

**UNIVERSIDADE ESTADUAL PAULISTA “JULIO DE MESQUITA
FILHO”
MESTRADO ACADÊMICO EM FILOSOFIA**

Leonardo Francisco Costa de Andrade

**CONSCIÊNCIA E INTENCIONALIDADE NA ERA DOS BIG DATA: A
TEORIA DA COMPETIÇÃO DOS MARCADORES SEMÂNTICOS DE
PAUL THAGARD**

MARÍLIA-SP

2021

Leonardo Francisco Costa de Andrade

**CONSCIÊNCIA E INTENCIONALIDADE NA ERA DOS BIG DATA: A
TEORIA DA COMPETIÇÃO DOS MARCADORES SEMÂNTICOS DE
PAUL THAGARD**

Orientador: Prof. Dr. Jonas Gonçalves Coelho

Dissertação de Mestrado Acadêmico
apresentada ao Programa de Pós-graduação
em Filosofia, da Faculdade de Filosofia e
Ciências, da Universidade Estadual Paulista –
UNESP – Campus de Marília, para exame de
defesa de mestrado em 2021. Pertencente à
área de concentração: Filosofia da
informação, da cognição e da consciência.

MARÍLIA-SP

2021

A553c Andrade, Leonardo Francisco Costa de
Consciência e intencionalidade na era dos big data : a teoria da
competição dos marcadores semânticos de Paul Thagard / Leonardo
Francisco Costa de Andrade. -- Marília, 2021
115 f. : il., tabs.

Dissertação (mestrado) - Universidade Estadual Paulista (Unesp),
Faculdade de Filosofia e Ciências, Marília
Orientador: Jonas Gonçalves Coelho

1. Ciências cognitivas. 2. Filosofia da mente. 3. Consciência. 4.
Intencionalidade. 5. Big Data.. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de
Filosofia e Ciências, Marília. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

AGRADECIMENTOS

Para que esta dissertação de mestrado fosse concluída, muitas pessoas estiveram envolvidas. Desse modo, devo ser justo e agradecer a todos aqueles, que de algum modo auxiliaram na conclusão deste projeto. Seja pela ajuda direta, com orientações e valiosas lições acadêmicas, ou indiretas, proporcionando apoio, afeto e experiências positivas, os quais possibilitaram muitos *insights* e equilíbrio emocional para que eu pudesse concluir mais essa etapa em minha vida.

Desse modo, gostaria de iniciar agradecendo aos meus pais Edson e Silvana pela excelente educação e amor que sempre dedicaram a mim e a meu irmão, durante toda a nossa vida. Aproveito para agradecer ao meu irmão Lucas, pelo apoio e amizade durante todos esses anos. Agradeço também a minha companheira Thauane, por estar ao meu lado e suportar minhas crises de ansiedade e *stress* ocasionadas pelas preocupações com a conclusão desta pesquisa.

Gostaria de agradecer também a todos os amigos do Dojo, especialmente, ao Sensei Clebert Eimori Kato e ao Prof. Dr. Paulo Ribeiro da Cunha, os quais por meio do projeto “Artes Marciais como Expressão de um Projeto Pedagógico” possibilitaram meu reencontro com a magnífica arte do Karate-Do, tornando essa atividade física em um excelente meio para o desenvolvimento físico e psicológico.

Além destes, gostaria de agradecer aos colegas do PPGFil por todas as experiências vivenciadas em conjunto, sejam nas riquíssimas discussões travadas nos eventos e nos arredores da FFC, até mesmo em meio às confusões nos momentos de solicitação de verba, agendamentos de transportes e estadia para participar de eventos em outras Universidades.

Gostaria ainda de, notadamente, agradecer as magníficas Prof.^a Dr.^a Maria Eunice Quilici Gonzalez e a Prof.^a Dr.^a Mariana Cláudia Broens, as quais contribuíram imensamente para minha formação em Filosofia. Além de sempre se portarem como pessoas dignas e idôneas, constantemente preocupadas com os interesses do programa, assim como com o bem-estar de seus alunos. Além disso, gostaria de enfatizar a minha gratidão pela oportunidade de participar do projeto “Compreendendo a dinâmica da opinião e da linguagem utilizando *Big Data*”, o qual tive a oportunidade de participar como bolsista FAPESP de treinamento técnico, experiência que me possibilitou participar de eventos interessantíssimos,

além de estabelecer contato com renomados professores. Por todos estes motivos, agradeço de coração por toda ajuda e auxílio ao longo desta jornada.

Gostaria também de agradecer ao meu orientador, o Prof. Dr. Jonas Gonçalves Coelho, o qual, desde o princípio, me acolheu auxiliando na elaboração do projeto de pesquisa e, conseqüentemente, no meu ingresso no Programa de Pós-graduação em Filosofia. Além de sempre tratar com muita educação e boa vontade todas as solicitações de auxílio e orientação, apresentando sempre direções e sugestões para que esta pesquisa fosse concluída da melhor maneira. Sem dúvida, esta é uma pessoa ímpar na conclusão deste trabalho, sem a qual, nada disso seria possível.

Por fim, mas não menos importante, gostaria de agradecer ao Prof. Dr. Alfredo Pereira Junior e ao Prof. Dr. João Antonio de Moraes pelas riquíssimas sugestões e orientações dadas em minha banca de qualificação e pelo aceite em participar novamente em minha banca de defesa. As sugestões elaboradas por estes professores possibilitaram que novos caminhos fossem traçados ao longo deste trabalho, bem como, com que as principais ideias presentes se conectassem, elevando o nível da pesquisa. Agradeço também ao Prof. Leonardo Lana de Carvalho por sua, sempre presente, cordialidade no aceite de participar como suplente da banca de defesa.

Agradeço ainda as agências de fomento CAPES e FAPESP pelo apoio financeiro a esta pesquisa. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. E foi concedida a bolsa de treinamento técnico (TT3) sob o processo nº 2019/09775-1, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP).

Concluo esta seção agradecendo a todos os indivíduos citados, com toda certeza, todos vocês foram de suma importância para que eu pudesse alcançar mais este objetivo em minha vida e para que esta pesquisa fosse concluída dignamente. Por todos estes motivos, enfatizo: Muito obrigado!!

“Às vezes, uma idéia muito complexa entra na consciência pública de uma forma tão altamente simplificada, que é pouco mais que uma caricatura do original; [...] a metáfora do computador repousa apenas no entendimento mais vago de um conceito científico difícil e complexo. [...] disso resulta que um computador possa ao menos imitar o homem, a natureza e a sociedade em todos os seus aspectos processuais.” Joseph Weizenbaum – Computer Power and Human Reason

RESUMO

Partindo da ideia geral sobre a analogia entre humano-máquina e sua interação, assunto presente nos debates sobre a relação mente/corpo, no desenvolvimento da Ciência cognitiva e, mais recentemente, no campo da ciência de dados com o advento dos *Big Data*. A presente pesquisa partindo da elaboração cartesiana sobre a relação mente/corpo, bem como o posterior debate na área da Filosofia da mente e da Ciência cognitiva, visa investigar de que modo os *Big Data* podem auxiliar na compreensão dos fenômenos Consciência e Intencionalidade e as possíveis implicações envolvidas. Segundo alguns posicionamentos dualistas contemporâneos, o fenômeno da Consciência e, por conseguinte, da Intencionalidade seriam incapazes de serem reduzidos ao seu substrato físico, uma vez destacado seu aspecto fenomênico evidenciando uma espécie de lacuna explicativa. Por outro lado, alguns posicionamentos funcionalistas tendem a encarar essa forma de distinção entre os aspectos fenomênicos e psicológicos da Consciência e da Intencionalidade enquanto obstáculos à compreensão científica destes fenômenos. Desse modo, será intentado investigar duas hipóteses principais: 1) Se o aspecto fenomênico da Consciência e da Intencionalidade é verdadeiro e se configura como um *hard problem*, então os *Big Data* não possuem potencial para auxiliar em sua investigação. Caso contrário, sua investigação pode ser aprimorada com sua aplicação. 2) No caso dos *Big Data* auxiliarem na compreensão destes fenômenos, quais seriam os possíveis impactos na manipulação/influência aos usuários? Para tanto, será empregada a abordagem da Competição de Marcadores Semânticos (CMS) de Paul Thagard com a finalidade de analisar e investigar as hipóteses destacadas, evidenciando a importância das moléculas que compõem a dinâmica bioquímica do cérebro, bem como o relevante papel desempenhado pela cognição e pelas emoções em meio à emergência de processos conscientes e intencionais.

