



## O LIMITE QUALITATIVO DE MODELOS QUANTITATIVOS<sup>1</sup>

*Oswaldo Pessoa Junior*

Professor do Depto. de Filosofia, FFLCH, USP. Mestre em Física Experimental pela Universidade Estadual de Campinas e doutor em História e Filosofia da Ciência na Indiana University, EUA. E-mail: [opessoa@usp.br](mailto:opessoa@usp.br)

### RESUMO

O presente artigo discute a limitação qualitativa de modelos matemáticos e computacionais. Por “qualitativo” entende-se não uma descrição linguística *fuzzy* ou sem o uso de matemática, mas sim a vivência de qualidades subjetivas e a extensão deste conceito para a materialidade das coisas, em oposição à suas relações estruturais. Explora-se didaticamente o “espectro funcionalista”, em que diferentes posições filosóficas a respeito da natureza da consciência são apresentadas. Esta análise opõe, de um lado, as visões comportamentalistas, funcionalistas e mecanicistas, e de outro, as concepções mentalistas, substancialistas e qualitativas. Mapeiam-se assim diferentes respostas à pergunta de se uma máquina pode ser consciente.

**PALAVRAS-CHAVE:** Comportamentalismo. Funcionalismo. Mentalismo. Modelos computacionais. Qualitativo.

### THE QUALITATIVE LIMIT OF QUANTITATIVE MODELS

### ABSTRACT

The paper discusses the qualitative limitation of mathematical and computational models. By “qualitative” one does not mean a fuzzy linguistic description, without the use of mathematics, but the experiencing of subjective qualities and the extension of this concept to the materiality of things, in opposition to structural relations. One explores didactically the “functionalist spectrum”, in which different philosophical positions about the nature of consciousness are presented. The analysis divides, on one

<sup>1</sup> A questão dos limites dos modelos computacionais e mentais foi desenvolvida em duas apresentações feitas em 2015, no *Workshop Modelos em Ciências*, realizado em 27 de abril na Fafich, UFMG, a convite de Patrícia Kauark-Leite, e no *Seminário de Pesquisa em Ontologias do Brasil*, em 8 de setembro no IMEUSP, a convite de Renata Wasserman. Já a exploração do espectro funcionalista foi feita inicialmente na disciplina de Filosofia das Ciências Neurais, dada no Depto. de Filosofia da USP em 2016.

side, behaviorist, functionalist and mechanistic views, and on the other, mentalist, substantialist and qualityist conceptions. Different answers to the question of whether machines can be conscious are explored.

**KEYWORDS:** Behaviorism. Functionalism. Mentalism. computational models. Qualityism.

## 1 MODELOS TEÓRICOS CAPTURAM ESTRUTURAS

Considerarei um “modelo” como um sistema construído por semelhança parcial com outro sistema, como um modelo plástico de zepelim, ou um modelo teórico de uma estrela. O que caracteriza um modelo teórico na ciência é o fato de ele ser expresso em uma linguagem que geralmente inclui a matemática ou a computação.

Pode-se dizer que um modelo matemático ou computacional “captura” a realidade? Tomemos como exemplo a simulação computacional de um furacão. O que está sendo capturado são as *relações* entre as partes da atmosfera, ou seja, a estrutura dinâmica da realidade. Com a informação armazenada no computador, podem-se prever aproximadamente os efeitos causais de situações reais e também de situações contrafactuais, como por exemplo, o que aconteceria com um avião que entrasse em um furacão.

Mas há algo na realidade que *não se pode* capturar através da modelagem computacional. No caso da simulação de um furacão, não se obtém um modelo realmente *molhado*, pois não há muita água dentro do computador. O computador pode ter um simulador do comportamento da água, a partir da representação de elementos moleculares, mas isso não é a mesma coisa que *ser* água.

O modelo captura apenas a organização ou estrutura do sistema, e não os elementos em si, não a materialidade do sistema representado. Todo modelo tem sua materialidade própria, sejam os dispositivos e circuitos elétricos de um computador, sejam as estruturas celulares em um encéfalo (segundo a concepção materialista do problema mente-corpo). Mas se a materialidade do sistema que representa for distinta da do objeto que é representado, esta distinção estabelece um limite “qualitativo” para a modelagem.

Consideremos agora um modelo ou uma teoria de como a mente emerge a partir do encéfalo. Por mais detalhada que seja esta teoria, ela não captura a matéria real do encéfalo, e também não captura as qualidades subjetivas que uma pessoa vivencia, os chamados “*qualia*”, como a amarelidão percebida ao se olhar para uma banana. Esta distinção entre a descrição

quantitativa-relacional fornecida pelo modelo e a materialidade do objeto representado é expressa de maneira clara no famoso experimento mental do quarto de Mary (JACKSON, 1982).

