria. Mas devemos agora perguntar: superior em que sentido? Eles podem ser superiores em mais sentidos do que um. O que quero dizer aqui, contudo – como aquilo em que o conceito piramidal se baseia em grande medida – é mais elevado em um sentido especial, do qual dependerá grande parte do meu tratamento.

Quando dois ou mais tipos de eventos, como os que mencionei anteriormente como A, B e C, coexistem em um sistema complexo de tal maneira que o tipo C envolve a coexistência de B, e B da mesma maneira envolve A, enquanto o tipo A não envolve a coexistência de B, nem de B a de C, nós podemos falar de C, como, neste sentido, superior a B, e B a A. Assim, para a evolução emergente, eventos conscientes no nível C (mente) envolvem eventos fisiológicos específicos no nível B (vida), e estes envolvem eventos físico-químicos específicos no nível A (matéria). Não há C sem B, e não há B sem A. Não há mente sem vida; e não há vida sem “uma base física”.

Observem que eu uso a palavra “envolver”. Falo de eventos em qualquer nível da pirâmide da evolução emergente como “envolvendo” eventos simultâneos em níveis inferiores. Agora, o que emerge em qualquer dado nível proporciona um exemplo daquilo que denomino de um novo tipo de relacionamento, do qual não existem exemplos em níveis inferiores. O mundo foi sucessivamente enriquecido pelo advento de relações vitais e conscientes. Nós devemos aceitar isto “com piedade natural”, como diz o Sr. Alexander. Se for encontrado como dado de alguma forma, deve ser tomado tal como o encontramos.

Mas quando algum novo tipo de relacionamento é superveniente (digamos, no nível da vida), o modo como os eventos físicos envolvidos seguem seu curso é diferente em virtude de sua presença – diferente do que teria sido se a vida tivesse sido ausente. Se for assim, com base nas evidências, também deve ser aceito com piedade natural. Parece-me que, com base nas evidências, é assim. Como, então, devemos dar-lhe expressão? Direi que esta nova maneira pela qual os eventos inferiores acontecem – este toque de novidade no avanço evolutivo – *depende do* novo tipo de relacionamento que é expresso naquilo de que o Sr. Alexander fala como uma qualidade emergente.

A posição então é esta: eventos do tipo que denominamos C *envolvem* eventos do tipo que denominamos B; e estes, por sua vez, envolvem eventos-*a*. Mas, em qualquer caso concreto, o modo específico como os eventos-*a* seguem o seu curso, aqui e agora, *depende da* presença específica de alguma fase de relacionamento-B vital; e da mesma forma, a maneira específica como esses eventos-*b* seguem seu curso – no comportamento, por exemplo – depende da relacionamento-C consciente que possa estar presente.

Devo implorar que este significado especializado associado às palavras “envolver” e “depende de”, respectivamente, seja constantemente mantido em mente. De forma alguma, estou apegado a esse modo de expressão verbal; mas acredito que, o que procuro expressar assim, é de muita importância. De qualquer forma, grande parte do que direi a seguir dependerá disso.

A ênfase na “dependência” não é menos essencial do que a na “involução”. Em um sistema físico onde a vida tenha emergido, a forma como as coisas acontecem é elevada a um plano superior. Em um organismo no qual a consciência é emergente, um novo curso de eventos depende da sua presença. Em uma pessoa em quem o pensamento reflexivo é emergente, o comportamento é sustentado em um nível mais elevado. Se a qualidade da deidade for superveniente, o plano de conduta será ainda mais elevado. Elimine a deidade e a conduta não será mais sustentada nesse nível. Elimine a consciência reflexiva e a ação será de ordem impulsiva inferior. Elimine toda consciência e comportamento orientadores que sejam apropriados ao nível de vida. Elimine a vida e o curso dos eventos descerá ao nível físico. As novas relações emergentes em cada nível superior orientam e sustentam o curso dos eventos característicos daquele nível, que, na fraseologia que sugiro, depende da sua presença contínua. Na sua ausência, segue-se a desintegração.

