



**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I
CENTRO DE EDUCAÇÃO
DEPARTAMENTO DE FILOSOFIA
CURSO DE LICENCIATURA PLENA EM FILOSOFIA**

MARIZIO JOSÉ BARBOSA JÚNIOR

**IMPLICAÇÕES ACERCA DA MENTE: UM DIÁLOGO ENTRE NATURALIDADE E
ARTIFICIALIDADE**

**CAMPINA GRANDE - PB
2024**

MARIZIO JOSÉ BARBOSA JÚNIOR

**IMPLICAÇÕES ACERCA DA MENTE: UM DIÁLOGO ENTRE NATURALIDADE E
ARTIFICIALIDADE**

Trabalho de Conclusão de Curso (Artigo)
em Filosofia da Universidade Estadual da
Paraíba, como requisito parcial à
obtenção do título de Graduação de
licenciatura plena em Filosofia.

Área de concentração: Filosofia
Analítica, Epistemologia e Cognição.

Orientador: Prof. Dr. Carlos Antônio de Souza.

Coorientador: Prof. Dr. Victor Pereira Gomes.

**CAMPINA GRANDE - PB
2024**

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

B239i Barbosa Júnior, Marizio José.
Implicações acerca da mente: um diálogo entre naturalidade e artificialidade [manuscrito] / Marizio José Barbosa Júnior. - 2024.
26 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Filosofia)
- Universidade Estadual da Paraíba, Centro de Educação,
2024.

"Orientação : Prof. Dr. Carlos Antonio de Souza,
Departamento de Filosofia - CEDUC".

"Coorientação: Prof. Dr. Victor Pereira Gomes,
Departamento de Filosofia - CEDUC".

1. Mente. 2. Funcionalismo. 3. Naturalismo biológico. I.
Título

21. ed. CDD 100

MARIZIO JOSÉ BARBOSA JÚNIOR

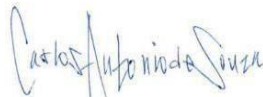
IMPLICAÇÕES ACERCA DA MENTE: UM DIÁLOGO ENTRE NATURALIDADE
E ARTIFICIALIDADE

Trabalho de Conclusão de Curso
(Artigo) em Filosofia da Universidade
Estadual da Paraíba, como requisito
parcial à obtenção do título de
Graduação de licenciatura plena em
Filosofia.

Área de concentração: Filosofia
Analítica, Epistemologia e Cognição.

Aprovada em: 19/11/2024.


BANCA EXAMINADORA




Prof. Dr. Carlos Antônio de Souza
(Orientador)
Universidade Estadual da Paraíba (UEPB)



Prof. Dr. Victor Pereira Gomes (Coorientador)
Universidade Estadual da Paraíba (UEPB)

Documento assinado digitalmente
 JOSE NILTON CONSERVA DE ARRUDA
Data: 21/11/2024 15:11:15-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. José Nilton Conserva de Arruda
Universidade Estadual da Paraíba (UEPB)

Documento assinado digitalmente
 EUGENIA RIBEIRO TELES
Data: 21/11/2024 10:36:14-0300
Verifique em <https://validar.iti.gov.br>

Profa. Dra. Eugênia Ribeiro Teles
Universidade Estadual da Paraíba (UEPB)

Aos meus pais, os quais me deram a vida,
conforto e atenção, DEDICO.

“Do trabalho de tuas mãos comerás,
serás feliz e tudo irá bem.” (BÍBLIA,
Salmos 128:2).

LISTA DE ILUSTRAÇÕES

Figura 1 – Teste de Turing.....	13
Figura 2 – Quarto Chinês.....	18

SUMÁRIO

1	INTRODUÇÃO.....	08
2	DAS DISPOSIÇÕES MENTAIS FUNCIONALISTAS.....	10
2.1	Experimento mental de Turing: o jogo da imitação.....	10
2.2	A mente como manipulação de símbolos	14
3	DO NATURALISMO BIOLÓGICO COMO ASSEGURAÇÃO DA	
	NATURALIDADE MENTAL.....	16
3.1	Experimento mental de Searle: o quarto chinês.....	18
3.2	A mente para além das funções sintáticas.....	20
4	METODOLOGIA.....	21
5	RESULTADOS E	
	DISCUSSÕES.....	22
6	CONCLUSÃO.....	24
	REFERÊNCIAS.....	25
	AGRADECIMENTOS.....	26

IMPLICAÇÕES ACERCA DA MENTE: UM DIÁLOGO ENTRE NATURALIDADE E ARTIFICIALIDADE

IMPLICATIONS ABOUT THE MIND: A DIALOGUE BETWEEN NATURALITY AND ARTIFICIALITY

Marizio José Barbosa Júnior^{1*}

RESUMO

O presente artigo consiste em uma abordagem analítica e investigativa acerca da natureza da mente, de suas problemáticas e da possibilidade de sua categorização artificial. A introdução apresenta a teoria cartesiana sobre a distinção fundamental entre conteúdos mentais e físicos, estabelecendo uma base teórica para a discussão. Em seguida, são analisadas duas teorias que se opõem ao Dualismo de Substâncias de Descartes: o Funcionalismo, que considera a mente como um processo cognitivo sintático, e o Naturalismo Biológico, que vê a mente como uma subjetividade natural pensante. Essas teorias da mente trazem diferentes perspectivas quanto à possibilidade de mentalidade em computadores digitais, levando às questões motivadoras do estudo: em que condições a mente é possível? Ela é exclusivamente um atributo natural dos seres vivos, ou é possível uma mente artificial? A conclusão aponta para a complexidade e a impossibilidade de definir a mente em sua totalidade.

Palavras-Chave: Mente; Funcionalismo; Naturalismo Biológico.

ABSTRACT

This article consists of an analytical and investigative approach about the nature of the mind, its problems and the possibility of its artificial categorization. The introduction presents the Cartesian theory on the fundamental distinction between mental and physical contents, establishing a theoretical basis for the discussion. Then, two theories that oppose Descartes' Dualism of Substances are analyzed: functionalism, which considers the mind as a syntactic cognitive process, and Biological Naturalism, which sees the mind as a thinking natural subjectivity. These theories of mind bring different perspectives on the possibility of mindset in digital computers, leading to the motivating questions of the study: under what conditions is the mind possible? Is it exclusively a natural attribute of living beings, or is an artificial mind possible? The conclusion points to the complexity and impossibility of defining the mind in its entirety.

Keywords: Mind; Functionalism; Biological Naturalism.

^{1*} Graduando em Filosofia (licenciatura plena), pela Universidade Estadual da Paraíba, Campus I - CG. Email: junior.barbosa.758399@gmail.com

1 INTRODUÇÃO

Ao que é percebido pelo decorrer dos fatos históricos, filósofos e cientistas têm apresentado visões distintas quanto à natureza do que é entendido por “mental”. Reale, em sua obra “História da filosofia antiga” (1993), dedica um capítulo extenso à análise da evolução da concepção de *alma* nos pensadores pré-socráticos. Ele demonstra como, desde as primeiras especulações sobre a origem do cosmos, a ideia de *alma* foi sendo gradualmente refinada no decorrer da história. O termo de origem grega *psyché*, que em sua tradução para o português significa “alma”, surge na antiguidade como uma primeira tentativa de definir o que viria a ser “mente” posteriormente. Nos filósofos pré-socráticos, a alma era vista como um princípio vital, uma força que “animava” os seres e os distinguia da matéria “inanimada”. Especialmente no entendimento de Aristóteles, a partir de seus escritos em “Da alma” (2006), o conceito de *alma* transcende a ideia de uma entidade espiritual separada do corpo, e em contraposição, ele nos traz uma distinção de alma com o sentido próprio de vida, sendo ela aquilo o que dá *anima* ao ser. Nesta visão, a alma se exprime em três níveis de categoria, sendo elas: (1) a vegetativa, que está presente nas plantas e é responsável pelas funções básicas da vida, a saber: a nutrição, crescimento e reprodução; (2) a sensitiva, que é encontrada nos animais e vai além das funções vegetativas, visto que inclui a capacidade de percepção, movimento e desejo; (3) a racional como nível mais elevado, que é exclusiva dos seres humanos, pois permite o pensamento abstrato, a linguagem, a moralidade e a consciência de si mesmo. Neste ensejo, a *psyché* racional traz consigo dotes de cunho cognitivo e essa seria a distinção entre os humanos (seres dotados do mais alto nível da cognição) e os outros seres vivos: plantas e animais². Assim, no viés aristotélico, a alma é entendida em seu sentido mais amplo, isto é, pelas “funções nutritivas, perceptivas e intelectuais e pelo movimento” (Aristóteles, 1986, p. 413b). Pelo que é relatado na história até então, não parece ser um problema afirmar que as pretensões de definir a mente surgem de uma característica imaterial. Seja no intuito de definir alma ou espírito, todas as explicações, provenientes deste percurso histórico, parecem distanciar a mente de qualquer tipo de materialidade.