Mary seria uma cientista do séc. XXIII que possuiria todo conhecimento teórico a respeito da neurofisiologia das cores, mas que nunca teria tido a vivência subjetiva dos *qualia*, pois teria vivido presa em um quarto sem cores. Ao finalmente sair do quarto e observar uma mancha amarela na parede, teria ela adquirido “conhecimento novo”? Num certo sentido sim, pois ela travou contato com uma nova experiência subjetiva. Este experimento mental tem sido interpretado das mais variadas formas (cf. NIDA-RÜMELIN, 2009), mas aqui o interpretamos como fazendo uma distinção entre conhecimento teórico, linguístico-quantitativo (que Mary possuiria dentro do quarto), e conhecimento por contatos direto (*acquaintance*), ambos pertinentes para o conhecimento científico.

A tese de que o mundo pode ser capturado de maneira completa pela descrição linguístico-quantitativa pode ser chamada de *mecanicismo*. Na filosofia da mente, a concepção mecanicista da consciência é chamada de “funcionalismo”.

## 2 DEFINIÇÃO DE CONSCIÊNCIA

O que é a consciência? Em que momento surgiu na evolução biológica? Como é possível a subjetividade em um mundo físico? Antes de prosseguir, tentemos definir o conceito geral de “consciência”, seguindo alguns dicionários de filosofia.

“Consciência” é a intuição, mais ou menos clara, que o sujeito tem dos seus estados e dos seus atos (LALANDE, 1999, p. 195). É a possibilidade que cada um tem de dar atenção aos seus próprios modos de ser e às suas próprias ações, de estar ciente dos próprios estados, percepções, ideias, sentimentos, volições etc. (ABBAGNANO, 2007, p. 217). É a mente consciente, em oposição à mente inconsciente ou subconsciente (RUNES, 1942, p. 64).

Consideremos, porém, uma citação clássica do filósofo escocês William HAMILTON (1877, p. 191):

A consciência não pode ser definida: nós mesmos podemos estar totalmente cientes do que seja a consciência, mas não conseguimos sem confusão transmitir para os outros uma definição do que nós mesmos apreendemos com clareza. A razão é simples: a consciência está na raiz de todo conhecimento.

Cientistas e engenheiros muitas vezes exigem que definamos claramente o conceito de “consciência”, por meio de palavras, antes de aceitarem prosseguir na discussão. Mas as palavras foram desenvolvidas para a comunicação intersubjetiva, e o fenômeno que queremos tratar é subjetivo, privativo a cada indivíduo. Para ilustrar esta situação, consideremos um experimento mental.

### 3 O TESTE DE TURING ROBÓTICO

Considere uma situação como a do filme *Metrópolis*, de Fritz Lang (1927), em que um robô metálico é construído pelo cientista malvado e transformado na ginoide (androide feminina) Maria, interpretada pela atriz Brigitte Helm, indistinguível de um ser humano. Vamos supor que, em seu interior, Maria é feita de válvulas (ou chips de circuito integrado), fios e motores metálicos, mas julgando por seu comportamento, suas expressões faciais e sua fala, todos a consideram um ser humano normal. Nesse sentido, pode-se dizer que ela passou no “teste de Turing robótico” (HARNAD; SCHERZER, 2008).

Alan Turing, o genial matemático e computeiro, escreveu em 1950 um artigo discutindo se máquinas poderiam “pensar”. Ao invés de tentar dar uma definição lexicográfica de “pensamento”, propôs um jogo ou teste para estabelecer se uma máquina pensa. Simplificando um pouco, o “jogo da imitação” envolveria um ser humano que faria perguntas ou dialogaria com a máquina ou com um humano escondido atrás de uma parede, e teria que fazer a identificação correta. TURING (1973, p. 61) fez a previsão de que no ano 2000 os computadores poderiam ser programados para “jogar o jogo da imitação tão bem que um interrogador médio não teria mais de 70% de probabilidade de chegar à identificação correta, após 5 minutos de interrogatório”. Ou seja, se uma máquina conseguir enganar um ser humano, de maneira que este julgue que está dialogando com outro ser humano, então a máquina deve ser considerada pensante ou inteligente.

Turing não está se referindo diretamente à “consciência”, mas na discussão das possíveis críticas à sua abordagem, considera o que chama “o argumento da consciência”:

Este argumento parece ser uma negação da validade de nosso teste. De acordo com a forma mais extremada de tal concepção, a única maneira de alguém estar seguro de que a máquina pensa é *ser* ele a própria máquina e sentir-se pensando. Poderia então descrever estes sentimentos [*feelings*] ao mundo, mas naturalmente não se justificaria que alguém lhe desse atenção. De modo semelhante, de acordo com a mesma concepção, a única maneira de saber-se

se um *homem* pensa é ser esse homem. Trata-se, de fato, do ponto de vista solipsista (TURING, 1973, p. 65).