Deixe-me exemplificar melhor dando passos progressivos na evolução mental em seu lado cognitivo. Eu distinguirei agora (*c*) pensamento contemplativo, (*b*) percepção ingênua e (*a*) percepção sensorial. A gênese evolutiva do pensamento contemplativo envolve aquilo que já foi desenvolvido no nível inferior da percepção ingênua; e a gênese de tal percepção envolve, como historicamente anterior, a percepção sensorial. Não se pode ter pensamento a menos que a percepção tenha fornecido alguns dos dados necessários; não se pode ter percepção a menos que seus fatores representativos tenham sido derivados da percepção sensorial precedente. Até agora involução. Mas, ao nível do pensamento contemplativo, *como* a percepção segue o seu curso depende da orientação da consciência reflexiva, na medida em que coexistem; e *como* o que é dado na percepção sensorial toma forma depende da orientação da percepção, se esse nível tiver sido alcançado. Não se pode,

portanto, aceitar o velho ditado: *Nihil in intellectu quod non prius in sensu*; se isso significa que não há nada *mais* no pensamento do que na percepção ingênua, e nada *mais* nisso do que é dado primariamente pelo sentido. Não, pelo contrário, aqueles que podem ser levados a aceitar a evolução emergente considerarão este velho ditado como radicalmente falso, apenas porque não leva em conta aquela emergência da qual o progresso natural depende tanto. Leibniz's pregnant rider, *sive intellectus ipse*, recebe justificação evolutiva, embora talvez não no sentido que pretendia.

§ IV. Em direção ao Espaço-Tempo

Nós temos agora que seguir o Sr. Alexander para baixo, em direção à base espaço-temporal da pirâmide. Mas devemos primeiro compreender claramente o uso que ele faz da palavra “qualidade”. Ele fala da emergência de novas qualidades. Ele diria que, em algum estágio da evolução inorgânica, surgiu esta ou aquela denominada qualidade secundária, como a cor; que, em algum estágio posterior do processo evolutivo, surgiu a qualidade de vida; e ainda mais tarde a qualidade da consciência. Eu usarei frequentemente a palavra “qualidade” neste sentido. Mas a minha própria interpretação segue mais as linhas do que denomino de relacionamento. A discussão de relacionamento, à qual dedicarei o terceiro capítulo, requer a consideração dos termos em relação ao interior de qualquer campo de relacionamento, e das relações desses termos. Relacionamento, no meu sentido da palavra, inclui ambos; não apenas os termos; não apenas as relações; pois eles nunca poderão ser divorciados se meu uso da palavra “termo” for provisoriamente aceito. Eu falarei do relacionamento que existe inteiramente dentro de qualquer sistema como *intrínseco*; e distinguirei o relacionamento deste sistema com algum outro sistema, ou sistemas, como *extrínseco*. O sistema de relacionamento intrínseco denominarei provisoriamente de entidade. Na medida em que o caráter de uma entidade natural é determinado pelo relacionamento intrínseco, falarei dela como uma qualidade que é uma expressão desse relacionamento intrínseco. Na medida em que o caráter de uma entidade natural é determinado pelo relacionamento extrínseco com outras entidades semelhantes, falarei dele como uma propriedade que expressa esse relacionamento extrínseco (cf. § XXXIII).

Neste entendimento, o que é superveniente em qualquer fase emergente do progresso evolutivo é um novo tipo de relacionamento – novos termos em novas relações – até então não existente. Em virtude desses novos tipos de relacionamento, as entidades naturais não apenas possuem novas qualidades dentro de seu próprio ser, mas também

novas propriedades em relação a outras entidades. As entidades superiores não são apenas diferentes em si mesmas; mas elas agem e reagem de maneira diferente na presença de outras. Em qualquer fase da evolução emergente, as questões são então: qual é o novo tipo de relacionamento que sobrevém? Quais são os novos termos e quais as relações? Que diferença intrínseca existe na entidade que atinge este nível superior, e que diferença existe no seu relacionamento extrínseco com outras entidades? Quando, por exemplo, uma entidade se torna um organismo, por mais baixa que seja sua condição, que diferença intrínseca é superveniente e que diferença extrínseca existe em relação ao seu “mundo”? Deveria tornar-se uma entidade superior na qual o relacionamento consciente está presente, além de tudo o mais que está envolvido – que diferença isso faz?