A partir das teses trabalhadas por Descartes nas “Meditações Metafísicas” (1983), no período moderno da história humana, temos uma caracterização que melhor se enquadra à definição do que seria a natureza da mente, de sua distinção e subjetividade, em contraponto aos componentes corpóreos. Nesta via de pensamento, o mundo é entendido pela dualidade das substâncias, “mente” e “corpo”, estas entendidas como substâncias distintas e independentes, que de alguma forma se relacionam e dão consistência ao mundo. Mais especificamente na segunda meditação de sua obra supracitada, ele estabelece a descoberta do *cogito* como assecuração da existência do *eu* pensante, o que enfatiza a sua máxima de que a essência da mente é o próprio pensar, e isto de forma independente de qualquer fisicalidade. Com relação à mente, sua natureza e distinção, Descartes (1983) manifesta um problema que se faz como tema central e polêmico nos

² Há estudos contemporâneos que indicam a possibilidade de existência de algum nível de consciência nos animais. Thomas Nagel, em seu artigo “Que é ser um morcego?” (2013), explora a natureza da consciência subjetiva e a impossibilidade de conhecermos completamente a experiência de outros seres, utilizando como exemplo a experiência de um morcego: através da ecolocalização, o morcego percebe o mundo de uma forma radicalmente diferente da nossa, o que torna sua experiência subjetiva inacessível à nossa intuição. Mas referente ao apontamento desta nota, Nagel não nega a possibilidade de que os animais tenham algum tipo de consciência. Todavia, ele questiona se podemos compreender a natureza dessa consciência em termos humanos.

estudos de Filosofia da Mente e que decaí no questionamento de como a relação entre mente e corpo é possível. Na meditação VI, Descartes aprofunda a sua dualidade entre mente-corpo, constatando que estas são substâncias distintas. Assim, ao passo que a mente deve ser entendida como substância pensante, o corpo deve ser entendido como uma substância extensa, ou seja, material. Após esta definição, Descartes analisa como estas duas substâncias convergem, tendo em vista as suas distinções tão agudas. A resposta cartesiana para tal problema consiste na sugestão de que a glândula pineal, que é localizada no cérebro, pode ser o ponto de interação entre essas duas substâncias. Tal resposta não foi tão convincente e vem sendo muito criticada ao longo da história, o que declama a dificuldade de tal empreitada e uma obscuridade quanto ao problema sobre a explicação de como os conteúdos mentais (*res cogitans*) “conversam” com os conteúdos físicos (*res extensa*).

Diante das divergências nas teorias que buscam explicar a relação entre mente e corpo, torna-se urgente encontrar respostas ou novas perguntas que nos permitam avançar em nosso conhecimento sobre nós mesmos e nossa espécie. Com o avanço da tecnologia e do acesso ao conhecimento, estas inquietações e respostas póstumas a Descartes firmam tal necessidade, pois muitas dúvidas ainda pairam quanto ao que podemos entender por mente, ao ponto de duvidarmos da natureza biológica que a mente intuitivamente demonstra ter. Há de pensarmos, até, se é possível, ou não, um tipo de mentalidade para além da natural.

Em decorrência de tais indagações, este artigo propõe uma análise lógico/crítica/filosófica de alguns argumentos sustentados por duas correntes contemporâneas de estudo da Filosofia da Mente, correntes estas que, em contraste, buscam, por um lado, defender a naturalidade da mente e, por outro, a artificialidade da mesma. Primeiramente, será feita a explanação e análise do Funcionalismo, teoria filosófica da mente que propõe que os estados mentais podem ser definidos em termos de suas funções causais e relações com outros estados mentais. Essa perspectiva, inspirada em modelos computacionais, sugere que a mente pode ser entendida, analogamente, como um sistema de processamento de informações, sem, necessariamente, se limitar aos aspectos cognitivos. Deste modo, a mente humana pode ser entendida como um programa de computador digital (*software*). Em vez de focar na substância física do cérebro (*hardware*), o funcionalismo se concentra nas funções que a mente realiza. Sendo assim, para o funcionalismo, o que define a mente não é o material de que ela é feita, (biológico ou artificial) mas o que ela faz (o seu papel funcional). Em segundo lugar, serão apresentadas e discutidas as respostas de John Rogers Searle sobre a distinção entre mente e software, que o Funcionalismo exprime. O Naturalismo Biológico, teoria da mente defendida por este autor, propõe que a mente é um fenômeno biológico emergente do cérebro. Essa teoria sugere que, embora a inteligência artificial (atribuída aos computadores digitais) possa simular muitos aspectos da cognição humana (lógico-sintáticos), a criação de uma inteligência artificial com consciência e experiência subjetiva semelhante à nossa é um desafio complexo e, segundo Searle, insuperável. Por último, usando como plano de fundo o embate travado entre essas duas teorias (Naturalismo Biológico x Funcionalismo), serão investigadas algumas questões que podem surgir a respeito da mente, a saber: 1) o que a mente exprime? 2) Ela existe? 3) Em que condições ela existe? 4) Ela é natural e inerente aos seres vivos? Com efeito, ao confrontar essas teorias, será pretendido contribuir para o debate sobre a natureza da mente e suas implicações para nossa compreensão de nós mesmos.

2 DAS DISPOSIÇÕES MENTAIS FUNCIONALISTAS

O modelo funcionalista surge em oposição ao dualismo cartesiano, buscando explicar a mente em termos de funções e processos. Embora algumas vertentes do funcionalismo possam parecer negligenciar o sentido fenomênico/subjetivo da mente, outras tentam explicá-la em termos de estados funcionais. Esta teoria defende que os estados mentais fenomênicos, sejam eles: dor, crenças, desejos, etc... Não são idênticos a estados físicos do cérebro (sua composição estrutural orgânica), mas sim, a estados funcionais. Neste ensejo, o que define um estado mental não é a sua composição física, mas o papel de ação que ele desempenha em um sistema. Diferentemente do dualismo cartesiano (1983) anteriormente citado, que postula uma separação radical entre mente e corpo, o funcionalismo, em termos gerais, concentra-se nos aspectos funcionais da mente, tratando-a como uma espécie de processo de computação simbólica. Ao focar na função, o funcionalismo busca compreender os estados mentais em termos de suas relações causais e suas entradas (inputs) e saídas (outputs), sem demonstrar uma expressa preocupação com sua natureza fenomênica e subjetiva que será trabalhada posteriormente neste artigo.

Existem vertentes distintas do Funcionalismo no tocante à mente, a saber: (1) o Funcionalismo que vai tratar da mente em aspectos psicológicos ao focar na relação entre os estados mentais e os processos psicológicos, e sendo assim, a mente é entendida como um sistema de processos que produzem comportamento; (2) o funcionalismo em linhas analíticas, onde busca analisar os estados mentais em termos de suas relações causais com outros estados mentais e comportamento; (3) o funcionalismo em termos de estado de máquinas inteligentes, no qual comparado a um programa de computador, a mente pode ser vista como um conjunto de estados funcionais que podem ser realizados por diferentes substratos físicos. Levando em consideração a proposta deste artigo, que se afirma especificamente na comparação dos aspectos naturais da mente aos possíveis aspectos artificiais que ela possa ter, será seguido o desenvolvimento das teorias funcionais que dão corpo a essa discussão, sendo elas, as que se encaixam no ponto 3.

Ao iniciarmos a explanação basilar dos objetivos que esta teoria defende, precisamos primeiro definir parâmetros. Ao falarmos de máquinas e de seu aspecto artificial, fugiremos da conceituação geral de máquina (que pode caracterizar um maquinário qualquer) e entenderemos pelo conceito de “máquinas digitais”. Ou seja, máquinas com dotes de múltiplas ações, capacidade de reprodução e assimilação sintática: computadores.