Palavras chave: Ciência cognitiva; Filosofia da mente; Consciência; Intencionalidade; *Big Data*.

ABSTRACT

Starting from the general idea about the analogy between human-machine and its interaction, a subject present in debates about the mind/body relationship, in the development of cognitive science and more recently in the field of data science with the advent of Big Data. This research starting from the Cartesian elaboration on the mind/body relationship, as well as the subsequent debate in the area of Philosophy of Mind and Cognitive Science, aims to investigate how Big Data can help to understand the phenomena of Consciousness and Intentionality and the possible implications involved. According to some contemporary dualist positions the phenomenon of Consciousness and, therefore, of Intentionality would be incapable of being reduced to its physical substrate, once it's phenomenal aspect is highlighted, evidencing a kind of explanatory gap. On the other hand, some functionalist positions tend to view this form of distinction between the phenomenal and psychological aspects of Consciousness and Intentionality as obstacles to the scientific understanding of these phenomena. Thus, it will be intended to investigate two main hypotheses: 1) If the phenomenal aspect of Consciousness and Intentionality is true and configures itself as a hard problem, then Big Data has no potential to assist in your investigation. Otherwise, your investigation can be improved with your application. 2) In the case of Big Data helping to understand these phenomena, what would be the possible impacts on manipulation/influence on users? Therefore, the Semantic Pointers Competition (SPC) approach of Paul Thagard will be used in order to analyze and investigate the highlighted hypotheses, highlighting the importance of the molecules that make up the biochemical dynamics of the brain, as well as the relevant role played by cognition and by emotions amid the emergence of conscious and intentional processes.

Keywords: Cognitive Science; Philosophy of Mind; Consciousness; Intentionality; Big Data.

LISTA DE QUADROS E FIGURAS

Quadro 1 – Critério de avaliação de teorias da representação mental.....	30
Figura 1 – Representação de populações de neurônios.....	52
Figura 2 – Representação da formação de marcadores semânticos.....	65
Figura 3 – Representação dos componentes funcionais do modelo de intenção	68

Sumário

LISTA DE QUADROS E FIGURAS.....	iii
Sumário	3
INTRODUÇÃO.....	3
CAPÍTULO 1:	
HISTÓRICO DE DESENVOLVIMENTOS: NA FILOSOFIA DA MENTE E NA CIÊNCIA COGNITIVA.....	
9	
1.1 Retomando antigos problemas: Descartes e a substância pensante.....	10
1.2 O problema mente/corpo: A retomada da Filosofia da mente no período contemporâneo	15
1.3 Ciência Cognitiva: Interdisciplinaridade, Paradigmas e Informação	23
1.4 A noção de Representação e a Cognição Incorporada e Situada.....	29
CAPÍTULO 2:	
OS FENÔMENOS CONSCIÊNCIA E INTENCIONALIDADE: CONCEITUAÇÃO, LIMITES	
INVESTIGATIVOS, DETALHES ESTRUTURAIS E MODELAGEM COMPUTACIONAL.....	
36	
2.1 O conceito de Consciência: Consciência psicológica e consciência fenomênica	37
2.2 Consciência afetiva: Composição, estrutura e processos.....	47
2.3 O conceito de Intencionalidade: Intencionalidade intrínseca e derivada	56
2.4 Intencionalidade, emoções e ação: uma teoria neural baseada em marcadores semânticos.....	63
CAPÍTULO 3:	
BIG DATA, CONSCIÊNCIA E INTENCIONALIDADE: OS RISCOS ENVOLVIDOS E OS	
DESENVOLVIMENTOS ALCANÇADOS	
71	
3.1 A promessa de revolução científica dos <i>Big Data</i>	72
3.2 Os riscos da era dos <i>Big Data</i> : a ditadura dos dados	78
3.3 <i>Big 'Brain' Data</i> : A aplicação dos <i>Big Data</i> na investigação da Consciência e Intencionalidade	84
3.4 Os <i>Big Data</i> da Neurociência e sua relação com a Neuroética.....	90
CONSIDERAÇÕES FINAIS.....	96
REFERÊNCIAS BIBLIOGRÁFICAS.....	102

INTRODUÇÃO

A presente pesquisa parte da ideia geral sobre a analogia entre humanos-máquinas e sua interação. Como destacado pela citação de *Joseph Weizenbaum (1985)*¹, logo na epígrafe deste trabalho, não é incomum, mesmo em contextos de pesquisa, que ideias complexas venham a ser tratadas de modo tão simplificado pelo entendimento comum, que seu sentido real é completamente deslocado do seu contexto de formação. No caso de Weizenbaum, este se referia a metáfora entre mente e computador, a qual após surgir no contexto da Filosofia da mente, passa a ganhar maior amplitude em meio à Ciência cognitiva. Já entre os temas tratados por esta dissertação, a citação de Weizenbaum permite-nos refletir sobre a promessa de revolução científica direcionada à aplicação de técnicas de *Big Data*.

A noção geral destacada acima constitui-se como parte importante dos debates travados no campo da Filosofia da mente, Ciência Cognitiva e, contemporaneamente, sobre a aplicação da tecnologia de *Big Data* no comércio, política e na ciência em geral. Para tanto, a fim de elucidar o desenvolvimento deste debate, foram selecionados os fenômenos da Consciência e da Intencionalidade. A justificativa para a seleção destes fenômenos na contextualização do debate reside em sua resistência a uma redução funcionalista, a qual pressupõe a possibilidade de uma redução dos fenômenos mentais a sua função, permitindo sua replicação artificial, bem como a íntima relação com questões éticas como a possibilidade de livre escolha, responsabilidade moral e garantia de direitos.

Os objetivos primários deste trabalho são: 1) A adoção de uma abordagem monista/naturalista que possibilite lidar com os fenômenos da Consciência e da Intencionalidade; 2) Investigar se o aspecto fenomênico da Consciência e da Intencionalidade é verdadeiro e se configura como um *hard problem*, pois, então, os *Big Data* não possuiriam potencial para auxiliar em sua investigação. Caso contrário, sua investigação poderia ser aprimorada com sua aplicação. 3) No caso dos *Big Data* auxiliarem na compreensão destes fenômenos, quais seriam os possíveis impactos na manipulação/influência aos usuários.