Agora, para que haja uma diferença no curso dos eventos, o relacionamento em questão deve ser o que denominarei de *efetivo*. Com isto quero dizer que, quando está presente, ocorre alguma mudança no curso dos eventos existentes, o que não ocorreria se estivesse ausente.

Eu terei a oportunidade de insistir daqui em diante, contra os behavioristas radicais, que a orientação mental dos eventos conta para o progresso e indica um tipo de relacionamento que é efetivo. Quando ele está presente, ocorrem mudanças que não ocorrem na sua ausência. A maneira de agir no sistema enriquecido é diferente. É isso que quero dizer quando falo da orientação como dependente do tipo de relacionamento superveniente ao nível da mente. Passando de um estágio, aceito com piedade natural a evidência de que há mais nos eventos que ocorrem no organismo vivo do que pode ser adequadamente interpretado em termos de física e química, embora eventos físico-químicos estejam sempre envolvidos. Mudanças ocorrem no organismo quando o relacionamento vital está presente, o que não ocorre quando a vida está ausente. Esse relacionamento é, portanto, efetivo. Descendo do nível da vida para o da matéria, ninguém provavelmente negará que os tipos de relacionamento das ordens química e física são solidamente efetivos, no sentido de que o curso dos eventos é diferente quando estão presentes daquele que ocorre em suas ausências.

Aqui alguém pode intervir e perguntar: por que esta fraseologia incômoda e pedante? Por que relacionamento? Por que não esta ou aquela força como a causa de tal ou tal mudança no que você denomina a maneira como os eventos acontecem? Estamos todos bastante familiarizados com as forças da natureza inorgânica. E os materialistas costumavam dizer-nos que estas são as únicas forças e que a vida, para não ir mais longe, é uma mera recombinação sutil de eventos puramente físico-químicos. Aparentemente,

você tem que confessar que eles estavam enganados; no entanto, você evita admitir que a vida é um tipo de força nova e diferente.

Eu procuro apenas evitar ambiguidade. Sei muito bem que os físicos falam da força de coesão — para citar apenas um exemplo. Mas o que eles querem dizer? Eles querem dizer mais, neste ou em qualquer outro exemplo, do que isso, dadas tais e tais entidades que funcionam como termos em certas relações descritíveis, esta ou aquela mudança ocorre. Isto, creio eu, e nada mais do que isso, é o que a maioria dos físicos modernos quer dizer. Mas o que muitos que *leem sobre* ciência entendem que isso significa é que existe alguma agência que torna as entidades coerentes. Esta agência é o que eles entendem por força da coesão. E então perguntam por que se presume negar que a vida, também, é uma agência — a força vital que faz os organismos viverem. Há, então, alguma ambiguidade na palavra “força”. E é isso que procuro evitar ao usar a palavra “relacionamento”, que pretende excluir o conceito de “agência” ou “atividade” de qualquer lugar na interpretação científica.

Neste entendimento, nós distinguimos mente, vida e matéria. Dentro de cada uma delas, é claro, existem muitas subordens emergentes de relacionamento. Cabe à ciência resolver os detalhes — à psicologia, à biologia, à química e à física. Uma filosofia construtiva, ao lidar apenas com resultados líquidos, não deve aceitar nada que seja discrepante com as descobertas destes ramos departamentais da ciência. Nem deve aceitar nada que seja contraditório com o resultado da crítica filosófica moderna dos conceitos fundamentais sobre os quais as ciências departamentais solidamente constroem as suas superestruturas.