2.1 Experimento mental de Turing: o jogo da imitação

Como supramencionado, o Funcionalismo, em busca de fundamentar suas premissas, propõe modelos que equiparam a mente natural a um sistema artificial (software). Mas a partir de qual fonte isso foi possível? Um exemplo seminal, que veio a moldar e dar possibilidade à concepção de sistemas computacionais modernos, é a *Máquina de Turing*, inventada pelo mesmo em 1936. Essa máquina teórica representa um modelo computacional que se consolidou como universal. Mas no que consiste tal máquina? Gomes bem resume:

Uma máquina de Turing (a-machine, máquina automática) é um mecanismo teórico composto por um cabeçote de leitura/escrita, e por uma fita infinita em sentido horizontal, tanto para a direita quanto para a esquerda, dividida

em quadrados de tamanhos iguais, que são examinados individualmente pelo cabeçote[...] (GOMES, 2018, p. 78).

A grosso modo, ao que podemos brevemente entender sobre o seu funcionamento, ela é uma máquina que opera seguindo regras simples que determinam como o cabeçote se move, o que é escrito na fita e para qual estado a máquina transita. Apesar de sua limitação devido à época, a *Máquina de Turing* é capaz de realizar qualquer cálculo que possa ser realizado por um computador moderno, configurando o seu status de máquina *standard* (padrão). Mas por qual razão a *Máquina de Turing* se configura como padrão? Podem ser destacadas duas características que firmam este cunho. A primeira delas, é o aspecto de “universalidade”: *Máquinas de Turing* são capazes de simular qualquer cálculo que possa ser realizado por qualquer computador digital. Isso implica que, em teoria, qualquer problema que possa ser resolvido por um computador pode ser resolvido por uma *Máquina de Turing*. A segunda característica se afirma no seu fundamento teórico: *Máquinas de Turing* fornecem uma base sólida para a teoria da computabilidade. Assim, ela nos permite delimitar o que é computável e o que não é, e nesta esteira, estabelece limites para o que um computador pode, ou não, fazer. Vale ressaltar que a *Máquina de Turing* é um conceito teórico, um modelo matemático idealizado para entender os limites da computação. Ela não é um computador físico, mas é o que dá margem para o desenvolvimento dos computadores atuais. A distinção que será feita posteriormente entre hardware e software será aplicada aos modelos de computadores modernos, e não diretamente a uma *Máquina de Turing*, mas ela servirá de alicerce inicial para a discussão.

Voltando ao escopo da artificialidade e naturalidade (motivação deste presente artigo), em que sentido este tipo de máquina pode se assimilar a um cérebro natural? Leclerc (2018) em “Introdução à Filosofia da Mente” brevemente nos explica:

Nosso cérebro recebe informações das mais diversificadas pelos órgãos sensoriais (sensações visuais, táteis, auditivas, propriocepções, etc.), mas tudo é “traduzido”, por assim dizer, em impulsos nervosos, sinapses, neurônios que disparam ou não disparam. Tudo se reduz a isso na verdade: neurônios disparam ou não disparam, 1 ou 0. [...] Nessa atividade eletroquímica, os vários estados neurais pelos quais passa o cérebro determinam o comportamento que o organismo irá adotar. Também, o que a máquina de Turing “lê” num quadro da fita infinita pode ser comparado a um estímulo sensorial, ao input; o resultado final da operação processada é comparável a uma ação ou atividade realizada (comportamento). Finalmente, as instruções da tabela determinam os estados internos da máquina. O que faz de um estado mental o estado que é, para o funcionalista, é seu papel causal relativo ao input (ativação de algum órgão sensorial), ao output (o comportamento), e aos outros estados mentais; da mesma maneira, o que torna um estado físico (do cérebro ou de um dispositivo eletrônico) uma realização física de uma determinada máquina de Turing são as relações causais que este estado físico mantém com o input do sistema físico, o seu output, e os outros estados internos do sistema (Leclerc, 2018, p.82).

Nesta passagem, o autor destaca a redução neural operada pelo funcionalismo. Sendo assim, a experiência consciente, por mais rica e diversificada que possa ser, é reduzida a processos neurais básicos, que são a ativação, ou não, de neurônios. Assim como a *Máquina de Turing* processa informações através de combinações de 0s e 1s, o nosso cérebro processa informações através de padrões

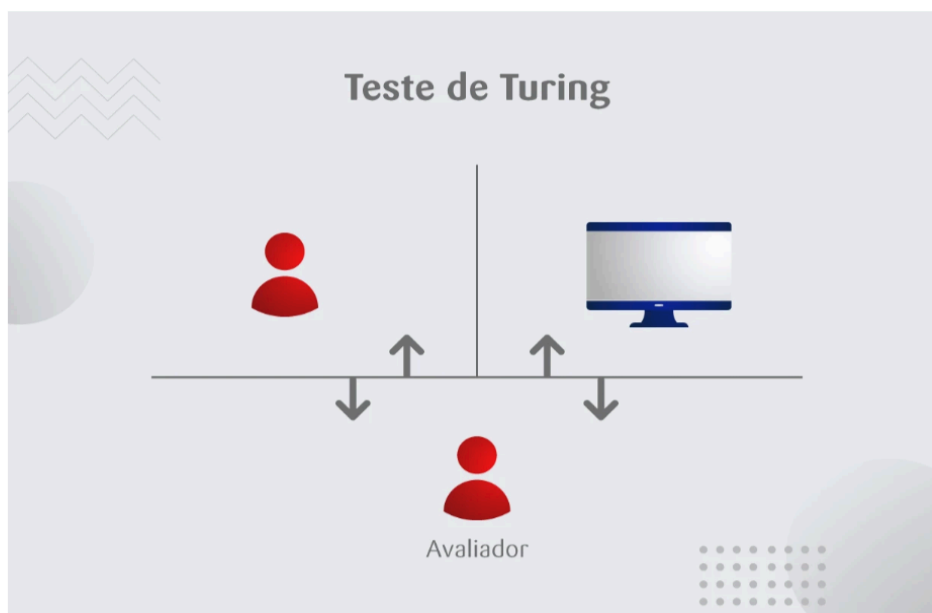
de neurônios que disparam, ou não. Assim, o que define um estado mental não é a sua natureza física de seus substratos, mas sim o papel funcional que desempenha no sistema neural, ou seja, como se relaciona com as entradas sensoriais, as saídas comportamentais e outros estados mentais.

No artigo intitulado "*Computing Machinery and Intelligence*" que foi publicado em 1950, Alan Turing instiga a célebre questão: "Podem as máquinas pensar?", contudo, ao entender que o processo de pensar possui uma definição controversa, ele reformula a questão inicial para a seguinte: "há como imaginar um computador digital que faria bem o jogo da imitação?", mas no que consiste esse jogo? Como isto seria possível?

Originalmente, o *jogo da imitação* de Turing possui duas etapas com intencionalidades de conclusão similares. A primeira, consiste em um *Quiz* de perguntas e respostas no seguinte cenário: imaginemos 3 salas separadas, sem nenhum contato visual ou audível entre elas, com o único meio de comunicação que é através de mensagens digitadas. Em cada uma delas existe um participante, sendo eles: (1) jogador A: um homem, (2) jogador B: uma mulher e (3) jogador C: um interrogador, de qualquer sexo. O jogo tem por objetivo principal que o jogador C, o interrogador, através de uma série de perguntas, determine qual dos outros dois jogadores (A, B) é o homem e qual é a mulher. Pela própria definição do jogo, o foco se afirma na imitação, e sendo assim, o jogador A, o homem, tenta imitar a jogadora mulher de forma a enganar o interrogador, e a mulher, por sua vez, responde verdadeiramente. Esta primeira etapa evoca as questões: considerando todo cenário no qual o jogo aconteceu, qual a taxa de probabilidade de acerto que o interrogador teria? Haveria como distinguir acertadamente sobre isso? Quem é o homem? Quem é a mulher?