Além destes objetivos, a presente pesquisa, ainda apresenta a intenção de evidenciar o desenvolvimento histórico do debate, partindo do *dualismo substancial* cartesiano, seguido

¹ “Nascido em 8 de janeiro de 1923, Berlim, Alemanha; inventor em 1965 de ELIZA, o psiquiatra mecânico, que primeiro alertou para os perigos de confundir computadores com pessoas e vice-versa. [...] No início dos anos 1960, trabalhando com a tecnologia emergente de computadores interativos de compartilhamento de tempo, ele desenvolveu um psiquiatra mecânico chamado ELIZA, que parecia ser capaz de conduzir conversas significativas. Na verdade, o sistema (Weizenbaum 1965) usava palavras-chave nas respostas do usuário para criar perguntas comumente usadas que davam a aparência de compreensão e capacidade de resposta. De certa forma, ELIZA pode ter sido uma das primeiras abordagens para a solução do teste de Turing para inteligência”. (LEE, 1995).

pela retomada do debate no campo da Filosofia da mente, sua relevância no surgimento da Ciência cognitiva e sua permanência em meio à emergência da era dos *Big Data*. Além de ser intentado evidenciar alguns modos de abordar os fenômenos Consciência e Intencionalidade, e apresentar algumas pesquisas contemporâneas na abordagem do tema.

Dessa forma, a fim de alcançar os objetivos destacados, a presente pesquisa foi dividida em três capítulos. O primeiro capítulo tem como objetivo elucidar o contexto de debate dos temas abordados. Partindo, desse modo, da segmentação ontológica cartesiana entre mente e corpo, seguido pelas críticas elaboradas, ainda na primeira metade do século XX, sob o contexto da Filosofia da mente, chegando ao desenvolvimento conceitual/tecnológico alcançado com o advento da Ciência Cognitiva e, por fim, elucidando o debate sobre a noção de Representação sob algumas perspectivas contemporâneas da Ciência Cognitiva.

Para isso, inicio a primeira sessão deste capítulo com a formulação do problema mente/corpo elaborada por René Descartes, evidenciando a linha de raciocínio traçada pelo autor na formulação do *cogito*, o qual postula uma segmentação ontológica entre a mente e o corpo. Sendo ainda apresentada à resposta cartesiana ao problema da causação mental, bem como o modo como o autor busca distinguir a humanidade de seres autômatos.

Após a formulação do *dualismo cartesiano*, o fim de século XIX e início do século XX, foi um marco para o levante de diversas críticas a esta concepção dual entre mente e corpo. Um dos principais fatores que influenciaram o advento de propostas críticas à posição cartesiana, foi a busca por métodos que permitissem abordar as questões relativas à mente e sua relação com o corpo de modo científico. Este objetivo de contornos científicos influenciou em grande parte o debate filosófico da época, ocasionando no levante de diversas teorias críticas sob o contexto da Filosofia da mente.

Dentre tais teorias críticas, podemos agrupá-las em alguns subgrupos, dentre eles, podemos destacar as abordagens *fisicalistas*, as quais podem apresentar objetivos *reducionistas* ou *eliminacionistas*; as abordagens *monistas* sob o escopo do *dualismo de propriedade*; e abordagens que, ao menos em princípio, seriam compatíveis com as propostas fisicalistas e dualistas, como é o caso do *Behaviorismo filosófico* e do *Funcionalismo*.

Ao longo de todo o debate travado pela Filosofia da mente, ainda na primeira metade do século XX, a abordagem que mais conquistou adeptos foi a proposta *Funcionalista*. Parte da popularidade desta abordagem teve como fator determinante o desenvolvimento dos

primeiros computadores, fato que também serviu de influência para a consolidação do *paradigma simbólico* da Ciência Cognitiva.

O campo de pesquisa da Ciência Cognitiva pode ser compreendido como um “espaço” de encontro de outras disciplinas independentes, como a Filosofia, Antropologia, Linguística, Inteligência Artificial, Neurociência e Psicologia. Interdisciplinar por excelência, ao longo do seu desenvolvimento, esta passou por duas mudanças paradigmáticas principais. Apesar de um início com pouco entusiasmo, o primeiro paradigma da Ciência Cognitiva foi o paradigma *conexionista*, o qual beneficiou-se das contribuições de Shannon, McCulloch e Pitts, mas logo foi abafado pelo advento do paradigma *simbólico*, até recobrar a sua força após a década de 80 com os desenvolvimentos no campo da Neurociência.

Em meio ao paradigma simbólico, a premissa suprimida era de que pensar seria o mesmo que computar, assim, o pensamento seria dependente do encadeamento adequado de representações simbólicas. Já no paradigma conexionista o pensamento é compreendido em analogia as computações em paralelo realizadas por redes de neurônios artificiais. As vantagens que a retomada do paradigma conexionista trouxe, em relação ao paradigma simbólico, é que este permitia a simulação dos processos em paralelo ocorridos no cérebro. Além de lidar com a informação de modo distribuído, permitindo o funcionamento da rede, mesmo no caso de haverem unidades danificadas.

No entanto, apesar das diferenças entre estes dois paradigmas da Ciência Cognitiva, ambos utilizam-se amplamente da noção de Representação. A noção de Representação assumida pela presente pesquisa é fornecida por Paul Thagard, a qual é compreendida como uma espécie de padrão de ativação entre neurônios ou populações neuronais. Para que estes padrões de ativação apresentem capacidade representacional, eles devem assumir correlações causais estáveis entre os disparos neuronais e os objetos da representação.

Contudo, esta noção de Representação não é comum a Ciência Cognitiva como um todo, sendo, então, brevemente comentado sobre algumas posições críticas a noção de Representação, bem como a resposta de Thagard a algumas dessas críticas.

No Capítulo 2, são retomados alguns pontos introduzidos ainda no primeiro capítulo, como, por exemplo, alguns posicionamentos sobre a relação mente e corpo, os quais, muitas vezes, alegam pela irreducibilidade da Consciência. Passando, então, a tratar diretamente dos fenômenos Consciência e Intencionalidade. Para tanto, as duas primeiras seções deste capítulo

são reservadas para tratar sobre o debate envolvendo o fenômeno da Consciência, enquanto as duas últimas são reservadas para tratar sobre a Intencionalidade.

Sendo este um capítulo mais técnico, em sua primeira metade, serão destacados alguns pontos principais como: A distinção entre a consciência fenomênica e psicológica elaborada por Chalmers, a crítica de Thagard ao argumento da múltipla realizabilidade, os posicionamentos de Searle e Dennett sobre a Consciência e a proposta da teoria da Consciência afetiva de Thagard e Aubie.

A conceitualização de Chalmer na distinção entre os termos Consciência fenomênica e psicológica é importante, uma vez que figura em meio a uma das hipóteses centrais da presente pesquisa. Segundo a distinção, é afirmado que um dos motivos que dificultam uma abordagem completa da Consciência tem relação com o seu *Hard Problem*, ou seja, a dificuldade de ser explicado por que razão todo estímulo está associado diretamente a uma experiência consciente. Sendo afirmado que descrições fisiológicas, comportamentais ou causais sobre a Consciência não diriam nada sobre seu aspecto fenomenal. Desse modo, considerando a legitimidade desta afirmação, podemos inferir que os *Big Data* em nada contribuiriam para aprimorar nosso conhecimento sobre este fenômeno.

Como visto anteriormente, a concordância com esta premissa tem forte relação com o modo como o problema mente/corpo é compreendido. Por este motivo são apresentadas as teorias da Consciência de Searle e Dennett. A finalidade desta apresentação é evidenciar como são comumente desenvolvidas as defesas segundo o *funcionalismo* e o *dualismo de propriedade*, sendo estes pontos de vistas concorrentes segundo as concepção fenomênica e psicológica da Consciência.