A posição que alcançamos, então, é a de que existem diferentes sistemas naturais a serem considerados — com sistemas mente-vida-matéria; sistemas matéria-vida; e sistemas matéria. No nível superior, existem modos de relacionamento efetivo que não estão presentes no nível médio; no nível médio, existem modos de relacionamento que não estão presentes no nível inferior.

Mas o que é o nível inferior? O Sr. Alexander nos convida a descer um degrau até o espaço-tempo. Um sistema integral sem mente é somente vida-matéria; um sistema integral sem vida é apenas material. Mas mais baixo ainda é aquilo que, na ausência de matéria, é somente espaço-tempo. Na ordem ascendente do progresso evolutivo, o espaço-tempo ainda não tem nenhum relacionamento físico efetivo; a matéria emerge, mas ainda não tem nenhum relacionamento vital, tem a matéria, mas ainda não tem nenhum relacionamento com a mente. O que então o espaço-tempo tem? Ele tem somente o relacionamento espaço-temporal em um *continuum* dentro do qual os pontos-instantes são termos em parceria que muda constantemente, proporcionando movimento puro.

Deve-se perguntar aqui: (1) o relacionamento espaço-temporal é capaz de existir independentemente de quaisquer eventos físicos? (2) Quando ele coexiste com eventos físicos que então se tornam espacial-temporal-e-físico, o fator espaço-temporal como tal é efetivo? A esta última questão não se pode dar resposta com base no *nosso* critério. Uma vez que o relacionamento espaço-temporal é ubíquo e universalmente presente, esse critério de presença ou ausência deixa de ser aplicável. Eu acredito que o Sr. Alexander responderia afirmativamente a ambas as perguntas. Espaço-tempo é, para ele, a base primordial da pirâmide e era existente antes da emergência de quaisquer eventos físicos. E o relacionamento espaço-temporal é efetivo de qualquer forma, no sentido de que permite o movimento fundamental do universo através do fluxo incessante do tempo. Os fundamentos metafísicos de seu esquema construtivo e suas implicações metafísicas multifacetadas estão estabelecidos em *Space, Time, and Deity*.

Muito mais modesto é o esquema construtivo que o alcance e a penetração mais limitados da minha visão especulativa me permitem considerar. Eu procuro em vão evidências de que o relacionamento espaço-temporal existe independentemente dos eventos físicos. Não posso ir mais fundo do que eventos que, em sua forma primordial, não são apenas espaço-temporais, mas também físicos. Além disso, embora reconheça o fluxo dos eventos físicos, sempre sujeito a relacionamentos espaço-temporais, duvido que o conceito da fluência do tempo, sobre o qual tantas coisas giram, resistirá ao teste da crítica filosófica. Que intrínseco a todo mínimo evento físico, e extrínseco como entre tais eventos, há (a) relacionamentos espaciais aqui-lá e que há (b) relacionamentos temporais agora-depois – sempre correlacionados tão inseparavelmente (ab) – me parece ser inegável; mas para a crença de que os relacionamentos aqui-lá ou agora-depois, solidamente e *como tais*, é efetiva na determinação do curso ou eventos físicos, não encontro nenhuma evidência satisfatória. Fluência existe; mas ela é a fluência dos eventos naquilo que é metodologicamente concebido como um enquadramento espaço-temporal.

Metafisicamente, meu modesto esquema não será comparável ao elaborado com admirável habilidade pelo Sr. Alexander, mas é tudo o que tenho a oferecer.

Quão longe então posso ir em direção a um nível basal de espaço-tempo? Somente até ao ponto de reconhecer um mundo físico em que o relacionamento espaço-temporal é ubíquo, mas como tal é não efetivo. Em outras palavras, eu aceito, como uma preocupação constante, tal mundo físico que possa proporcionar uma base para aquilo que tem sido descoberto no decorrer da pesquisa científica. Mas, como terei ocasião de confessar, considero a existência independente de tal mundo físico em seu direito próprio como não

suscetível de comprovação sob rígida crítica filosófica. Portanto, aceito-o sob o que denomino de reconhecimento. Aceito-o, em resumo, como parte de um esquema construtivo de evolução emergente.