Na segunda etapa, que consiste no mesmo cenário da etapa anterior, mas vamos direcionar o teste à *Máquina de Turing* a fim de responder à questão inicial proposta no artigo mencionado. Portanto, agora temos como participantes: jogador A: um computador digital, jogador B: um homem e jogador C: um interlocutor humano. O teste tem por finalidade última constatar que o computador digital possui a capacidade de responder questões de forma humana, sem que o interlocutor consiga distinguir quem é o humano ou quem é a máquina. O jogo segue no mesmo padrão da etapa anterior, só que agora, a máquina é quem vai tentar negar a sua identidade, enquanto o homem vai tentar afirmar a sua. O teste prossegue para a fase que consiste nas interações por texto: os participantes A, B e C entram em uma conversação exclusivamente por mensagens de texto em linguagem natural, tudo isto para isentar de qualquer identificação no tocante à aparência ou timbre vocal, e assim, dando veracidade ao teste. Por fim, adentramos na avaliação do teste: na condição do interrogador não conseguir distinguir quem está falando a verdade sobre a sua identidade, acarreta que o computador passou no teste. Assim, constatando que o computador, de fato, conseguiu simular a inteligência humana de forma convincente. Para uma melhor compreensão, o teste pode ser assimilado pelo fluxograma a seguir:

Figura 1 – Teste de Turing.



Fonte: <https://www.caiena.net/blog/teste-turing-o-que-e-historia-desdobramentos> (Acesso em 08/11/2024).

O *Jogo da Imitação* de Turing destaca a importância da imitação e da linguagem na avaliação da inteligência, e isso nos diz que podemos entendê-la dessa forma, haja vista as suas similitudes. Embora o foco inicial do teste tenha sido a imitação de gênero (quem é o homem? Quem é a mulher?) e, secundamente, o de espécie (Quem é o humano? Quem é a máquina?), o conceito se mostrou extremamente relevante para o desenvolvimento e comprovação do que entendemos por uma ação inteligente e serve de combustível para a definição de uma *inteligência artificial* mais avançada. Se um interrogador humano não conseguir distinguir as respostas de uma máquina daquelas de um humano em uma conversa textual, então a máquina pode ser considerada como tendo passado no teste. Isto levanta a hipótese das *Máquinas de Turing* serem inteligentes³, pois se elas conseguiram enganar o interlocutor, elas demonstram, ao menos, que respondem como um humano. Nos perguntamos, então: algo que seja capaz de realizar uma ação inteligente é inteligente? E mais, algo que simule a ação de algo, torna-se algo? O teste mental de Turing abre margem para muitas indagações. Ao tratar do tipo de inteligência atribuída aos computadores mais avançados, que tiveram por base o modelo de Turing, é notório citar as algumas capacidades: (1) aprender com dados: podem ser treinadas com grandes quantidades de dados, a partir dos quais elas identificam padrões e aprendem a realizar tarefas específicas; (2) resolver problemas: podem analisar informações e encontrar soluções para problemas

³ Se faz necessário ressaltar que o *teste mental de Turing* não define o que é inteligência, mas sim, estabelece um padrão para avaliar a capacidade de uma máquina de simular comportamento inteligente. Essa proposta de Turing gerou debates intensos sobre a natureza da inteligência e a possibilidade de criar máquinas pensantes. Embora o Funcionalismo, a teoria da Filosofia da Mente anteriormente explicitada, tenha sido influenciada pelas ideias de Turing, o *Teste de Turing* não é, em si mesmo, uma definição do Funcionalismo, mas está sendo usado como um parâmetro comparativo ao teste mental que se seguirá mais à frente no presente artigo.

complexos, de forma direcionada ou autônoma; (3) tomar decisões: Com base nas informações que possuem e nos algoritmos que as guiam, as máquinas inteligentes podem tomar decisões, como escolher a melhor rota em um mapa ou identificar um objeto em uma imagem; (4) Se adaptar a novas situações: Elas podem ajustar seu comportamento e suas respostas com base nas experiências que elas adquirem; (5) a capacidade de interação com os humanos, que foram demonstradas à partir do teste em questão. Estas são características possíveis de, intuitivamente, atribuir à inteligência humana. Mas até que ponto?

2.2 A mente como manipulação de símbolos

Leclerc, na obra “Uma Introdução à Filosofia da Mente” (2018), livro que aborda as problemáticas referentes à mente e de como elas se desenvolveram no decorrer da história, ao tratar especialmente das teses funcionalistas, nos resume a principal pretensão delas sobre a mente:

A identidade de um estado mental não depende do tipo de material (se é biológico, à base de carbono e proteínas, ou à base de silício, ou qualquer outra matéria orgânica, ou se é uma estrutura de circuitos eletrônicos, como nos computadores atuais, ou ainda uma estrutura mecânica, como nas antigas calculadoras). A identidade dos estados mentais depende de relações causais entre as experiências sensoriais (a entrada – input), de relações causais entre estados internos do organismo, e o comportamento causado (a saída – output) (Leclerc, p. 76-77, 2018).

Ou seja, nossos pensamentos e ações são resultado de um complexo processo de processamento de informações. Quando recebemos uma informação do mundo exterior (input), nosso cérebro a interpreta, relaciona com conhecimentos prévios e gera uma resposta (output). Essa resposta pode ser um pensamento, uma emoção, uma ação ou uma combinação destes fatores. É como se o cérebro fosse um computador biológico, constantemente recebendo dados, processando e produzindo o que resulta deles. Não há, portanto, um vínculo necessário entre um cérebro complexo e a mentalidade (enquanto aspecto subjetivo), de modo que o essencial vem a residir no papel funcional dos fenômenos mentais, ou seja, nas relações causais que determinado fenômeno mental possui. Portanto, seguindo o ideal funcionalista, a mente está para o cérebro assim como o software está para o hardware. Mas em que sentido essa assimilação se faz como sólida?

Um dos principais defensores desta doutrina de pensamento é o Hilary Putnam⁴ com a defesa do argumento de realizabilidade múltipla exposto explicitamente em seus artigos: “A natureza dos estados mentais” (1967a, título traduzido pelo autor) artigo seminal onde ele introduz a ideia de que os estados mentais não são idênticos a estados físicos específicos, mas sim a estados funcionais; “Cérebros e comportamentos” (1967b, título traduzido pelo autor) artigo onde Putnam aprofunda sua defesa do funcionalismo computacional, utilizando o argumento da realizabilidade múltipla para sustentar a ideia de que os estados mentais podem ser reproduzidos por computadores.

⁴ Em sua trajetória intelectual, Putnam inicialmente (1960) mostrou cautela ao comparar estados mentais e estados de máquinas de Turing, devido à possível abertura causal da mente. No entanto, em trabalhos posteriores (1967a e 1967b), ele evoluiu para uma posição mais funcionalista, sugerindo uma correspondência entre os estados mentais e os estados de uma máquina de Turing, embora reconhecendo a complexidade intrínseca da mente. Em sua obra final sobre o tema (1967b), ele adotou plenamente o funcionalismo computacional.

O fundamento para tais afirmações consiste na observação da diversidade da vida na Terra. Tal diversidade demonstra que organismos distintos podem desempenhar funções idênticas. A fotossíntese, por exemplo, é realizada por plantas e bactérias, embora seus mecanismos celulares sejam variados. Decorre que diferentes tipos de plantas realizam o mesmo processo da fotossíntese, apesar de possuírem a sua diferença categorial. A partir deste entendimento, para Putnam, nós podemos aplicar este mesmo princípio aos estados mentais. Exemplifiquemos a defesa desta afirmação com a dor física: Nós sentimos dor, temos o breve entendimento de que a dor não é uma coisa física que podemos tocar ou ver. Devido a isto, a dor configura um estado mental, que desemboca na sensação de dor. Seguindo o argumento de Putnam, esse estado mental (a dor) pode ser causado, ou expressado, por diferentes processos físicos em diferentes organismos. Se imaginarmos a possibilidade de um alienígena poder experimentar dor, apesar de suas diferenças biológicas com relação aos humanos, isto indicaria que a dor não é exclusiva de uma determinada estrutura física, que no caso, seria de um sistema nervoso ligado a um cérebro que é constituído de massa cinzenta. Nesta esteira, a função da dor - sinalizar danos e orientar o comportamento - é o que a define, independentemente do substrato biológico. Ocorre que, segundo este argumento, é possível que ambos experimentem a mesma sensação subjetiva de dor, mesmo que os processos neurológicos que causam essa dor sejam completamente diferentes em cada um deles, a composição material não importa. Com este argumento, Putnam desafia a ideia de que cada estado mental é idêntico a um estado físico específico, pois se a mesma dor pode ser realizada por diferentes estados físicos, então não há uma correlação direta e necessária entre um estado mental e um estado físico particular.