No entanto, apesar da elaborada argumentação deste dois autores, nenhum deles pareceu apresentar uma descrição completa, que pudesse fundamentar a aplicação de técnicas de *Big Data*. Por outro lado, a abordagem da Consciência afetiva de Thagard e Aubie apresenta-se como uma descrição mais detalhada sobre muitos aspectos da Consciência, sendo, portanto, mais compatível com a proposta da presente pesquisa.

Seguindo a segunda metade do capítulo, passamos a tratar sobre o fenômeno da Intencionalidade. A Intencionalidade pode ser compreendida como a capacidade dos estados mentais de se dirigirem a objetos e estados de coisa do mundo. Além disso, este fenômeno assume uma forte relação com a Consciência, dado que, na maioria dos casos, sempre estamos conscientes de algo ou de alguma coisa, sendo este o conteúdo intencional.

Novamente, passamos a apresentar os pontos de vista de Searle e Dennett. Com visões claramente conflitantes, suas divergências se destinam às acepções dos termos *intencionalidade intrínseca e derivada*. Basicamente, podemos distinguir esses termos da seguinte forma: As palavras redigidas nesta dissertação apresentam um tipo de *intencionalidade derivada* das ideias e interpretações da mente deste que vos escreve, as quais seriam, portanto, dotadas de *intencionalidade intrínseca*.

Contudo, apesar desta interessante distinção, a proposta teórica assumida para lidar com este fenômeno é desenvolvida por Thagard, Schröder e Stewart, conhecida como *teoria neural baseada em marcadores semânticos*.

A compreensão das intenções enquanto uma espécie de marcador semântico, a qual baseia-se na concepção de marcadores semânticos proposto por Elia Smith (2013), propõe que as intenções seriam como padrões de atividade dos disparos neurais, os quais mantêm em sua estrutura informações comprimidas de padrões de disparos de outras regiões cerebrais. Essa abordagem é relevante e endossada pela presente pesquisa devido ao seu alto grau de explicação sobre as lacunas entre intenção-ação, ações automáticas e deliberadas, bem como a sua ampla descrição dos mecanismos envolvidos na produção das intenções.

Dessa forma, o que pode ser concluído ao fim deste capítulo é que as abordagens da Consciência afetiva e a Teoria neural baseada em marcadores semânticos de Paul Thagard são as mais indicadas para averiguar o papel da Consciência e da Intencionalidade em meio à era dos Big Data. Isso, pois, dentre as abordagens apresentadas, são as que apresentam maior grau de explicação, detalhamento e previsibilidade do comportamento.

No terceiro Capítulo, inicio com uma breve definição dos *Big Data*, alguns fatores relacionados à sua emergência e a promessa de revolução científica ocasionada pela sua aplicação. Posteriormente são apresentados alguns riscos envolvidos na era dos grandes dados. Partindo da discussão sobre o papel dos fenômenos Consciência e Intencionalidade e, por fim, apresentando alguns projetos em Neurociência com aplicação de técnicas de *Big Data* e a contribuição da Neuroética no desenvolvimento destes projetos.

O termo *Big Data* se refere à tecnologia contemporânea desenvolvida no campo da ciência de dados. Segundo seus entusiastas, é suposto que a grande vantagem dos *Big Data* seria que, com o aumento na quantidade de dados, o desenvolvimento qualitativo de suas análises também seria aprimorado. Sendo afirmado que ao serem enfatizadas as correlações entre dados, ao invés, da busca pela causalidade, seria a grande contribuição dos *Big Data*

para uma revolução no campo científico. Os principais fatores envolvidos na sua emergência são o aprimoramento das técnicas de *datificação*, o desenvolvimento da computação ubíqua e, consequentemente, a popularização das TIC's (Tecnologias Informacionais da Comunicação).

Entretanto, apesar da promessa dos *Big Data* ter o potencial para ocasionar uma revolução científica na ciência em geral, são ainda muitos os riscos envolvidos. Dentre eles, os principais riscos que a aplicação de técnicas de *Big Data* oferecem a sociedade, são riscos a privacidade, liberdade e autonomia humana. Os riscos a privacidade envolvem as questões do consentimento informado, dificuldades no processo de anonimização e possibilidade de reidentificação dos usuário. São também destacados os riscos a liberdade ocasionado pela aplicação de técnicas preditivas no contexto penal. Além dos riscos a autonomia humana por meio da vigência de uma ditadura dos dados.

Já nos *Big Data* da Neurociência algumas dificuldades são comuns as suas aplicações gerais, enquanto outras configuram-se particularmente segundo este contexto.

O problema geral da privacidade também é um aspecto importante em meio aos *Big Data* da Neurociência, uma vez considerada a sensibilidade destes dados. Contudo, uma proposta para amenizar os riscos a privacidade seria a aplicação da *governança responsável dos dados* no contexto neurocientífico. Além disso, existem questões técnicas e culturais que necessitam ser averiguadas para que os *Big 'Brain' Data* possam apresentar resultados promissores. Dentre as questões técnicas podemos citar a variedade de métodos e técnicas na coleta de dados, bem como a diversidade de modelos animais, além da correlação com dados comportamentais. Sendo estes elementos responsáveis pelo aumento na quantidade de dados e da complexidade das análises. Para tanto, é sugerido que mudanças de ordem culturais também são necessárias como modos mais eficazes de compartilhamento e integração de dados, bem como padronização dos dados e reflexão sobre o papel da teoria em meio as análises automatizadas.

Por fim, reservo a última sessão, uma descrição de alguns projetos de pesquisa em curso, dos quais utilizam-se de técnicas de *Big Data* voltadas ao campo da Neurociência. Dentre os projeto descritos, são selecionados o *Human Brain Project* e o *BRAIN Initiative* para uma análise mais detalhada, averiguando o progressos alcançados e as dificuldades envolvidas. Sendo ainda apresentada algumas colocações, de dois membros do subgrupo de Neuroética do *BRAIN Initiative*, sobre a importância do debate de questões éticas em projetos deste porte.

CAPÍTULO 1

**Histórico de desenvolvimentos: Na Filosofia da mente
e na Ciência Cognitiva**

1.1 Retomando antigos problemas: Descartes e a substância pensante

O debate sobre a natureza dos estados mentais não é recente. Contudo, o que é relevante para introduzir a discussão proposta nesta dissertação, é a reformulação desse antigo problema realizada por René Descartes no período Moderno, em meados do século XVII.

Descartes compreendia que as diversas áreas do conhecimento seriam análogas as ramificações de uma única árvore, composta por único tronco, seu propósito seria, então, estabelecer um método de investigação geral o suficiente para ser aplicado a qualquer problema que pudesse ser colocado ao *espírito humano*². Para isto, seria necessária a constituição de uma *metafísica* que permitisse deduzir os princípios de sua física. Uma vez pressuposto que toda a ciência tem como origem o espírito humano, a dificuldade residiria no estabelecimento de um método que permitisse deduzir um princípio originário de todos os outros. Por assim dizer, sendo o conhecimento matemático aquele que apresenta um raciocínio satisfatório ao uso normal da razão, a filosofia cartesiana tem como objetivo descrever a maneira que o espírito pensa, quando pensa matematicamente, transpondo-o para quaisquer outros domínios. Este foi o objetivo percorrido por Descartes por muitos anos. Todavia, devido à força hegemônica da doutrina escolástica durante este período, o projeto de publicação de tal método passou por diversas prorrogações.