Dito de outra forma, um mundo físico de forma alguma inconsistente com o resultado positivo das pesquisas departamentais nos vários ramos da ciência, por um lado, e de uma filosofia crítica que lida com os conceitos fundacionais nestes ramos da ciência, por outro – isso parece estar na base da nossa pirâmide. E se não puder ser estabelecida com base em evidências positivas, contento-me em reconhecer a sua existência como parte daquilo que pretende ser uma filosofia construtiva da evolução emergente. Sr. Alexander vai além.

§ V. Deidade

Ao tentar avançar em direção à deidade que, deve ser lembrado, é uma qualidade emergente dos sistemas naturais mais elevados que conhecemos, ou seja, algumas pessoas humanas – comecemos pelo espaço-tempo do Sr. Alexander. Mesmo neste nível basal, ele fala do tempo como a mente do espaço. Ele nos adverte, entretanto, que ele não “quer dizer que o tempo é mente ou qualquer grau inferior de mente” (STD II. p. 44). O que então ele quer dizer? Ele quer dizer, penso eu, que por todo o universo, da base ao ápice da pirâmide, existem dois “atributos” diversos, como Spinoza os denominou, ou “aspectos”, às vezes chamados de “internos” e “externos”. Todas estas palavras, por falta de melhores, nomeiam um tipo bastante singular de diversidade ou dualidade, que, afirma-se, é inerente à natureza de todos os eventos. Eu denominarei sua união inseparável de “correlação” (cf. Huxley, ii. p. 163) em um sentido da palavra que inclui o que o Sr. Alexander denomina de identidade (S.T.D. II. p. 5).

Sem subscrever a doutrina do Sr. Alexander sobre o tempo como, em qualquer sentido, a mente do espaço – isto é, excluído pela minha atitude em relação ao relacionamento espaço-temporal –, eu aceito plenamente a correlação de sistemas irrestritos e universais e não psíquicos como um reconhecimento – declaradamente especulativo, e reconhecidamente além da prova positiva (ou refutação), mas essencial para a minha filosofia construtiva da evolução. Isto significa, para mim, que não existem sistemas físicos, de status integral, que não sejam também sistemas psíquicos; e nenhum sistema psíquico que não seja também sistema físico. Todos os sistemas de eventos são, em seu grau, psicofísicos. Ambos os atributos, inseparáveis em essência, estão presentes em todo o universo das

entidades naturais. Isto é grosseiramente representado pela linha pontilhada *c.....n*, cruzando o diagrama anterior [Figura 1]. Toda entidade natural, digamos do átomo ao homem, expressa ambos os atributos enquanto ainda preserva sua identidade substancial, em algum sentido desta frase. O conceito é familiar aos estudantes de Spinoza, que diz com efeito que o aspecto físico do correlato que constitui a mente humana é o corpo (*Eth.* Pt. ii. Prop. 13), e tem o cuidado de acrescentar no *scholium* que isto se aplica “não mais para os homens do que para outras coisas individuais, todas as quais, embora em graus diferentes, são animadas”, isto é, “gozam” à sua própria maneira do correlato psíquico.

Eu repito que isso decorre da natureza das coisas e não é suscetível de prova positiva; mas também está, eu insisto, além da refutação.

Seja como for que possamos expressá-lo, há duas maneiras bastante diferentes pelas quais nós, seres humanos, estamos familiarizados com os eventos psicofísicos. De certa forma, estamos familiarizados com a sua natureza física. Por outro lado, com sua natureza psíquica. M. Bergson fala deste último modo como intuição; Sr. Alexander como gozo. E esta última forma é restrita a cada sistema integral – a você, a mim ou a outro.