Turing conclui então que o proponente do argumento da consciência não iria querer assumir uma posição “solipsista”, ou seja, uma posição que assume que só existiria a sua própria mente. Finalizando a discussão deste argumento:

Não quero dar a impressão de que penso que não exista nenhum mistério no que respeita à consciência. Existe, por exemplo, algo assim como que um paradoxo vinculado às tentativas de localizá-la. Mas não acredito que tais mistérios tenham de ser necessariamente resolvidos antes de podermos responder à pergunta que nos preocupa neste artigo (TURING, 1973, p. 67).

As condições específicas colocadas por Turing para o ano 2000 ainda não foram plenamente satisfeitas, mas a cada ano o Prêmio Loebner é oferecido para a melhor performance de máquina. O futurólogo Ray Kurzweil (2005, p. 295) prevê que o teste de Turing será plenamente superado em torno de 2029.

Voltemos agora para o teste de Turing robótico envolvendo a ginoide Maria. A pergunta a ser considerada é a seguinte: a ginoide é consciente? Por hipótese, ela se comporta como uma humana, fala como uma humana, é maquiavélica como nós, tem a pele igual a de um humano e tem cheiro de humano, como Ava, do filme *Ex machina* (2015). Mas Ava tem um cérebro feito de matéria especial, ao passo que a ginoide que estamos discutindo é feita de circuitos integrados. Maria teria consciência?

#### 4 COMPORTAMENTALISMO FILOSÓFICO

A concepção de que a natureza de uma coisa se esgota com sua aparência, ou com seu comportamento externo, acoplado a informações sobre a sua história, pode ser chamada de “comportamentalismo”, ou *behaviorismo*. Como é bem sabido, este termo está associado a uma corrente específica da psicologia, que na verdade se divide em diversas correntes. O’Donahue e Kitchener (1999) descrevem pelo menos catorze versões do comportamentalismo, dez psicológicas e quatro filosóficas (cf. ABIB, 2015, p. 283).

Não exploraremos este tema aqui, mas definiremos sucintamente o “comportamentalismo filosófico” (OPPY; DOWE, 2016) como a visão que afirma que um estado mental só pode ser atribuído a um sistema baseado em seu comportamento observável.

Isso é resumido na frase de Wittgenstein (1979, p. 580): “Um ‘processo interno’ necessita de critérios externos”. Esta concepção também é expressa na seguinte citação do psicólogo John Staddon (1999, p. 230), referindo-se à corrente que propõe, chamada “behaviorismo teórico”:

O behaviorismo teórico adota a visão do teste de Turing da consciência. Porém, esta visão não é aceita por todos. John Searle (1992), se o compreendi corretamente, argumenta que se um dispositivo fosse encontrado que conseguisse passar no teste de Turing, mesmo assim ele não seria consciente. Tenho três reações a isso. Primeiro, a suposição de que tal dispositivo *pode* ser criado somente a partir de hardware pode ser falsa, e neste caso nada mais precisa ser dito. Em segundo lugar está uma pergunta óbvia: supondo que ele possa ser criado, *como você sabe* que ele não é consciente? A única resposta a esta pergunta é: porque ele não passa no teste de Turing, o que contraria a primeira suposição. Em outras palavras, se a única maneira de saber que alguém (ou alguma coisa) é consciente é porque ela responde às nossas perguntas de maneira apropriada, então, por definição, uma máquina que consegue passar no teste de Turing deve ser consciente. Minha terceira reação é simplesmente esperar para ver. Se uma máquina for criada que passe no teste de Turing, as pessoas logo vão tratá-la como alguém de seu tipo. Se estivermos dispostos a atribuir consciência para um cachorro, ou para alguém cuja habilidade para comunicar esteja tão prejudicada como para Helen Keller, será que deixaremos de atribuí-la para um dispositivo que fala e responde de maneira indistinguível de um ser humano?

## 5 MENTALISMO

A posição que nega que a ginoide tenha consciência, na falta de um nome melhor, pode ser chamada de “mentalismo”, em oposição ao comportamentalismo filosófico. O *mentalismo*, então, seria a tese de que há uma perspectiva qualitativa ou subjetiva que caracteriza a consciência, e que vai além do comportamento externamente verificável. Esta tese está encapsulada na citação de Hamilton da seção 2.

Um debate oposto mentalistas e comportamentalista se deu na década de 1980 com relação à consciência animal. O estudioso do comportamento animal e psicólogo gestaltista Wolfgang Köhler havia colhido na ilha de Tenerife, durante a Primeira Guerra Mundial, indícios de comportamento inteligente em chimpanzés, que conseguiam empilhar caixas para subir e pegar bananas presas no alto da jaula. Ele concluiu que isso teria sido feito não por mera tentativa e erro, mas por “insight” (*Einsicht*), sugerindo que os chimpanzés possuem consciência como nós (a menos da linguagem e outras funções cognitivas superiores). Em 1945, Herbert G. Birch mostrou que experiência prévia com os instrumentos era necessária

para a ocorrência deste comportamento inteligente. Em 1981, Epstein, Lanza & Skinner também criticaram a conclusão de Köhler, mostrando que é possível condicionar pombos com um repertório de comportamentos de maneira que eles possam reproduzir o comportamento sofisticado observado em chimpanzés. Para eles, isso sugeria que era leviano atribuir uma mente aos macacos, pois o seu comportamento poderia ser completamente explicado pelo seu repertório de condicionamentos, conforme sugerido por Birch (SHETTLEWORTH, 2012). Este resultado é muito interessante, mas não afeta a tese de que há um estado mental que sirva de causa intermediária na cadeia causal ligando condicionamentos e comportamento. Além disso, hoje em dia há fortes evidências de que várias espécies de pássaros têm inteligência comparável à de chimpanzés (EMERY, 2006).