Em seu texto, Discurso do Método, Descartes apresenta as dificuldades impostas à razão humana no julgamento satisfatório entre o falso e o verdadeiro, não atribuindo esta dificuldade a deficiências intelectuais dos indivíduos, mas devido aos desvios nas etapas de verificação das evidências. Além de discorrer sobre as controversas apresentadas na elaboração do conhecimento filosófico e científico:

Nada direi da filosofia, a não ser que, vendo que foi cultivada pelos mais excelentes espíritos que viveram desde há vários séculos, e que, não obstante, nela não se encontra coisa alguma sobre a qual não se discuta e, por conseguinte, que não seja duvidosa, eu não tinha tanta presunção para esperar me sair melhor do que os outros; e que, considerando quantas opiniões diversas pode haver sobre uma mesma matéria, todas sustentadas por pessoas doutas, sem que jamais possa haver mais de uma que seja verdadeira, eu reputava quase como falso tudo o que era apenas verossímil. [...] Quanto às outras ciências, na medida em que tiram seus princípios da filosofia, eu julgava que nada de sólido se podia ter construído sobre fundamentos tão pouco firmes. (DESCARTES, 2007, p.17)

²

A definição cartesiana de espírito humano pode ser equiparada à concepção contemporânea de 'mente', tendo em vista, a distinção elaborada pelo autor entre espírito e *razão*. O espírito engloba além da razão, a capacidade de distinguir o verdadeiro do falso, a memória, imaginação e todas as outras faculdades do pensamento.

A diversidade do debate filosófico, apresentado nesta passagem, implica em uma diversidade semelhante ao conhecimento científico, uma vez que seus princípios seriam derivados do primeiro. Visando então o estabelecimento de um princípio filosófico indubitável que o permitisse deduzir todos os outros princípios, Descarte toma como falso tudo aquilo que se apresenta apenas como verossímil. Ao adotar essa postura, o autor intenta destituir o princípio hegemônico que servia de fundamento a tese de que todo o conhecimento era advindo dos sentidos, buscando provar que o próprio pensamento é capaz de produzir formas indubitáveis de conhecimento que jamais tenham passado pelos sentidos.

Na segunda parte do discurso são apresentadas quatro regras simples que compõe seu método, sendo elas: 1) nunca admitir como verdadeiro algo que não seja possível conhecer de forma evidente, evitando juízos precipitados e preconceitos, os quais sejam possíveis de se analisar clara e distintamente. 2) dividir em parcelas o problema central derivando em problemas menores e mais simples. 3) elencar os problemas mais simples para partir aos mais complexos, sempre supondo uma ordem dedutiva entre eles. 4) atribuir enumerações, as quais permitam uma revisão da cadeia dedutiva. Neste ponto, a dúvida hiperbólica entra em vigor logo na primeira regra. Contudo, devemos nos atentar que o método de por em dúvida todas as antigas opiniões, não se trata de uma imposição a toda a ciência e seu ensino, mas um método utilizado pelo autor para deduzir os princípios da primeira filosofia.

Pressupor a veracidade aos dados do sentido é uma atitude contraindicada na aplicação do método, visto que as opiniões mantidas até aquele momento seriam apreendidas dos sentidos ou pelos sentidos e tendo sido enganado pelos mesmos em outro momento, não seria prudente confiá-los novamente tal tarefa. Os sentidos sendo, portanto, um meio questionável a aquisição do conhecimento, poderiam ser postos à prova por meio da dúvida metódica. Ao levar o método da dúvida ao ápice de sua formulação, seria possível questionar toda a realidade sensível, como se toda a experiência do mundo não passasse de um sonho.

Ora, quem me pode assegurar que esse Deus não tenha feito com que não haja nenhuma terra, nenhum céu, nenhum corpo extenso, nenhuma figura, nenhuma grandeza, nenhum lugar, e que não obstante eu tenha os sentimentos de todas essas coisas, e que tudo isso não me pareça existir de modo diferente do que eu o vejo? [...] A tais razões nada tenho, por certo, responder, mas sou forçado a confessar que, de todas as opiniões que outrora recebera em minha crença como verdadeiras, não há uma da qual não possa agora duvidar, não por alguma inconsideração ou leviandade, mas por razões muito fortes e maduramente consideradas. (DESCARTES, 2000, p. 36)

Porém, argumenta o autor, mesmo que seja possível que em todo o conhecimento esse gênio maligno o engane, é certo, que se o gênio o engana, então, o enganado existe, pois “jamais poderá fazer com que eu não seja nada, enquanto eu pensar ser alguma coisa” (DESCARTES, 2000, p.43). Assim, mesmo supondo duvidar de toda a realidade circundante, o fato de que este existente pensa é indubitável. “Mas logo depois atentei que, enquanto queria pensar assim que tudo era falso, era necessariamente preciso que eu, que o pensava, fosse alguma coisa” (DESCARTES, 2007, p.58). Sendo assim, o autor conclui que a essência de sua existência era existir como uma substância pensante.

Ao passo que, se apenas eu parasse de pensar, ainda que tudo o mais que imaginara fosse verdadeiro, não teria razão alguma de acreditar que eu existisse; por isso reconheci que eu era uma substância, cuja única essência ou natureza é pensar, e que, para existir, não necessita de nenhum lugar nem depende de coisa alguma material. (DESCARTES, 2007, p.60)

Ao alcançar esta conclusão, o conhecimento advindo da razão e o emprego da introspecção são privilegiados ao direcionarem o espírito humano na busca por ideias claras e distintas. Entretanto, ao evidenciar a distinção entre a natureza da mente (*espírito*) e do corpo, a proposta cartesiana além de necessitar explicar a relação causal entre corpo e mente, deve também demonstrar como podemos estar cientes da existência de “outras mentes”, dado que o método introspectivo permite analisar apenas o conteúdo da própria mente.

Para tanto, o autor salienta que a natureza distinta de cada uma destas substâncias não implica em uma simples relação de conjunção entre a mente e o corpo, “como um piloto em seu navio, mas, além disso, que lhe sou estreitamente conjunto e tão confundido e misturado que componho como que um único todo com ele” (DESCARTES, 2000, p. 122). Sendo esta a *primeira modalidade da união* da mente com o corpo, segundo a teoria cartesiana. Neste sentido, uma vez que a mente (*espírito*) não possui extensão alguma, esta também não informaria, tão pouco estaria localizada, em qualquer parte específica do corpo.

Entretanto, mesmo salientando a estreita união entre estas duas substâncias distintas, ainda resta à explicação sobre a relação causal entre elas. Sendo assim postulada a *segunda modalidade da união*, a qual afirma que a mente (*espírito*) medeia a sua relação com o corpo

por intermédio da *glândula pineal*³, localizada na base do cérebro, governando o movimento dos *espíritos animais*⁴.

Após apresentar por dedução a existência ontológica da mente imaterial, o passo seguinte seria a dedução de uma física baseada no *cogito*. Fato que o permitiu fundamentar a distinção da humanidade em relação aos outros animais e as máquinas, concebendo-os como *autômatos*⁵.

Descartes ao deduzir, pelo emprego de seu método, a distinção entre a substância pensante (alma imaterial) e o corpo físico, parte na elaboração de uma argumentação que permita atribuir ao ser humano uma posição privilegiada no estabelecimento das bases de sua física. O autor inaugura sua argumentação expondo a capacidade do coração, responsável pelo movimento corpóreo, de transmitir aos outros órgãos os *espíritos animais*. Segundo o autor, este argumento permite demonstrar “que os animais sem razão a nós se assemelham sem que por isso encontrasse nenhuma das que, dependentes do pensamento, são as únicas que nos pertencem enquanto homens” (DESCARTES, 2007, p.81).