Agora, quando discutimos a mente, no seu nível evolutivo apropriado, nós nos instalamos no atributo psíquico; e então reconhecemos os correlatos físicos de todos os eventos mentais nesse sistema psíquico. Mas quando nós discutimos sistemas físicos, como tais, então reconhecemos correlatos psíquicos do curso dos eventos nesses sistemas. Nós nomeamos o nível da mente do ponto de vista da abordagem psíquica, reconhecendo (entre parênteses, por assim dizer) correlatos físicos. Nós nomeamos o nível de vida a partir da abordagem física, reconhecendo (entre parênteses) correlatos psíquicos. E o mesmo acontece no nível da matéria. Portanto, nosso esquema abrangente funciona assim:

C, Mente (com correlatos físicos).

B, Vida (com correlatos psíquicos).

A, Matéria (com correlatos psíquicos).

É no nível C que parece haver evidência positiva de *alguma* correlação. Mas mesmo aqui nós devemos confessar que não há nenhuma prova positiva de que *todos* os eventos mentais tenham tais correlatos. Nós aceitamos, portanto, isto sob um reconhecimento

que vai além da evidência, mas não é, pensamos, contraditório a ela – assumindo assim uma posição filosófica bastante familiar. Nós insistimos que, pela natureza do caso, só podemos “gozar” dos correlatos psíquicos da vida e da matéria que estão envolvidos em todo o sistema psíquico integral em nosso nível de mente. Com correlatos psíquicos da vida apenas no nível B, e da matéria apenas no nível A, nós não podemos ter conhecimento direto; pois não podemos *ser* uma ameba em um nível, ou uma molécula em outro nível, de modo a estarmos *assim* familiarizados com o atributo psíquico que só ela pode “gozar”.

Não é necessário acrescentar que não há relação causal entre um atributo e outro. Para modernizar Spinoza: o plano ordenado de avanço no atributo psíquico está estritamente correlacionado com o do atributo físico. Nós temos “uma e a mesma coisa [evolução], embora expressa de maneiras diferentes” (Eth. Pt. II. Prop. vii. cf. *scholium*).

Afastei-me do meu texto – tratamento do Sr. Alexander – para colocar a posição à minha maneira e não a dele. Volto agora ao texto para levantar algumas questões – em parte verbais, mas certamente também algo mais.

Começando com o tempo como, no sentido pretendido, a mente do espaço, o Sr. Alexander considera cada qualidade na hierarquia ascendente como um aspecto mental superior superveniente no curso do progresso evolutivo e aquilo que está abaixo dele – aquilo que denomino como envolvido – como jogar para esta mente a parte de seu corpo. Assim, uma qualidade secundária é, diz ele, a mente de seu substrato primário (S.T.D. II. p. 60); vida é a mente dos eventos físico-químicos; consciência é a mente do organismo vivo nos quais ela emerge. Além disso, se não me engano, cada qualidade superior desempenha também o papel de deidade para aquela que está abaixo dela.

Eu não posso seguir aqui seu exemplo. De qualquer forma, eu deveria colocar a posição de maneira diferente. A palavra “mente” parece ser usada em pelo menos dois sentidos: (1) como o nome de uma qualidade em um nível distintamente emergente; (2) significar o correlato daquilo que também é um processo corporal ou físico – tal correlato implica identidade substancial. Se for usada no primeiro sentido, parece desaconselhável aplicá-la a qualquer coisa que não seja essa qualidade distintiva. O máximo que pode ser dito é que a vida está para a matéria no mesmo tipo de relação que a mente está para a vida. E isto, creio, pode ser melhor expresso dizendo que a vida envolve uma base de matéria, assim como a mente envolve uma base de vida. A relação comum a ambas é aquela que chamo de involução.

E se a palavra “mente” for usada neste último sentido, devo insistir que, *qua* correlata, ela se situa *no mesmo nível* daquela com a qual está correlacionada – e não acima dela.

Eu penso, contudo, que o desacordo aqui reside, no fundo, mais nos modos de afirmação do que nos princípios de interpretação. Em todo caso, o que é comum a ambas é aquela hifenização dos dois atributos que eu denomino de correlação e ele de identidade substancial.