## 6 FUNCIONALISMO DE ESTADOS MENTAIS

Na filosofia da mente, o funcionalismo é a posição que define estados mentais por suas *funções*, e não por sua natureza material. Assim, os estados mentais seriam constituídos pelas relações causais existentes entre eles, além de pelas entradas sensoriais e pelas saídas comportamentais. Não importaria qual é o substrato da mente – matéria orgânica, matéria inorgânica, espírito –, o que importa seria a organização do sistema, ou o estado informacional. O chamado “funcionalismo de máquina” compara estados mentais com estados lógicos (*software*) de um computador, que existiriam independentemente da natureza física do computador (*hardware*).

O funcionalismo foi articulado por autores como Hilary PUTNAM (1967) e Jerry FODOR (1981) (ver discussão em BLOCK, 1980), mas a sua tese principal remonta a Aristóteles (cf. SHIELDS, 1990).

A intuição por trás do funcionalismo é que o que determina o tipo psicológico ao qual um particular mental pertence é o papel causal do particular na vida mental do organismo. Individuação funcional é diferenciação com respeito ao papel causal. Uma dor de cabeça, por exemplo, é identificada com o tipo de estado mental que, entre outras coisas, causa uma disposição para tomar aspirina em pessoas que acreditam que a aspirina alivia uma dor de cabeça, causa um desejo de se livrar da dor que se está sentindo, frequentemente causa alguém que fala português a dizer coisas como “Estou com dor de cabeça”, e surge a partir de excesso de trabalho, fadiga ocular e tensão. Esta lista presumivelmente não está completa. Mais será conhecido sobre a natureza de uma dor de cabeça, à medida que pesquisas em psicologia e fisiologia descobrirem mais sobre o seu papel causal (FODOR, 1981, p. 128).

Diversas faculdades mentais são definidas de maneira funcional, como por exemplo, “capacidade de cálculo”. Um indivíduo com “síndrome de *savant*” tem uma capacidade de cálculo assombrosa, mas um computador de última geração tem uma capacidade ainda maior. Neste exemplo, a faculdade mental é definida pela eficiência de realizar um cálculo matemático, não importando como e nem qual o substrato material.

A tese funcionalista forte é que todos os estados mentais são definíveis em termos funcionais. Isso incluiria a própria consciência. A tese de que máquinas mecânicas ou computadores de circuitos integrados de silício podem ter consciência é chamada “funcionalismo de máquina” ou “inteligência artificial forte” (vimos, porém, que o debate se complica pelo fato de diferentes definições de consciência serem usados).

Uma maneira de exprimir esta posição é dizer que computadores não apenas podem “simular” a mente humana, como se tenta fazer com o *Blue Brain Project* (Projeto Cérebro Azul) de Lausanne, mas poderão “emular” a consciência, ou seja, fazê-la emergir a partir de uma computação altamente complicada. Certamente um computador só pode *simular* um furação, não emulá-lo, pois o seu interior não fica molhado; será que um computador poderia ir além da simulação de um cérebro, emulando a consciência subjetiva? O funcionalismo de máquina defende que sim.

## 7 ASPECTO QUALITATIVO DO MENTAL

O maior problema para o funcionalismo é dar conta do aspecto qualitativo do mental, o que Fodor chama de “conteúdo qualitativo”:

Não é fácil dizer o que é o conteúdo qualitativo. De fato, segundo algumas teorias, não é possível sequer dizer o que ele é porque ele não pode ser conhecido por descrição, mas apenas através da experiência direta. Apesar disso, eu tentarei descrevê-lo. Tente imaginar-se olhando para uma parede em branco através de um filtro vermelho. Agora, troque o filtro vermelho por um verde e deixe o resto exatamente como estava antes. Alguma coisa relativa ao caráter de sua experiência se altera quando o filtro é trocado, e é esse tipo de coisa que os filósofos chamam de conteúdo qualitativo. [...] A razão pela qual o conteúdo qualitativo representa um problema para o funcionalismo é evidente. O funcionalismo está comprometido com a definição dos estados mentais em termos de suas causas e seus efeitos. Parece possível, no entanto, que dois estados mentais possam ter as mesmas relações causais, diferindo, porém, em seu conteúdo qualitativo (FODOR, 1981, p. 130).