Em resumo, os espíritos animais seriam transmitidos do cérebro aos músculos e ao coração, sendo transmitidos pelo sangue e pelos nervos, dando movimento ao corpo. Por meio desta explicação, Descartes pretende demonstrar duas teses principais: A primeira, que todo o funcionamento do corpo segue, exclusivamente, as leis da mecânica, dado que seriam elas também as leis da natureza. A segunda, seria que os humanos, sendo animais racionais compartilhariam com os outros animais as mesmas regras de funcionamento do corpo, todavia, uma vez dotados de razão, poderiam premeditar e deliberar suas ações, além de desfrutar da vida após a morte, em razão à imortalidade da alma.

Além disso, são apresenta duas formas de distinção entre ações mediadas pela razão de meros movimentos mecânicos realizados por autômatos. A primeira maneira seria o teste da linguagem, uma vez que máquinas e alguns animais poderiam proferir palavras ou até frases. Ao aplicar-se o teste, seria então verificado que o domínio do significado é

³ Há duas razões principais que levam Descartes a deduzir que a sede da alma seja esta glândula. A primeira delas é referente à duplicidade dos hemisférios cerebrais, assim como o resto dos órgãos dos sentidos. A segunda se refere às cavidades existentes, as quais permitiriam a mobilidade dos espíritos animais.

⁴ “Os espíritos animais são as partículas mais tênues do sangue; eles circulam continuamente através dos nervos, que Descartes considera dutos, e, de lá, vão para os músculos, cujos movimentos determinam. Sua produção pelo simples calor do coração tem de notável o fato de explicar por essa mesma causa todos os movimentos do animal” (GILSON, 2007, apud DESCARTES, 1989, p. 91).

⁵ Descartes concebe enquanto autômatos seres destituídos de substância pensante (ou mentes), sendo movidos por funções puramente mecânicas.

competência exclusiva dos seres humanos. O segundo modo de distinção seria mediante a análise do comportamento. Por vezes, alguns desses autômatos poderiam apresentar mais êxito na realização de tarefas específicas em relação aos humanos, mas falhariam em tantas outras, dado que apenas os humanos, enquanto possuidores de uma ferramenta universal como a razão, são capazes de uma ampla gama de êxito em ações diversas.

Imaginar que a alma dos animais seja da mesma natureza da nossa, e que, por conseguinte, nada temos a temer nem a esperar depois desta vida, como ocorre com as moscas e as formigas; ao passo que, quando se sabe o quanto elas diferem, compreendem-se muito melhor as razões que provam que a nossa é de uma natureza inteiramente independente do corpo e que, por conseguinte, não está sujeita a morrer com ele; depois por não vermos outras causas que a destruam, somos naturalmente levados a julgar que é imortal. (DESCARTES, 2007, p.99)

Concluindo esta Seção, é possível notar que o método cartesiano é imbuído de uma física amplamente mecanicista associado a uma metafísica que garante, além da exclusividade à atribuição de razão à espécie humana, a distinção ontológica entre mente e corpo.

Contudo, o compromisso de sua tese com esta distinção ontológica foi o principal alvo das mais diversas críticas no início do século XX, servindo de motivação à construção de teorias que evitassem este compromisso, reconfigurando a dualidade ontológica entre o mental e o físico. Na sessão seguinte serão apresentadas as principais reformulações contemporâneas do problema proposto por Descartes, introduzindo novos modos de explicar a controversa relação causal entre corpo e mente.

1.2 O problema mente/corpo: A retomada da Filosofia da mente no período contemporâneo

Uma das principais críticas direcionada ao dualismo substancial cartesiano é elaborada por Gilbert Ryle. Em sua crítica é argumentado que a teoria cartesiana pressupõe a existência de uma vida dupla, sendo uma pública e outra privada. A característica central da vida pública são os comportamentos corporais regidos por leis mecânicas, dotados de poder causal e passíveis de observações públicas. Enquanto os conteúdos da vida mental seriam privados, imateriais e destituídos de poder causal. Desse modo, nosso acesso as outras mentes seriam reduzidos a “inferências problemáticas” (RYLE, 2000, p.16) realizadas mediante a observação do comportamento, por meio de analogia ao próprio comportamento do observador. Por outro lado, nossos estados mentais, acessados por intermédio da faculdade da introspecção, seriam analisados de “maneira direta e autêntica dos atuais estados e operações de sua mente” (RYLE, 2000, p.14).

A crítica de Ryle a concepção cartesiana de mente e corpo, enquanto participantes da mesma categoria lógica, é subscrita enquanto um tipo de erro categorial. Conceber “os fatos da vida mental como se pertencessem a um tipo lógico ou categoria (ou domínio de tipos ou categorias), quando efetivamente pertence à outra categoria” (RYLE, 2000, p.17).

Um exemplo de erro categorial é o do estrangeiro, que ao comparecer pela primeira vez a Universidade de Oxford e após visitar os prédios, salas, biblioteca e laboratórios, ainda afirma manter o desejo de conhecer a Universidade, incidindo erroneamente ao pressupor que o conceito de Universidade participa da mesma categoria lógica que os conceitos de prédios, salas, biblioteca e laboratório. Contudo, é importante salientar que:

Estes exemplos de erros de categoria têm uma característica comum que deve ser notada. Os erros foram feitos por pessoas que não sabiam utilizar os conceitos [...]. A sua confusão surgiu da inabilidade de usar certos vocábulos de determinada língua. Os erros de categoria de interesse teórico são produzidos por pessoas perfeitamente competentes para aplicar conceitos, [...] mas que estão ainda sujeitas, nos seus pensamentos abstratos, a situar esses conceitos em tipos lógicos a que eles não pertencem. (RYLE, 2000, p.19)

O intuito é demonstrar que a teoria da vida dupla é fundada em um erro de categoria, representando a mente como um fantasma na máquina. Todavia, o autor é concordante com o fato da linguagem da física, química e fisiologia, serem insuficientes na descrição dos estados mentais, portanto, sendo sugerida sua descrição em uma linguagem paralela.

O autor alega que a ocorrência de erro categorial, na teoria cartesiana, tem como objetivo garantir os princípios mecânicos estabelecidos por Galileu sem entrar em conflito com os princípios teológicos dos escolásticos, incorrendo no emprego de propriedades mentais inversas as propriedades mecânicas. “Aderindo, ainda que involuntariamente, à gramática da mecânica, tentou evitar o desastre, descrevendo as mentes no que era meramente um vocabulário invertido” (RYLE, 2000, p.21).

Dessa maneira, ao evidenciar o erro na teoria cartesiana o autor busca descrever as consequências positivas de sua argumentação, afirmando que:

Se a minha argumentação for válida, terá interessantes conseqüências. Primeiro, o consagrado contraste entre Mente e Matéria será dissipado, não pelas igualmente consagradas absorções da Mente pela Matéria ou da Matéria pela Mente, mas sim de um modo bastante diferente. [...] A crença de que há uma oposição diametral entre Mente e Matéria vem da crença de que são termos do mesmo tipo lógico. Seguir-se-á também que tanto o Idealismo como o Materialismo são respostas a uma pergunta inadequada. A “redução” do mundo material a estados e processos mentais, assim como a “redução” dos estados e processos mentais a estados e processos físicos, pressupõe a legitimidade da disjunção. (RYLE, 2000, p. 24)

Alguns anos após a publicação do texto de Ryle, um de seus discípulos apresenta a defesa de outra abordagem, partindo da solução apresentada.

A teoria em questão é conhecida como *Materialismo reducionista* ou *Teoria da identidade mente/cérebro*. A *Teoria da identidade* postula o cérebro como o aparato físico de instanciação da mente, concentrando seus esforços em estabelecer relações de identidade entre os estados mentais e cerebrais. Contudo, o que está pressuposto no estabelecimento de identidade não se equivale a uma tradução entre conceitos mentais a conceitos cerebrais, mas em afirmar que ao se referir a um estado mental, refere-se, igualmente, a um estado cerebral.