Há outra razão pela qual considero indesejável usar a palavra “mente” nos dois sentidos, (1) como uma qualidade que emerge em um nível atribuível na nossa hierarquia, e (2) como atributo correlacionado a *todos* os níveis. Muitos críticos da nossa tese insistirão que a *mente não emerge*. Isto é verdade para a mente (2) como correlata; *essa* mente não emerge. Mas existem níveis emergentes dessa mente, como correlatos; e é, em um nível atribuível, que a mente (2) *emerge*. É uma qualidade emergente da ordem psíquica correlacionada em um estágio aproximadamente definível de avanço evolutivo. Portanto, se dissermos que a mente emerge nesta fase, enquanto outros afirmam que a mente não emerge e não pode ser tratada como emergente, isto *pode* ser porque a palavra “mente” é usada nestes dois sentidos diferentes.

No que diz respeito à deidade, para a qual estamos trabalhando, devemos voltar ao *nisus*. Nós vimos que de um nível na hierarquia de níveis um novo tipo de existência emerge. Este fato de emergência progressiva é *nisus*, que é, portanto, algo mais do que o *conatus* de Spinoza (cf. §XXIV.). “Assim, o *nisus* do mundo se reflete na transformação de tipos que ocorrem, como atestado pela observação e pela teoria, dos níveis inferiores para os níveis superiores”. E é compartilhado por tudo. Dentro da nossa consciência reflexiva, dentro das mentes dos organismos inferiores, e mesmo das coisas materiais, ela “é sentida como um *nisus* em direção a algo não alcançado” (Sp. T. p. 72-7).

Aqui novamente eu tenho dificuldades. Em primeiro lugar, eu devo dizer que o *nisus* em direção à deidade – se a deidade for uma qualidade superveniente à consciência reflexiva – segue *uma* linha de avanço bastante específica. Caso contrário, não consigo ver que resposta pode ser dada à pergunta: como você propõe caracterizar esta qualidade de deidade? Não deveria ser suscetível pelo menos de indicação, se não de definição? Se assim for, podemos dizer: deidade é aquilo que é exemplificado nesta ou naquela pessoa, ou grupo de pessoas, que atinge este nível emergente. Se *todas* as linhas de avanço exemplificam um *nisus* em direção à deidade – então a deidade deve ser caracterizada de alguma maneira diferente – digamos, aquela que é mais elevada em qualquer linha de avanço.

Agora eu quero colocar pressão sobre *uma* linha de avanço e interpretá-la literalmente; e eu quero enfatizar o que isso pode implicar. Parece, conforme penso com base nas evidências, que quanto mais nós ascendemos na hierarquia – e especialmente quando alcançamos as pessoas humanas – a complexidade emergente é tal que parece justificável dizer que não existem duas pessoas exatamente iguais. Cada pessoa é um produto exclusivamente individual ao longo de uma das muitas linhas de avanço – digamos, Shakespeare, Goethe, Newton e Darwin. Se assim for, o *nisus* em direção à deidade em sua linha estritamente central deveria culminar em uma pessoa única, no próprio ápice da pirâmide. Se uma pesquisa histórica imparcial deve levar à conclusão de que o *nisus* em direção à deidade culminou em um indivíduo único, não há, até onde posso ver, nada na interpretação naturalista da evolução emergente que impeça a aceitação desta conclusão.

Tomemos, no entanto, a visão de que *todas* as linhas de avanço exemplificam, em algum sentido, um *nisus* em direção à deidade, e que a deidade deve ser definida de tal maneira que esteja de acordo com esta visão – digamos “algo mais elevado ainda não alcançado” – mesmo assim o inatingível, como tal, parece implicar aquilo que ainda não é; e isto, penso eu, só pode ser previsto, ainda que obscura e vagamente, quando o nível da mente consciente for atingido.

Mas, à parte algumas questões em que sou levado a discordar, eu concordo plenamente que, ao longo de multifacetadas linhas de avanço – todas, sem dúvida, interligadas por linhas ramificadas em um sistema da natureza – há, como resultado líquido, um progresso ascendente e para frente (apesar de muitos retrocessos), que é o que procuro enfatizar no âmbito da evolução emergente.