Algo que é possível assimilar aos computadores no momento em que entendemos os fenômenos enquanto essa simplificação sintática. Assim, a “dor” nada mais é do que um aspecto funcional, uma excitação na fibra nervosa, e ao compararmos com um computador, muito pode se assimilar a ação dele ao detectar um erro de funcionamento: o erro foi computado de forma *input* e exprimido de forma *output*. Por assimilarmos esta ação computacional de acusar um erro de funcionamento, não parece ser importante adentrarmos em qualquer possibilidade de aspecto fenomênico à ação de sentir dor, já que podemos a entender pelo que ela exprime. Sendo assim, o “argumento da realizabilidade múltipla” sugere que a inteligência não está limitada a substratos fenomênicos, fato este que abre caminho para a possibilidade da inteligência artificial enquanto subjetividade pensante no momento em que compreendemos a subjetividade no seu aspecto funcional. Considerando que a mente é definida por suas funções e não por sua realização física, então, em princípio, uma máquina poderia realizar as mesmas funções mentais que um ser humano.

Ao sintetizarmos as máximas funcionalistas, nós obtemos que a identidade de um estado mental não está vinculada a um substrato material específico. Em vez disso, é determinada por seu papel funcional dentro de um sistema. Assim sendo, as relações causais entre as entradas sensoriais (*input*), os estados internos e as saídas comportamentais (*output*) são o que define a natureza de um estado mental, independentemente de sua realização física em um cérebro biológico, um circuito eletrônico ou qualquer outro sistema.

3 DO NATURALISMO BIOLÓGICO COMO ASSEGURAÇÃO DA NATURALIDADE MENTAL

Partindo das postulações da teoria do Naturalismo Biológico, a analogia entre mente e software, ao enfatizar os aspectos computacionais, tende a desconsiderar a dimensão qualitativa da experiência consciente, reduzindo-a a um conjunto de regras e algoritmos. Seguindo as condições que as teses funcionalistas propuseram, onde os fenômenos mentais não seriam essencialmente naturais (já que eles podem ser reproduzidos artificialmente por um computador), acabamos por desconsiderar o fator semântico e a noção de *qualia* que estão fortemente atrelados à mente.

De acordo com John Searle, os *qualia*, ou experiências subjetivas e qualitativas, representam um desafio fundamental ao funcionalismo. Mas em que condição isso se exprime? Enquanto os funcionalistas buscam explicar os estados mentais em termos de funções causais, Searle argumenta que a consciência envolve uma dimensão qualitativa irreduzível. Os computadores, neste caso, por mais sofisticados que eles possam ser, podem simular qualquer processo computacional em seu sentido sintático-lógico, mas *nunca* poderão ter a experiência subjetiva de ver a cor vermelha ou ter a sensação de sentir dor, por exemplo. Essa lacuna entre a sintaxe (estrutura formal) e a semântica (significado) da linguagem, segundo Searle, é intransponível por qualquer sistema meramente formal. Ainda na visão do autor mencionado, a semântica está relacionada diretamente ao significado das palavras e das sentenças. Por este ensejo, ela é a parte da linguagem que nos permite entender o que as palavras significam e como elas se relacionam com o mundo. Assim, podemos resumi-la ao conteúdo próprio da linguagem.

Ao introduzir o conceito de *qualia*, John Searle desafia a tese de que a mente possa ser reduzida a um programa de computador, ou entendida como um. Para ele, a consciência envolve mais do que a *mera* manipulação de símbolos de acordo com regras sintáticas. A experiência qualitativa, ou *qualia*, é uma característica intrínseca da mente que não pode ser capturada por nenhum programa. Assim, a semântica, ou o significado, que está ligado à intencionalidade⁵ dos estados mentais, é outro aspecto fundamental da mente que Searle acusa de escapar à explicação funcionalista.

Em “A redescoberta da mente” (1997) John Searle demonstra a inadequação de explicações que reduzem a consciência a aspectos externos, como comportamento ou processos internos como estados cerebrais. Para Searle, a consciência é uma experiência interna, singular e qualitativa, inerente à primeira pessoa. Ao ignorar essa dimensão subjetiva, na visão do autor, qualquer teoria sobre a consciência acaba por explicar algo diferente dela mesma. O argumento de Searle, que veremos a seguir, é particularmente persuasivo quando ele demonstra que a consciência não pode ser reduzida a um fenômeno mais fundamental, pois ela possui características únicas que a distinguem de outros fenômenos. Vejamos:

Parte da questão da redução no caso do calor era distinguir entre o aparecimento subjetivo deste por um lado e a realidade física subjacente no outro. Realmente, é uma característica geral de tais reduções o fenômeno ser definido em termos da 'realidade' e não em termos do 'aparecimento'.

⁵ O sentido de “intencionalidade” se exprime de forma distinta de filósofo para filósofo, mas o que John Searle trata, se define como o dote fundamental da mente de estabelecer relações com o mundo. Neste ensejo, a intencionalidade é a característica que permite que nossos pensamentos, crenças e desejos (dotes subjetivos/fenomênicos) sejam “sobre” algo, de modo que estejam direcionados a alguma coisa no mundo.

Mas nós não podemos fazer este tipo de distinção de aparecimento-realidade para consciência porque a consciência consiste nestes próprios aparecimentos. Quando o aparecimento é a questão não podemos fazer a distinção aparecimento-realidade porque o aparecimento é a realidade (Searle, 1992, p.121-122).

A citação de Searle indica um ponto crucial na filosofia da mente, especialmente no debate sobre o reducionismo e a natureza da consciência. Ele inicia a citação destacando um problema comum em abordagens que, de alguma forma, buscam reduzir, ou desconsiderar a mente ou seus estados: a tendência de separar o fenômeno como ele se “apresenta” da realidade física subjacente. Em outras palavras, o reducionismo busca explicar a experiência subjetiva em termos de processos físicos mais básicos, quase como se a experiência fosse uma “ilusão” criada por esses processos. Em contraposição, Searle argumenta que a consciência é diferente, existe algo à mais nela. Ao contrário de outros fenômenos, a consciência é o “aparecimento”. A experiência consciente não é algo que ocorre “por trás” de um processo físico, mas sim, é ela mesma a “realidade” que estamos tentando explicar. Para uma melhor assimilação, resta a exemplificação das cores: imagine a possibilidade de estar tentando explicar a cor vermelha para alguém que nunca viu essa cor (uma pessoa que nasceu privada da visão). Ao tentar explicar, pressupondo que tenha a noção sobre características da cor, você poderia descrever o comprimento de onda da luz, as reações químicas no olho, etc... Mas, no fim, você nunca conseguiria realmente transmitir a experiência de ver o vermelho. Ocorre que, a cor vermelha é mais do que apenas um conjunto de propriedades físicas, e assim, Searle defende que ela é uma experiência subjetiva única. Seguindo a sua linha de argumentação, assim como no caso da cor vermelha, a “consciência” é uma experiência fundamental que não pode ser completamente reduzida a processos físicos, como assim defende o “argumento da realizabilidade múltipla” de Putnam (1967ab) supracitado. Deste modo, tentar explicar a “consciência” apenas em termos de neurônios e sinapses seria como tentar explicar a beleza de uma pintura apenas descrevendo a composição das tintas.