Ullin T. Place apresenta à hipótese científica da consciência enquanto um processo cerebral, sugerindo que o *argumento lógico*⁶ sobre a identidade não seria suficiente para destituir a hipótese por completo. O autor ainda indica que a consciência poderia ser identificada como um padrão da atividade cerebral, uma vez que fosse possível explicar observações introspectivas subjetivas por meio de processos cerebrais correlatos.

Place afirma que seu objetivo não é afirmar que descrever uma sensação ou imagens mentais é o mesmo que descrever estados cerebrais, nem mesmo que afirmações sobre a

⁶ O argumento em questão afirma que as identificações de estados cerebrais com estados mentais gerariam afirmações literalmente ininteligíveis, ou em outras palavras “erros categoriais”. Desse modo, caso fosse encontrado uma única propriedade que fosse verdadeira para os estados cerebrais, mas falsa para estados mentais (ou vice-versa) a teoria da identidade seria refutada. (PLACE, 1956).

consciência são afirmações sobre estados cerebrais, seu objetivo é defender que “a afirmação, ‘a consciência é um estado cerebral’, embora não seja necessariamente verdadeira, também não seria necessariamente falsa” (PLACE, 1956, p.45).

Para isso, o autor argumenta, que a alegação de que a teoria da identidade poderia ser descartada por motivos lógicos, está baseada no erro de se considerar a relação de identidade expressada pelo verbo “é”, assumindo apenas um único sentido. O autor afirma que a aplicação do verbo permite conectar um objeto a sua definição, do mesmo modo, que pode ser aplicado conectando um objeto a descrição de sua composição. Assim, no primeiro caso a predicação é uma definição universal, ao se referir a um objeto à veracidade da definição é atribuída a todos os objetos dotados das propriedades descritas na definição. Já no segundo caso, a predicação é particular e o significado deve ser analisado mediante a observação, tanto ao objeto referido, quanto do predicado elencado.

No entanto, a expressão “a consciência é um processo cerebral” é o caso de uma expressão universal, a qual deveria ser aplicada a todos os casos de consciência. Assim, como seria possível argumentar em favor da veracidade desta expressão? Para isso, Place argumenta sobre a regra da gramática usualmente aplicada, onde havendo uma independência lógica entre duas expressões, deve haver também uma independência ontológica dos estados de coisas do qual eles se referem. Contudo, é afirmado que o caso da consciência é um dos poucos casos de exceção da regra. “As operações as quais têm de ser realizadas a fim de verificar dois conjuntos de características inerentes ao objeto ou estados de coisas em questão, raramente podem ser realizadas simultaneamente” (PLACE, 1956, p.46). Nestes casos, além de não ser possível verificar simultaneamente ambos os casos, também não podem ser analisados por meio de observação no sentido convencional, devido às distinções entre os níveis de análise. A introdução de níveis de análise distintos é relevante ao considerar o papel de hipóteses científicas de elementos não-observáveis, sendo compatíveis com uma teoria à medida que auxiliam na explicação e previsão do fenômeno, como ocorre no caso da consciência.

Place finaliza seu artigo argumentando que a principal dificuldade na compreensão de como processos introspectivos poderiam ser processos cerebrais, estão pautados sobre o que o autor chama de “falácia fenomenológica”, segundo o autor:

A falácia fenomenológica, na qual esse argumento é baseado, depende da suposição errada de que nossa habilidade de descrever coisas em nosso ambiente depende de nossa consciência deles, nossas descrições das coisas são primeiramente descrições

de nossas experiências conscientes, e posteriormente, descrições realizadas indireta e inferencialmente, dos objetos e eventos em nosso ambiente. (PLACE, 1956, p.49)

Entretanto, a denominada “falácia fenomenológica” criticada por Place, é fundamentada nas principais alegações sobre a irredutibilidade da consciência ao físico, elaborada pelos adeptos do *Dualismo de propriedade*.

Os adeptos ao dualismo de propriedade consideram esta uma abordagem participante do escopo das teorias *monistas*, ao passo que estão de acordo com a existência de uma única substância, a matéria física. A divergência entre o *Dualismo de propriedade* e o *fisicalismo* clássico é a sua reivindicação da existência de propriedades intrínsecas ao mental e não redutíveis a matéria física, como a consciência, a experiência subjetiva e os qualia.

*Uma defesa controversa ao dualismo de propriedade*⁷ é realizada por Thomas Nagel, em *What it is like to be a bat?* Neste artigo, Nagel apresenta o caráter subjetivo do ponto de vista de primeira pessoa, ou seja, *do que seria ser como um determinado organismo*. Argumentando que mesmo ao compreender toda a estrutura biológica, o funcionamento dos processos fisiológicos e os comportamentos de um organismo, tudo o que se pode conhecer sobre o aspecto subjetivo seria por meio de recursos mentais limitados as nossas próprias capacidades cognitivas. Podendo até mesmo supor como seria ser esse organismo, contudo, nunca poderíamos estar em posse de saber *o que é ser esse organismo, segundo a perspectiva do próprio organismo*. A escolha de morcegos é justificada da seguinte maneira:

Escolhi morcegos em vez de vespas ou linguados porque, se alguém viaja muito longe na árvore filogenética, as pessoas gradualmente perdem a fé de que existe alguma experiência lá. Os morcegos, embora mais intimamente relacionados conosco do que as outras espécies, apresentam uma gama de atividades e um aparato sensorial tão diferente do nosso que o problema que quero colocar é excepcionalmente vívido (embora certamente possa ser aplicado com outras espécies). (NAGEL, 1974, p.438)

Assim, este exemplo permite explorar as semelhanças e diferenças das experiências entre morcegos e humanos. Entre nós e os morcegos a principal semelhança está no fato de ambos sermos mamíferos. Contudo, as funcionalidades de nossos sistemas perceptivos são bem distintas, enquanto nos localizamos quase exclusivamente pela visão, os morcegos são dotados de ecolocalização. Para facilitar a compreensão, poderíamos usar uma metáfora e afirmar que seria *como se* eles pudessem enxergar através do som.

⁷ A subscrição de Thomas Nagel ao escopo do dualismo de propriedade é um tema de amplo debate. Parte dos críticos o concebe sob este escopo, enquanto alguns defensores afirmam a existência de uma importante distinção entre a existência de propriedades ontológicas irredutíveis e a distinção de perspectivas distintas (1º e 3º pessoa), considerando, portanto, sua teoria enquanto um dualismo de perspectiva.

Esta é a principal questão defendida por Nagel, pois apesar de podermos articular, por meio de metáforas e analogias, ou, até mesmo, replicarmos artificialmente o sistema perceptivo desses animais, como no caso dos sonares, nunca poderíamos compartilhar as mesmas experiências conscientes.

O que Nagel intenta na elaboração de sua argumentação é demonstrar que “é inútil basear a defesa do materialismo em qualquer análise de fenômenos mentais que não lide explicitamente com seu caráter subjetivo” (NAGEL, 1974, p. 437). Considerando a maneira que as abordagens fisicalistas elaboram suas teorias, o intuito de um programa reducionista é exatamente a elaboração de uma teoria mais objetiva, eximindo ao máximo a presença do aspecto subjetivo. No entanto, “a razão é que todo fenômeno subjetivo está essencialmente conectado a um único ponto de vista, e parece inevitável que uma teoria física objetiva abandone esse ponto de vista.” (NAGEL, 1974, p. 437).