Nós chegamos então finalmente ao conceito de Atividade.

Uma questão antiga – não na ciência, mas na filosofia – assume para nós uma nova forma. O que faz surgir os emergentes? O que dirige o curso dos eventos em que uma linha saliente é o *nisus* em direção à deidade? Alguns podem dizer que não sabemos e não podemos saber. Outros podem perguntar que necessidade existe de uma Fonte de emergência diretiva. Por que não deveria prosseguir sem uma? No entanto, outros podem insistir que é inútil colocar em uma Fonte ativa exatamente o que se diz que sai dela. Pois se houver menos, resta alguma coisa a ser explicada; e se houver mais (*eminenter*, como disseram os escolásticos), que evidência disso ainda está disponível?

Por outro lado, pode-se afirmar que aquilo que mencionei no início como o plano abrangente de sequência de todos os eventos naturais é certamente, por si só, evidência

suficiente do Propósito; e isto implica, diz-se, alguma Mente através de cuja Actividade (eu uso letras maiúsculas iniciais para conceitos deste tipo), o curso dos eventos é dirigido. Nós temos, no entanto, neste conceito de Mente, (3) algo diferente da mente (1) como uma qualidade emergente, pois esta é diferente da mente (2) como correlato universal. É difícil evitar o uso da palavra “Mente” neste sentido. Talvez a letra maiúscula seja suficiente para indicar que ela compreende indefinidamente mais do que a mente no sentido (1); ou pode ser diferenciada como Espírito.

Agora, além da correlação, que eu aceito sob o que chamo de reconhecimento, aceito, também sob reconhecimento, um mundo físico que existe por si só, de forma bastante independente de qualquer mente humana ou subumana. Por que aceito isso *sob reconhecimento*? Porque não estou convencido de que a sua existência possa ser irrefutavelmente estabelecida sob a luz da crítica filosófica moderna. Eu admito então que, ao aceitá-lo, vou além da evidência positiva. Mas afirmo que ele não incorpora nada que seja discrepante com ou contraditório a essa evidência. Como, então, posso alcançar este mundo físico reconhecido? Seguindo para baixo a linha da “involução” até chegar ao que é, para a minha filosofia construtiva, o conceito limitante. Mas se, da mesma maneira, eu sigo para cima a linha da “dependência”, eu alcanço novamente (para a minha filosofia construtiva) um conceito limitante – o da dependência última em termos do qual todo o curso da evolução emergente é explicado (e não meramente interpretado) dentro de um esquema consistente e equilibrado. Isso também eu aceito sob reconhecimento. Também ele está, penso eu, além da prova pela evidência positiva que a crítica filosófica exige e, dentro do seu âmbito, tem razão em exigir. Mas é discrepante com ou contraditório a qualquer evidência positiva que somos obrigados a aceitar com piedade natural? Eu acho que não. E sinto-me, portanto, livre para exortar a sua legitimidade sob reconhecimento. Isto, para mim, conduz para cima em direção a Deus, como atividade diretiva dentro de um esquema que visa a consistência construtiva.

Muito mais, é claro, está nos bastidores desta maneira divulgada. Devemos buscar a relação do Deus assim mal reconhecida com aquelas pessoas nas quais existe alguma medida da qualidade da deidade. Para o Sr. Alexander, a deidade, não menos que a mente (no sentido 1), é uma qualidade emergente. Ele distingue entre “deidade como uma qualidade e Deus como um ser”. E ele diz que “Deus, como realmente possuidor de deidade, não existe, mas é um ideal, está sempre se tornando; mas Deus, como todo o universo tendendo à deidade, existe” (*Mind*, XXX. p. 428). De acordo com a segunda parte desta afirmação, com o seu anel de Spinoza, Deus, como ser, é o *nisus* do universo avançando para níveis ainda não alcançados; ou, como eu preferiria dizer, é o *Nisus* diretivo do curso

dos eventos. No que diz respeito à primeira parte, surge a questão crucial se, e em caso afirmativo, em que sentido, tal ideal é verdadeiramente Real.