Pelos computadores demonstrarem carência deste fator de consciência, a mente, nesta via de conceituação, não se demonstra como possível à elas. Assim sendo, o Naturalismo Biológico evoca a tese de que uma mente artificial não é possível, pois o que vem a constituir a mente são justamente as suas características fenomênicas, naturais, semânticas, a consciência autônoma (particular) e as suas propriedades de primeira pessoa. Antes de definir a sua crítica às teorias funcionais, Searle primeiro caracteriza a existência da IA fraca e da forte:

Que significado psicológico e filosófico devemos atribuir aos esforços feitos recentemente para simular capacidades cognitivas humanas através do computador? Para responder esta questão, considero útil distinguir entre o que denomino de IA no sentido “forte” e IA no sentido “fraco” ou “cautelosa”. De acordo com a IA no sentido fraco, o principal valor do computador para o estudo da mente reside no fato de que este nos fornece uma ferramenta extremamente poderosa. Por exemplo, ele nos permite formular e testar hipóteses de maneira mais rigorosa e precisa do que antes. Mas de acordo com a IA no sentido forte, o computador não é meramente um instrumento para o estudo da mente. Muito mais do que isso, o computador adequadamente programado é uma mente, no sentido de que, se lhe são dados os programas corretos, pode-se dizer que eles entendem e que eles têm outros estados cognitivos. Conforme a IA no sentido forte, uma vez que o computador programado tem estados cognitivos, os programas não são meros instrumentos que nos capacitam a testar explicações psicológicas: os

programas constituem as próprias explicações. Não tenho objeções à IA no sentido fraco, pelo menos no que diz respeito ao escopo deste artigo. Minha discussão será dirigida às afirmações que defini como caracterizando a IA no sentido forte, especificamente a ideia de que computadores adequadamente programados têm estados cognitivos e que os programas, a partir disso, explicam a capacidade cognitiva humana. Quando eu me referir à IA, estarei considerando a IA no sentido forte, definida através das duas afirmações acima (Searle, 2021, p.281).

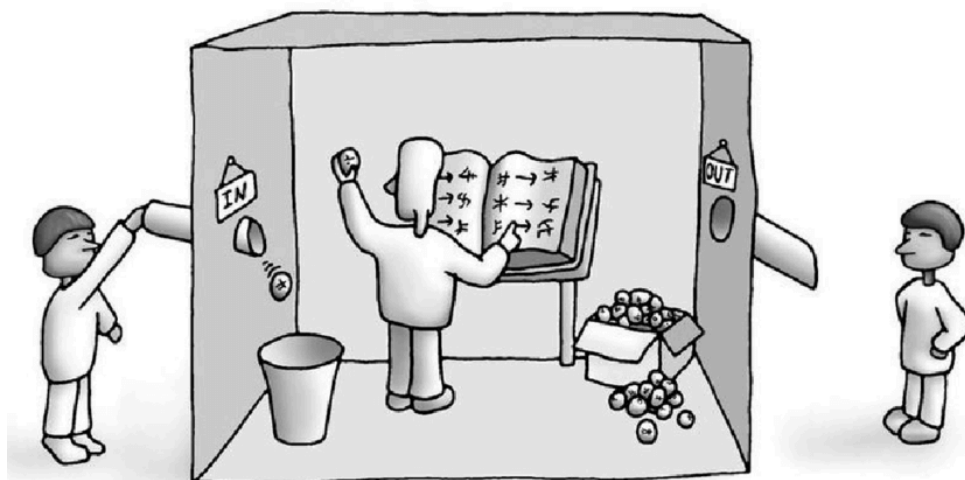
A partir desta citação, Searle apresenta uma distinção fundamental na abordagem da inteligência artificial: a IA forte e a IA fraca. Essa dicotomia se fará como central para o debate proposto por este trabalho sobre a natureza da mente e a possibilidade de criar máquinas com consciência. A defesa da IA forte postula que um computador adequadamente programado poderia ter estados mentais genuínos, como compreensão e intencionalidade. Nesta via, os programas não seriam apenas ferramentas para explicar a mente, mas eles constituem a própria explicação da mente. A partir disso, a mente poderia ser reduzida a processos computacionais.

Seguido desta distinção dos níveis de IA proposta por John Searle, podemos perfeitamente direcionar a sua crítica ao Teste de Turing e da comparação entre mente-cérebro = software-hardware mencionados anteriormente.

3.1 Experimento mental de Searle: o quarto chinês.

De um modo geral, a crítica que John Searle direciona às teses Funcionalistas da mente se efetiva no que ele denomina de “argumento do quarto chinês” e que se exprime como contundente ao tentar nos demonstrar os limites das ações dos computadores e de suas impossibilidades substanciais de entendimento. O “argumento do quarto chinês”, publicado em seu artigo “*Minds, Brains and Programs*” (1980), que traduzido para a língua portuguesa significa: “Mentes, Cérebros e Programas”, possui a máxima geral de que *nenhum* tipo de mentalidade é possível através de *mera* manipulação de símbolos. Referente ao *experimento* mental, imaginemos a seguinte situação: Uma pessoa sozinha em uma sala fechada, com conhecimento apenas de sua língua materna e com o acesso à área externa por uma entrada (*input*) e uma saída (*output*). Dentro da dita sala, o personagem está de posse de um manual de instruções, na sua língua materna, sobre como manipular ideogramas chineses, de modo a compatibilizá-los mecanicamente. Por conseguinte, de forma de entrada (*input*), chega uma informação escrita no idioma chinês, e o personagem, por fazer uso do manual de instruções, manipula a informação e entrega essa informação manipulada no idioma chinês pela saída (*output*). O *experimento* pode ser melhor assimilado pela ilustração a seguir:

Figura 2 – Quarto chinês.



Fonte: <https://www.researchgate.net/figure/Figura-1-Experimento-Quarto-chines-de-John-Searle-Filosofia-na-mente-sd-fig20-332775462> (Acesso em 08/11/2024).

Tal exemplo nos instiga a indagação: o personagem do *experimento* mental do quarto chinês obteve conhecimento do idioma chinês ao manipular os ideogramas em Chinês? Tal manipulação de símbolos operada configura um entendimento sobre a ação feita? Searle indica que este não é o caso. Ao levarmos em conta o aspecto natural da mente, o processo no qual o personagem operou foi puramente sintático. Ocorre que, a sua manipulação feita pelo uso do manual nada mais foi do que uma manipulação de símbolos, desprovida de qualquer entendimento, de fato, do novo idioma. Este simples exemplo expõe a debilidade do argumento funcionalista no tocante à sua comparação entre mente-cérebro tal como software-hardware, pois a pura operação sintática *não* diz nada quanto ao entendimento da própria operação. Se considerarmos que o personagem não obteve entendimento do idioma chinês ao fazer uso do do manual de instruções de como manipular ideograma chineses, o mesmo ocorre com as informações que o computador manipula. Assim, no entendimento de Searle, uma ação sintática carece de semântica (sentido). Partindo da constatação desse exemplo, Searle nos pontua implicações lógicas dotadas de premissas e conclusões:

Premissas:

- (P1): O cérebro causa a mente;⁶
- (P2): A sintaxe não é suficiente para a semântica;
- (P3): Programas de computador são totalmente definidos por sua estrutura formal ou sintática;
- (P4): A mente possui conteúdos semânticos.

Conclusões (Considerando as premissas anteriores como verdadeiras):

(C1) Em decorrência de: P2,3,4 = Nenhum programa é, por si mesmo, suficiente para produzir uma mente; além de suas características distintas de

⁶ Premissa dos avanços provenientes da neurociência que visam reduzir os estados mentais em estados causais.

qualquer tipo de mentalidade, um programa é insuficiente, por si próprio, para gerar uma mentalidade. (Fazer comparação entre entendimento e replicação).

(C2) Em decorrência: P1, C1 = O modo como o funcionamento cerebral causa a mente não pode ser unicamente em decorrência da implementação de um programa de computador.

(C3) Em decorrência: P1 = Qualquer coisa que cause uma mente tem que ter capacidades causais pelo menos equivalentes às do cérebro.

(C4) Em decorrência: C1, 3 = A implementação de um programa de computador por si mesma não seria capaz de dotar uma máquina de estados mentais semelhantes aos nossos, já que a máquina teria que ter capacidades causais equivalentes às do nosso cérebro.

É obtido que as premissas e conclusões apresentadas defendem a ideia de que a mente humana é um fenômeno complexo e emergente, que não pode ser reduzido a simples processos computacionais. A criação de uma IA verdadeiramente consciente continua sendo um desafio que exige uma compreensão mais profunda da natureza da mente e do cérebro.