Este é o ponto central da presente discussão, a natureza “qualitativa” da vivência subjetiva consciente, cujos elementos têm sido chamados de “dados dos sentidos” ou “*qualia*”. Argumentos como o do “espectro invertido”, que remonta a John Locke (1999, p. 518-9), são usados para questionar a validade do funcionalismo de estados mentais, como colocado por Fodor na última oração citada acima.

Ao discutirmos os “limites qualitativos dos modelos”, é preciso distinguir dois sentidos de “qualitativo”. Qualitativo-1 estaria associado a uma descrição linguística de um fenômeno sem o uso de matemática, por exemplo: “ocorreu uma transição de fase ao esfriar a água, que se solidificou”.

Não é este o sentido que estamos utilizando no presente texto, mas sim o de qualitativo-2, que se refere às vivências subjetivas, aos *qualia* associados a percepções e emoções, conforme colocado por Fodor. Modelos matemáticos e computacionais podem se aproximar bem de uma descrição qualitativa-1, por exemplo, utilizando às técnicas associadas à lógica *fuzzy*, mas não capturar o qualitativo-2 do objeto modelado.

## 8 EXPERIMENTO MENTAL DA TROCA DE CÉLULAS CEREBRAIS POR CHIPS

Uma boa maneira de explorar as posições em filosofia da mente é apresentando experimentos mentais (*Gedankenexperimenten*, *thought experiments*, também traduzido por “experimentos de pensamento”), ou seja, experimentos que não podem ser realizados, pelo menos não hoje em dia (por razões técnicas), ou então por uma questão de princípio.

Já examinamos o experimento mental da ginoide feita de circuitos integrados, que definiu duas posições: o comportamentalismo filosófico e o mentalismo. Consideremos agora o experimento mental da troca de células cerebrais por chips de silício, explorada por John SEARLE (1997, p. 98). Suponha que cada célula de encéfalo é substituída por um chip que reproduz as mesmas relações causais (de entrada e saída) conhecidas hoje em dia, e que os chips tenham a mesma plasticidade que as células reais, estabelecendo novas conexões sinápticas, etc.

Ao final da substituição, a “pessoa” terá consciência ou não? Novamente, não temos certeza do que aconteceria, pois o experimento é apenas imaginado. Mas podemos definir duas posições antagônicas com relação ao resultado esperado do experimento. A suposição de que a consciência se manteria intacta caracteriza um *funcionalismo no nível de células biológicas*:

a organização entre as células (desprezando os detalhes internos) seria suficiente para fazer emergir a consciência subjetiva.

Esta é uma posição mentalista, pois não define “consciência” apenas a partir do comportamento externo. Em oposição a este tipo de funcionalismo, temos a posição que Searle chama de *naturalismo biológico*: alguma coisa que ocorre nos processos biológicos dentro das células seria essencial para a emergência da consciência. Assim, das três alternativas colocadas na citação de Staddon, a primeira seria a correta.

## 9 O ESPECTRO FUNCIONALISTA

Para fechar este capítulo, consideremos com a Tabela 1 as várias escalas em que se dá a oposição entre, de um lado, o comportamentalismo e funcionalismo, e de outro, o mentalismo e o naturalismo biológico. Representamos cinco posições (linhas na Tabela), mas outras intermediárias são possíveis.

O que caracteriza o lado funcionalista do espectro é considerar que a consciência possuída por um sistema define-se somente pelas *relações* entre partes do sistema (ou seja, pela organização do sistema, o “arranjo” dos atomistas, a “forma” aristotélica), ou simplesmente pelo comportamento do sistema, sem importar qual é a constituição material do sistema (ou a constituição espiritual).

No limite, a posição metafísica mais radical tem sido chamada de “estruturalismo ontológico” (ou realismo estrutural ontológico), e considera que tudo em nosso Universo são relações ou estruturas, e que no fundo a própria matéria (e outras grandezas físicas) emerge de relações sem *relata*, ou seja, de relações de relações de relações, etc., sem a presença de elementos básicos (STEINLE, 2010).

A negação desta concepção pode ser chamada de *substancialismo*, ou seja, a noção de que há uma substância, que pode ser a matéria (ou outra entidade física) ou outra categoria de entidade, irreduzível a relações. O hilemorfismo de Aristóteles seria uma visão substancialista (mesmo que mínima), na medida em que na natureza não haveria forma sem matéria. O naturalismo biológico seria também uma expressão do substancialismo, de maneira que o espectro complementar ao espectro funcionalista pode ser chamado de “espectro substancialista”.

| EXPERIMENTO MENTAL  | PERGUNTA  | SIM  | NÃO                              |
|---|---|--|----------------------------------|
| Ginoide de circuitos integrados   | A ginoide é consciente?   | Comportamentalismo filosófico                                    | Mentalismo (genérico)            |
| Programa que simula estados mentais ou células cerebrais<br>( <i>Blue Brain Project</i> ) | Um simulador da mente ou cérebro pode ser consciente?                       | Funcionalismo de máquina (ou de estados mentais)<br>(I.A. forte) | Mentalismo anti-funcionalista    |
| Substituição de células cerebrais por chips   | Substituindo cada célula por chips com mesma função preserva a consciência? | Funcionalismo de células biológicas                              | Naturalismo biológico celular    |
| Substituição de partes das células (biologia sintética)                                   | Substituindo protoplasma e organelas de células preserva a consciência?     | Funcionalismo subcelular   | Naturalismo biológico subcelular |
| -----   | A realidade consiste só de relações, sem <i>relata</i> ?                    | Estruturalismo ontológico radical                                | Substancialismo mínimo           |

**Tabela 1:** O espectro funcionalista é representado na terceira coluna, e seu complementar na quarta coluna (o espectro substancialista).