3.2 A mente para além das funções sintáticas.

Em “A redescoberta da mente” (1997), John Searle justifica a sua crítica sobre a insuficiência das respostas das relações causais que o Funcionalismo postula. Ao passo que são negadas as ações fenomênicas, as teorias funcionais da mente, que buscam esse viés, não conseguem abarcar a relação existente entre estrutura-função. O que implica que a descrição das relações causais defendidas por essa teoria funcional não são suficientes para uma explicação da complexidade da mente. Vejamos o argumento na íntegra:

[...] suponhamos que uma parte da população tivesse seus espectros de cores invertidos de tal maneira que, por exemplo, a experiência que eles designam por “ver vermelho” fosse chamada de “ver verde” por uma pessoa normal; e o que designam por “ver verde” fosse chamada de “ver vermelho” por uma pessoa normal (Block e Fodor, 1972). Ora, podemos supor que esta “inversão de espectro” seja inteiramente não-detectável por quaisquer dos testes usuais de discernimento de cores, já que o grupo anormal faz exatamente as mesmas discriminações de cores em resposta a exatamente os mesmos estímulos tal como o resto da população (SEARLE, 1997, p.65).

Em outros termos, podemos entender o argumento da seguinte maneira: há possibilidade de seres humanos distintos, com entendimentos distintos sobre as cores, a partir de um mesmo *input* de informação (ver a cor), terem um *output* de exteriorização diferente (dizer que cor é essa). Considerando esse cenário em que pessoas têm compreensões distintas sobre as cores (a exemplo: pessoas com daltonismo possuem um modo de distinguir as cores diferente das pessoas sem essa condição) se usarmos como base o fenômeno qualitativo interno diferente (usar o *qualia* de cor verde ao ver a cor vermelha) obtemos que há algo para além das relações causais e que deve ser considerado ao tratarmos da mente.

Ao considerarmos esta teoria do Naturalismo Biológico, a noção de compreender parece buscar algo além. Mas o que seria a *compreensão*? De antemão, Searle nos explica que ela é um fenômeno biológico complexo e intrinsecamente ligado à natureza do cérebro humano. Assim, ela não é algo que possa ser reduzido a um conjunto de regras ou algoritmos, mas sim, um produto da

evolução biológica e da interação entre o organismo e seu ambiente. Vejamos a definição de “compreensão”, para Searle, na íntegra:

Newell e Simon (1963) escrevem afirmando que o sentido de “compreensão” para os computadores é exatamente o mesmo que para os seres humanos. Gosto do modo incisivo desta afirmação e é este tipo de asserção que analisarei. Argumentarei que, em um sentido literal, o computador não compreende nada, da mesma maneira que o carro e a máquina de somar também não compreendem nada. A compreensão do computador não é como minha compreensão de alemão, ou seja, parcial ou incompleta, ela é zero (Searle, 2021, p286).

Searle está dizendo que o conceito de "compreensão" é muito mais complexo do que parece à primeira vista. Muitas vezes, nós simplificamos demais essa ideia no momento em que pensamos que alguém, ou algo, pode simplesmente "compreender" ou "não compreender" alguma coisa. Podemos compreender algo superficialmente ou profundamente, parcialmente ou completamente. A compreensão, nesta via de pensamento, é um espectro, não uma categoria binária. Podemos compreender um conceito, uma emoção, uma situação, uma pessoa de maneiras diferentes. Decorre que, cada tipo de compreensão exige diferentes habilidades e diferentes conhecimentos.

John Searle está nos alertando para não cairmos na “armadilha” de pensar na *compreensão* como algo simples e absoluto. Ele nos instiga a pensá-la como um processo complexo, influenciado por diversos fatores, a saber: conhecimento prévio: quanto mais sabemos sobre um assunto, melhor e mais profunda será a nossa compreensão; Contexto: o significado de algo pode variar de acordo com o contexto em que é apresentado; Experiência: nossas experiências pessoais moldam nossa capacidade de compreender o mundo; Intencionalidade: a compreensão envolve a atribuição de significado e intenção aos atos e objetos.

Assim sendo, ao contrapormos as teses que John Searle defende, ao afirmar o seu Naturalismo Biológico em detrimento do experimento mental de Turing, ainda que o computador digital obtenha a sua taxa de sucesso e passe no teste, simulando perfeitamente um humano em uma conversa, ele não nos constata nada mais do que uma simulação de comportamento de respostas, não uma “compreensão” sobre elas. Para a tese naturalista da mente, não cabe qualquer fator cognitivo de *qualia* à simples ação de resposta provável. O que demarca, de fato, é que um computador manipula sintaticamente e dá respostas prováveis, e sendo prováveis, vão entrar em similitude às respostas humanas, que também podem se exprimir em funções sintáticas, mas Searle acusa que as respostas humanas constam o fator do entendimento. Sendo assim, uma IA forte declama sua inteligência neste quesito: computadores são capazes de responder como humanos, e que, portanto, simulam o nosso raciocínio inteligente, mas isso não parece ser o caso de simular o entendimento em si. Ao passo que manipular não é o mesmo que pensar, a sintaxe é diferente de semântica. Assim, computadores modernos configuram ações inteligentes, e devido a isso, são artificialmente inteligentes, mas, segundo Searle, não podemos atribuir mentalidade a eles no momento em que consideramos mente aquilo que é dotado de conteúdos semânticos, qualitativos e subjetivos.

4 METODOLOGIA

Foi proposto um método investigativo-analítico que parte da exposição e análise das teorias relativas à natureza da mente por via da revisão literária de livros

e artigos científicos pertinentes à temática. O livro "Uma Introdução à Filosofia da Mente" (Leclerc, 2018) serviu como base para compreender o desenvolvimento histórico das questões sobre mente e cérebro. Dado este alicerce teórico inicial a exploração da concepção cartesiana sobre a mente como uma substância distinta do corpo, que desemboca na problemática mente-corpo, adentramos à problemática central deste trabalho, que consiste na exploração das defesas da naturalidade ou artificialidade da mente. Para isso, foram introduzidas duas doutrinas teóricas contemporâneas: Funcionalismo e Naturalismo Biológico, estas julgadas como as mais pertinentes à discussão. O trabalho seguiu com a análise argumentativa que cada doutrina expôs, também confrontou as teses das mesmas a fim de obter respostas acerca da mente.

5 RESULTADOS E DISCUSSÕES

Resulta, assim, que a problemática acerca da mente, sua condição de existência e relação com o mundo material é uma questão passível de discussão. Mesmo após o desenvolvimento da atual ciência e filosofia, ainda que venhamos a caracterizar aspectos similares entre os softwares mais avançados e a mente, ou pelo contrário, defender a mente como exclusivamente natural, não podemos postular uma posição definitiva, pois nenhuma das teorias expostas no decorrer deste artigo conseguiram comprovar a mente a partir das suas propostas teóricas.

Sendo assim, por não haver consenso no tocante ao que a mente pode ser, não há como cunhar uma definição universal para ela, visto que nenhuma das teorias da mente existentes (até o presente momento) comprovou que ela é essencialmente natural, ou pelo contrário, que é possível atribuir algum nível de mentalidade artificial, o que temos são pontos de vista distintos: uns que diferenciam as coisas imateriais das coisas físicas por categorias, outros pelo que a mente exprime pelos seus processos cognitivos, ou suas funções dentro de um sistema, e outros por sua naturalidade de interpretação e entendimento sobre as coisas. A história da Filosofia nos demonstra essa dificuldade de consenso: provar algo imaterial, que se manifesta de alguma forma, parece ser algo inconcebível, o ponto de decisão estará no consenso prévio do qual se pretende partir (como "eu" entendo a mente?). Sendo a mente algo que configura a impossibilidade de representação no mundo por uma caracterização definitiva com relação à sua essência, não podemos dizer o que ela é em sua plenitude, mas também não parece assertivo negar o seu sentido fenomênico/subjetivo, haja vista que, enquanto seres pensantes, sentimos (ainda que não possamos explicar) sua natureza qualitativa das experiências conscientes. O que declama que tal discussão volta novamente ao seu lugar de origem, à filosofia. Nos questionamos, então: O que é mente? Pressuponho que esta será uma discussão que estará em aberto por longos anos.

Ao analisar a subjetividade sob a lente do Naturalismo Biológico de John Searle, é entendido que os fenômenos mentais são de natureza biológica e se caracterizam pela subjetividade, *qualia* e outros aspectos fenomenológicos. O que parece não se aplicar como possível a um programa de computador (software), pois até o presente momento, não se tem o conhecimento da atribuição de características emocionais e fenomênicas ao mesmo. A subjetividade intrínseca à mente natural, que Searle defende, carrega consigo a natureza do que *realmente* ela apresenta ao nosso entendimento. Em contrapartida, pelo que podemos considerar do *experimento* mental de Turing, é possível identificar aspectos sintáticos que os computadores assimilam através de seus softwares (manipulação de símbolos) e

que, também, os seres humanos operam através de sua cognição, mas não pode ser o caso em que podemos implicar que uma ação similar (neste caso, o ato de manipular símbolos) caracterize uma igualdade entre a identidade dos seres naturais pensantes e dos computadores, isso seria um *problema lógico de identidade*⁷.