Fonte: Autor, 2016.

## 10 UMA MÁQUINA PODE SER CONSCIENTE?

Vemos que pergunta sobre se “máquinas podem ter consciência” pode obter diversas respostas. Se definirmos consciência no sentido do comportamentalismo filosófico, então parece muito plausível que atribuiremos consciência a robôs em torno de 2030 (segundo a estimativa de

Kurzweil). Porém, consideramos filosoficamente mais útil uma definição mentalista de consciência, que considera que vivenciamos subjetivamente um estado qualitativo que escapa (num certo sentido) às palavras e que não pode (pelo menos até hoje em dia) ser observado diretamente do ponto de vista externo, objetivo, por outras pessoas.

Da perspectiva mentalista, coloca-se então a questão de como surge a consciência, a partir do corpo (materialismo) ou a partir de duas substâncias separadas (o espiritualismo). Será que a consciência é fruto da organização entre as partes, sendo irrelevante a natureza física (ou espiritual) que compõem essas partes, ou será que a natureza das partes é essencial para a

consciência? A posição de que a natureza da matéria é essencial (opondo-se ao funcionalismo) é uma forma de “substancialismo”.

Vimos que este debate se dá em diversas escalas (Tabela 1), e em cada uma delas há uma resposta diferente à questão de se máquinas podem ter consciência. Para o funcionalismo de estados mentais, bastaria a máquina reproduzir as relações causais entre estados mentais (o chamado “funcionalismo de maquina”). Para o funcionalismo de células biológicas, seria preciso reproduzir as relações causais entre as células do encéfalo para que emergisse a consciência em um robô.

A posição substancialista que se opõe a isso (chamada por Searle de “naturalismo biológico”) supõe que há algo dentro da célula biológica que é essencial para a consciência. Mas se descobríssemos que uma parte da célula é essencial, então as outras partes poderiam ser substituídas por partes sintéticas, como explorada hoje em dia na área chamada de “biologia sintética”. Neste caso, um robô artificial poderia ter consciência se ele carregasse em suas células esta parte material essencial. Esta seria uma solução substancialista.

O funcionalismo, porém, nunca é completamente derrotado. Nessa altura, o funcionalista poderia descer um nível abaixo, e argumentar que aquela parte da célula, essencial para a consciência, é também composta de partes, por exemplo, macromoléculas (proteínas, etc.), e que seria possível substituir essas macromoléculas por elementos artificiais, conservando-se as relações entre as partes. Nesta escala, o funcionalismo subcelular argumentaria que a consciência é fruto apenas da organização das partes do sistema, sendo irrelevante a natureza do substrato material. Neste caso, seria possível uma máquina completamente artificial ter consciência. Mas o substancialista nega que isso seja possível.

## 11 O LIMITE QUALITATIVO DE MODELOS QUANTITATIVOS

Retornemos à questão que abriu este artigo. Argumentamos que a modelagem computacional não consegue capturar a “materialidade” do objeto modelado. Uma simulação eletrônica só consegue capturar as relações *quantitativas* entre as partes do objeto modelado. A única exceção a esta regra é se um modelo computacional simular outro programa de computador: neste caso haverá compartilhamento de materialidade, e poder-se-á falar que o modelo “emula” o seu objeto. Computadores analógicos estendem essa noção de emulação, e a classe de computadores analógicos feitos de circuitos elétricos não só simula circuitos

elétricos em geral, mas também os emulam. Nesse sentido, é razoável considerar que um modelo mental de uma parte da mentalidade de outra pessoa poderia envolver a recriação (emulação) desta mentalidade, e não só uma simulação estrutural. Ao nos envolvermos por empatia com outra pessoa, não só simulamos seus sentimentos, mas o emulamos (até certo ponto).

Concluimos que a simulação computacional de um encéfalo muito provavelmente não “emulará” a consciência, porque o computador de silício careceria das capacidades materiais para gerar a consciência (naturalismo biológico). Porém, quando esta parte material essencial for identificada no tecido biológico, aí sim ginoides poderão ser construídas com consciência, desde que elas contenham esta parte material essencial para a consciência, e não apenas chips de silício.

Nesse sentido, não se poderá fazer um “*upload* da consciência” em um computador digital, como alguns funcionalistas de máquina apregoam, apesar de esta informação poder ser usada para reconstruir um ser consciente no material apropriado (cf. KURZWEIL, 2005, p. 377-86; SANDBERG; BOSTROM, 2008).

Apesar desta possibilidade de construir ginoides com consciência, mesmo assim permanecerá o “problema difícil” da consciência. Uma ginoide com consciência resolve um problema de engenharia, mas por si só não explica teoricamente a emergência da consciência a partir da matéria. Em outras palavras, a mera descrição linguístico-quantitativa (que Mary possuía dentro do quarto) não é suficiente para gerar as qualidades subjetivas (com as quais Mary só entra em contato fora do quarto).

A solução científica será postular leis psicofisiológicas fundamentais como *princípios*, de maneira análoga a como a teoria da relatividade restrita postula a constância da velocidade da luz como princípio. Assim, a pesquisa neurocientífica futura fará um mapeamento de quais estados mentais surgem a partir de quais classes de estados cerebrais, e a sistematização e simplificação deste mapeamento constituirá a teoria das leis psicofisiológicas. Esta teoria poderá ser testada mais facilmente em seres humanos, a partir de experimentos em que cientistas alterem os estados mentais de um paciente a partir de manipulação cerebral, e confiando-se nos relatos verbais dos pacientes. A teoria a ser construída envolveria um conjunto de princípios, cuja combinação geraria todos os estados subjetivos possíveis de um

ser humano. A extensão para outros animais seria bem mais difícil de testar, assim como uma teoria geral da consciência, válida para qualquer ser com consciente primária.

## 12 QUALITATISMO

Para finalizar, é importante indicar qual seria a conexão entre as qualidades subjetivas e a materialidade do mundo físico, dado que ambos se colocam como limites da descrição funcionalista da mente e do mundo (neste último caso, como descrição mecanicista). Tal tema será explorado em outro artigo (PESSOA, 2016), mas vale a pena indicar aqui a conexão.

O monismo neutro antirrealista, defendido em certos momentos das carreiras de David Hume, Ernst Mach, Richard Avenarius, William James e Bertrand Russell, afirma que as qualidades subjetivas são a única realidade que existe. Outra tese distinta, defendida por influência kantiana por William Hamilton e Henri Poincaré, e depois também por Moritz Schlick em 1918 e Russell em 1927, é conhecida como “realismo estrutural” (ou relacional), ou simplesmente “estruturalismo”, e é a tese de que a ciência só captura as relações entre as coisas, e não as coisas em si. A junção dessas duas teses resulta na afirmação de que a única parte da realidade que conhecemos diretamente, por contato (*acquaintance*), são as sensações subjetivas, ao passo que as teses a respeito da realidade externa (que está para além das sensações subjetivas) só são conhecidas por inferência (ver discussão em FEIGL, 1975). Tal afirmação pode ir além do monismo neutro, se reconhecer que há realidade para além dos dados dos sentidos. Tal realidade envolveria coisas que apresentam as relações ou estruturas capturadas por inferência, mas qual seria a natureza dessas coisas externas?

Uma posição que pode ser chamada de “panqualitismo” (*pan-quality-ism*, FEIGL, 1975, p. 326), ou simplesmente “qualitismo”, defende que a realidade externa é constituída, pelo menos em parte, de qualidades semelhantes àquelas que vivenciamos subjetivamente. Se isso for verdade, então pode-se tentar explicar (parcialmente) a consciência como tendo surgido a partir da existência prévia de tais qualidades, que a partir de um certo nível de organização resultariam em qualidades conscientes. Esta concepção é chamada de “panprotopsiquismo” (CHALMERS, 2015), e esteve presente em autores como William Clifford e Durant Drake. Outro termo usado para exprimir a concepção de que o mundo consistiria de coisas-em-si “inescrutáveis”, que no cérebro gerariam a consciência, é “monismo russeliano” (ALTER; NAKASHIMA, 2012). Se tais inescrutáveis, fora do cérebro, já possuíssem alguma

propriedade propriamente mental, ter-se-ia versões do “pampsiquismo” (o qual não é defendido aqui).