Entramos, portanto, no impasse, no momento em que nos questionamos sobre o que devemos considerar por mente. Levando em conta a comparação entre as similitudes das operações sintáticas operadas pelos computadores e os humanos, muito podemos entender sobre como a mente se expressa a partir de engenharia reversa⁸, mas ocorre que estaremos limitados ao aspecto puramente sintático. A mente natural, que John Searle acusa, traz consigo aspectos subjetivos que estão para além de qualquer redução lógica ou material, e que em sua plenitude, são inacessíveis. Nesta esteira, havendo a consideração prévia das teorias da mente trabalhadas neste artigo, podemos constatar que existe algo que os computadores possam expressar através dos seus softwares (de forma sintática) enquanto conteúdos mentais, ao passo que, também, podemos distinguir o software da mente natural no momento em que consideramos os fatores fenomênicos atribuídos ao conceito de mente natural. De um lado, temos a mente natural como a experiência subjetiva, os pensamentos, sentimentos e percepções que emergem de alguma forma do cérebro. Ela é intangível, subjetiva e abstrata. O cérebro é um órgão físico, composto por bilhões de neurônios que se comunicam através de sinapses. Ele é concreto, material e biológico. Por outro lado, hardware e software também apresentam características distintas. O hardware é a parte física de um computador, como o processador, a memória RAM, o disco rígido e os periféricos. Ele é tangível, material e responsável por executar as instruções. O software, por sua vez, é a parte intangível, composta por programas que fornecem instruções ao hardware. Ele é puramente lógico, abstrato e determina as funções que o computador pode realizar. Percebemos certa similitude no modo como mente/cérebro operam em comparação ao software/hardware, mas precisamos pontuar até onde esta similitude se faz como coerente.

6 CONCLUSÃO

Dado o percurso argumentativo trabalhado, conclui-se que: (1) Pelo que entendemos do Naturalismo Biológico de John Searle, a constituição de uma mente natural exige a semântica para além de manipulação de símbolos; (2) Pelo que entendemos do Funcionalismo, mais especificamente do “Jogo da Imitação” de Alan Turing, o processo de manipulação de símbolos caracteriza uma ação inteligente no mesmo sentido que a inteligência considerada natural. A defesa da naturalidade da mente que Searle faz, em seu “*Experimento do Quarto Chinês*”, expõe a distinção entre sintaxe e semântica, demonstrando que a compreensão *requer* mais do que a simples manipulação sintática de símbolos. Em contrapartida, Turing também demonstra, a partir do seu “Jogo da imitação”, que o seu modelo computacional pode ser inteligente na condição de conseguir simular um diálogo humano, e sendo

⁷ Seguindo os princípios lógicos aristotélicos explícitos na sua obra “Organon”, o conceito de identidade afirma que uma proposição é sempre verdadeira quando se refere a si mesma. Ou seja, o que esse princípio nos diz é que algo é o que é, e o contrário disso implica em uma falsidade, um erro. Simplificando em termos lógicos $A=A$ (V), $A=\sim A$ (F).

⁸ Podemos entender o aprendizado por engenharia reversa da seguinte forma: analogamente como a partir de um desmonte de um relógio, podemos entender como ele demarca as horas. Este é um processo de análise que parte do produto final (o relógio) para “compreender” seus componentes, sua estrutura, e, assim, como ele funciona.

assim, sustenta a sua inteligência neste sentido. O problema decai no que devemos considerar sobre o que é mente, se uma operação inteligente (manipular símbolos) pode conferir um *status* de mentalidade, ou do contrário, que só poderá conferir mentalidade àquilo que é provido de fator biológico e de entendimento. Decaímos, portanto, em um problema conceitual, e por isso, até o presente momento, não há como cunhar uma resposta definitiva sobre a mente.

Ademais, a mente é um fenômeno complexo que desafia qualquer tentativa de definição única e universal. A ausência de consenso sobre o que entendemos por mente é resultado da natureza subjetiva da experiência, das diferentes perspectivas disciplinares, das questões fundamentais ainda não respondidas e da influência da cultura e da história. Nós voltamos, portanto, à nossa “mãe”, à Filosofia. A busca por uma compreensão mais profunda da mente continua sendo um dos maiores desafios da humanidade, talvez seja uma busca inconclusa, mas a própria busca configura Filosofia, pois ela é o que nos motiva desvendar aquilo que parece ser inconcebível.

REFERÊNCIAS

ARISTÓTELES. **Da Alma: livros I, II e III**. Trad. Maria Cecília Gomes dos Reis. São Paulo: Editora 34, 2006.

ARISTÓTELES. **Órganon: Categorias, Da Interpretação, Analíticos anteriores, Analíticos posteriores, Tópicos, Refutações sofisticas**. Trad. textos adicionais e notas de Edson Bini. Bauru: EDI-PRO, 2005.

BÍBLIA, A. T. Salmos. In BÍBLIA. Português. **Sagrada Bíblia Católica: Antigo e Novo Testamentos**. Tradução de José Simão. São Paulo: Sociedade Bíblica de Aparecida, 2008.

DESCARTES, R. **Meditações. Os Pensadores**. São Paulo: Abril Cultural, 1983.

GOMES, V. P. **Sobre a existência ou inexistência de um função universal para a classe das funções recursivas primitivas**. Revista instante, Campina Grande-PB, v.1, n. 2, p. 76-91, 2018.

LECLERC, A. **Uma introdução à filosofia da mente**. 1. ed. Curitiba: Appris Editora, 2018.

NAGEL, T. **Que é ser um morcego?**. Cadernos de História e Filosofia da Ciência. Campinas, v. 20, n. 1, p. 11-22, 2013.

Putnam, H. (1960). **Minds and Machines**. Dimensions of Mind: A Symposium. S. Hook. New York, Collier: 138-164.

Putnam, H. (1967a). **The Mental Life of Some Machines. Intentionality, Minds, and Perception**. H. Castañeda. Detroit, Wayne State University Press: 177-200.

Putnam, H. (1967b). **Psychological Predicates**. Art, Philosophy, and Religion. Pittsburgh, PA, University of Pittsburgh Press.

REALE, Giovanni. **História da filosofia antiga. Volume I: das origens a Sócrates**. Trad. Marcelo Perine. São Paulo: Loyola, 1993.

SEARLE, J. R. **Mentes, Cérebros e Programas**. Trad. Cléa Regina de Oliveira Ribeiro. Fortaleza: Revista Reflexões, 2021.

SEARLE, J. R. **O mistério da consciência: e Discussões com Daniel C. Dennett e David J. Chalmers**. Paz e Terra, 2008.

SEARLE, J. R. **A Redescoberta da Mente**. São Paulo, SP: Martins Fontes, 1997.

SEARLE, J. R. **The Rediscovery of the Mind**. Cambridge: MIT Press, 1992.

TURING, A. M. **Computing Machinery and Intelligence**. 1. ed. Mind, 1950.

AGRADECIMENTOS

Aos meus pais, *Maria de Lourdes e Marizio José*, por seu amor e apoio incondicional, que no sol, tornaram a minha jornada acadêmica possível, na sombra. Aos meus amigos: *Maauri Neto, Erivaldo Pereira, Eduardo Gomes, Emanuele Dantas, Fábio Alexandrino, Luan Borba e Milene Vanessa*, que estiveram, de uma forma ou outra, presentes em cada etapa desta jornada acadêmica, tornando todo processo árduo em algo suave. Aos meus mestres, a todos meus professores do departamento de filosofia, em especial: *Irio Coutinho, Otacílio Neto, Simone Marinho, Ramon Bolivar, Eugênia Teles, Nilton Conserva, Márcio Correia e Alysson Pereira*. Aos meus orientadores, *Carlos Antônio e Victor Pereira*, aqueles que compartilharam todo conhecimento e tiveram a paciência de me mostrar o caminho a seguir. Este trabalho é dedicado a vocês, que sempre acreditaram em mim.