## REFERÊNCIAS

- ABBAGNANO, N. *Dicionário de Filosofia*. Tradução de Alfredo Bosi e Ivone Castilho Benedetti. 5ª ed. revista e ampliada. São Paulo: Martins fontes, 2007.
- ABIB, J.A.D. Psicologia, comportamentalismo e subjetividade. In: Chitolina, C.L.; Pereira, J.A. ; Pinto, R.H. (orgs.), *Mente, cérebro e consciência: um confronto entre filosofia e ciência*. Jundiaí: Paco, p. 279-96, 2015.
- ALTER, Torin et al. What is Russellian monism?. *Journal of Consciousness Studies*, v. 19, n. 910, p. 67-95, 2012.
- BLOCK, N. (org.). *Readings in the philosophy of psychology*, vol. 1. Cambridge: Harvard University Press, p. 171-306, 1980.
- 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, p. 246-76, 2015.
- EMERY, N. J. Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, v. 361, n. 1465, p. 23-43, 2006.
- FEIGL, H. Russell and Schlick: a remarkable agreement on a monistic solution of the mind-body problem. In: Roberts, G.W. (org.). *Bertrand Russell memorial volume*. Oxfordshire: Routledge, p. 321-38, 1975.
- FODOR, J.A. “The mind-body problem”, *Scientific American*. Trad. para o português de S.F. Araujo. v. 244, n. 1, p. 124-32, 148 (janeiro), 1981. Disponível em: <<http://www.fflch.usp.br/df/opessoa/Fodor-Port-4.pdf>>. Acesso em: 6 de mai. 2016.
- HAMILTON, W. *Lectures on metaphysics and logic*. 6ª ed., vol. I. Org. H.L. Mansel & J. Veitch. Edinburgh: Blackwood, 1877. Original de 1837.
- HARNAD, S.; SCHERZER, P. First, scale up to the robotic Turing test, then worry about feeling. *Artificial Intelligence in Medicine*, v. 44, n. 2, p. 83-89, 2008.
- KURZWEIL, R. *The singularity is near*. New York: Penguin, 2005.
- JACKSON, F. Epiphenomenal qualia. *The Philosophical Quarterly*, v. 32, n. 127, p. 127-136, 1982.
- LALANDE, A. *Vocabulário técnico e crítico da filosofia*. 3ª ed. Trad. F.S. Correia, M.E.V. Aguiar, J.E. Torres; M.G. Souza. São Paulo: Martins Fontes, 1999.
- LOCKE, J. *Ensaio sobre o entendimento humano*. 2 vols. Trad. E.A. de Soveral et al. Lisboa: Fundação Calouste Gulbenkian, 1999. Original em inglês, 1890.
- NIDA-RÜMELIN, M. Qualia: the knowledge argument, in: Zalta, E.N. (org.), *Stanford Encyclopedia of Philosophy*, 2009. Disponível em: <<http://plato.stanford.edu/entries/qualiaknowledge/>>. Acesso em: 6 de mai. 2016.
- O’DONAHUE, W.; KITCHENER, R. (orgs.), *Handbook of behaviorism*. San Diego: Academic, 1999.
- OPPY, G.; DOWE, D. The Turing test, in: Zalta, E.N. (org.), *Stanford Encyclopedia of Philosophy*, 2016. Disponível em: <<http://plato.stanford.edu/entries/turing-test/>>. Acesso em: 6 de mai. 2016.

- PESSOA Junior, O. A tese do encéfalo colorido. Em preparação, 2016.
- PUTNAM, H.W. Psychological predicates. In: Capitan, W.H. & Merrill, D.D. (orgs.), *Art, mind and religion*. Pittsburgh: University of Pittsburgh Press, p. 37-48, 1967. Republicado com o título “The nature of mental states”.
- RUNES, D.D. *The dictionary of philosophy*. 4ª ed. New York: Philosophical Library, 1942.
- SANDBERG, A.; BOSTROM, N. *Whole brain emulation: a roadmap*. Technical Report. Future of Humanity Institute, Oxford University, 2008. . Disponível em: < <http://www.fhi.ox.ac.uk/brain-emulation-roadmap-report.pdf> >. Acesso em: 6 de mai. 2016.
- SEARLE, J. *A redescoberta da mente*. São Paulo: Martins Fontes, 1997. Original em inglês, 1992.
- HETTLEWORTH, Sara J. Do animals have insight, and what is insight anyway?. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, v. 66, n. 4, p. 217, 2012.
- SHIELDS, C. The first functionalist. In: Smith, J.-C. (org.), *Historical foundations of cognitive science*. Dordrecht: Kluwer, p. 19-33, 1990.
- STADDON, J.E.R. Theoretical behaviorism. In: O’Donahue, W. & Kitchener, R. (orgs.), *Handbook of behaviorism*. San Diego: Academic, p. 217-41, 1999.
- STEINLE, W. O Realismo estrutural ontológico e o problema das relações sem os relata. *Analytica. Revista de Filosofia*, v. 14, n. 1, p. 29-51, 2013.
- TURING, A.M. Computadores e inteligência. Trad. M. Epstein. In: Epstein, I. (org.), *Cibernética e comunicação*. São Paulo: Cultrix, p. 45-82, 1973. Original em inglês, 1950.
- WITTGENSTEIN, L. *Investigações filosóficas*. Coleção Os Pensadores, 2ª ed. Trad. J.C. Bruni. São Paulo: Abril Cultural, 1979. Original em inglês, 1953.



---

PESSOA JUNIOR, Oswaldo Frota. O limite qualitativo de modelos quantitativos. *Complexitas - Rev. Fil. Tem.*, Belém, v. 1, n. 1, p. 101-117, jan/jun. 2016. Disponível em: <<http://www.periodicos.ufpa.br/index.php/complexitas/article/view/3417>>. Acesso em: 04 jul. 2016.

---