Outro ponto relevante na argumentação de Nagel é de um lado, a distinção entre fatos presentes na realidade fenomênica de esquemas conceituais, e representações sistemáticas de outro. “Meu realismo sobre o domínio subjetivo em todas as suas formas implica uma crença na existência de fatos além do alcance dos conceitos humanos” (NAGEL, 1974, p. 441). Porém, o autor afirma que, o que é entendido por característica subjetiva da experiência, não é limitado ao acesso privado da experiência individual, nem que seria limitado a própria individualidade, uma vez que poderíamos descrever as qualidades da experiências. Entretanto, segundo Nagel, a característica essencial de um ponto de vista é:

Eles são subjetivos, no entanto, no sentido de que mesmo esta atribuição objetiva da experiência é possível apenas para alguém suficientemente semelhante ao objeto de atribuição, podendo, assim, adotar seu ponto de vista - para entender a atribuição na primeira pessoa e também em terceira. (NAGEL, 1974 p.442)

Desse modo, o autor conclui que se nos afastarmos temporariamente da relação entre mente e cérebro, poderíamos elaborar um entendimento mais objetivo baseando-nos apenas no mental. Sendo encarado como um estímulo para a construção de uma fenomenologia objetiva, independente do método da imaginação e da empatia.

A consideração sobre o aspecto subjetivo da consciência é uma das maneiras de se garantir o aspecto genuíno e exclusivo de capacidades cognitivas dos organismos, sendo distinguidas de capacidades cognitivas de sistemas cognitivos artificiais. Esse modo de distinguir capacidades cognitivas genuínas de derivadas, torna-se mais evidente nas

afirmações do *Naturalismo biológico* de Searl, como veremos em maiores detalhes no próximo capítulo.

Em meio às objeções tanto da abordagem reducionista, quanto do dualismo de propriedade, a segunda metade do século XX presenciou a emergência de uma nova abordagem na análise da natureza da mente, conhecida como *Materialismo eliminacionista*.

O eliminacionismo compreende que os conceitos mentais, também denominados de conceitos da *psicologia popular*⁸, mantêm seu uso em meio às teorias psicológicas devido ao elevado grau de complexidade de seu objeto de estudo, mas que, no entanto, estes conceitos não apresentariam qualquer referente físico capaz de abarcar as exigências necessárias a uma redução interteórica. Postulando, portanto, que, em vez de uma redução interteórica, os conceitos mentais deveriam submeter-se a uma eliminação por completo, priorizando conceitos da estrutura neurofisiológica.

Os principais argumentos em favor do eliminacionismo são os seguintes: Primeiro, é argumentado o fracasso da psicologia popular em termos de explicação, previsão e controle dos fenômenos mentais. “Em resumo, as coisas mais fundamentais sobre nós permanecem inteiramente misteriosas no âmbito da psicologia popular” (CHURCHLAND, 2004, p.83).

A segunda razão em favor do eliminacionismo argumenta sobre a tentativa de “extrair uma lição indutiva dos conceitos” (CHURCHLAND, 2004, p. 84). São apresentados casos onde conceitos de teorias populares apresentam inexatidão em relação aos fenômenos explicados, culminando na sua eliminação em detrimento a uma teoria mais sofisticada.

O terceiro argumento é construído afirmando que o eliminacionismo teria uma vantagem *a priori*, uma vez que considera negativa a possibilidade de corroboração dos conceitos da psicologia popular com uma neurociência amadurecida, diferentemente da teoria da identidade e do funcionalismo.

Até este ponto, todas as abordagens sobre o problema mente/corpo, com exceção do behaviorismo filosófico, argumentam favorecendo a relação entre cérebro e estados mentais.

Entretanto, a última abordagem apresentada destitui o cérebro de sua posição privilegiada, argumentando que ao definirmos a mente pela sua *função*, podemos admitir que esta possa ser instanciada nos mais diversos tipos de sistemas cognitivos. “É também com base nesta tese de que estados mentais se definem pelo seu papel funcional que um sistema

⁸ O termo psicologia popular se refere aos termos psicológicos de *senso comum* utilizados no cotidiano para explicar e prever o comportamento. Crenças, desejos, intenções e vontades são exemplos de termos da psicologia popular.

pode apresentar predicados mentais *independentemente* do tipo de substrato físico do qual eles poderiam eventualmente resultar” (TEIXEIRA, 1998, p.49, grifo nosso)

A teoria em questão é conhecida como *Funcionalismo* e foi incitada, principalmente, pelo artigo de Alan M. Turing *Computing Machinery and Intelligence*. O artigo de Turing é um dos primeiros a propor a construção de uma Inteligência Artificial. Turing sugere que poderíamos substituir a questão “Poderia uma máquina pensar?” ao nos questionarmos, sobre a possibilidade de construção de um computador digital de estados finitos, programado com uma grande, ou finita, tabela de instruções, a qual poderia levar um interrogador a acreditar tratar-se de um ser humano.

A característica central da abordagem funcionalista, é que, o que constitui um estado mental, não é dependente de seus estados internos, ao invés disso, o que constitui um estado mental é a *função*, ou o papel funcional que desempenha, no sistema do qual faz parte. O principal argumento da teoria funcionalista contra as *teorias naturalistas tradicionais*⁹ é o argumento da *múltipla realizabilidade*. O argumento admite que um estado mental, pode ser instanciado em diferentes tipos de sistemas físicos, sejam eles naturais ou artificiais. Desse modo, as teorias que apresentam uma descrição detalhada da estrutura neurofisiológica, não desvelariam a natureza fundamental da mente, dado que esta seria apenas uma dentre as possibilidades de instanciação das funções mentais.

O sucesso da teoria *funcionalista* no desenvolvimento da Ciência Cognitiva tradicional ocorreu devido à aplicação do dispositivo teórico conhecido como *Máquina de Turing*¹⁰, sendo este o princípio de qualquer computador moderno. A construção dessa nova ferramenta foi motivo de grande entusiasmo pelos cientistas cognitivos, sendo pressuposto o seu emprego como modelo “mental”. A possibilidade de manipular um modelo “mental” artificial presumia uma vantajosa ferramenta epistêmica, possibilitando avanços na compreensão da mente.

Hilary Putnam, um obstinado defensor da teoria funcionalista, em seu texto de 1960, intitulado *Minds and Machines*. Argumenta que em geral as questões que fundamentam o

⁹ Cabe aqui incluir uma ressalva quanto às abordagens naturalista que são criticadas pela teoria funcionalista. O que denomino como *naturalismo tradicional* engloba, segundo a proposta de Abrantes (2004), o naturalismo nato e o naturalismo metodológico. Enquanto o naturalismo conceitual-linguístico, também subscrito em meio aos ditos “naturalismo”, não seria criticado pela múltipla realizabilidade do mental, proposto pela abordagem funcionalista.

¹⁰ “As máquinas de Turing, descritas pela primeira vez por Alan Turing 1936/1937, são dispositivos computacionais abstratos simples destinados a ajudar a investigar a extensão e as limitações do que pode ser calculado. As “máquinas automáticas” de Turing, como ele as denominou em 1936, foram concebidas especificamente para o cálculo de números reais. Elas foram chamadas pela primeira vez de “máquinas de Turing” por Alonzo Church em uma revisão do artigo de Turing (Church 1937). Hoje, eles são considerados um dos modelos fundamentais da computabilidade e da ciência da computação (teórica)” (DE MOL, 2019, p. 1).

problema mente/corpo são de natureza lógica e linguística, todavia, alguns teóricos afirmam que “fatos” empíricos poderiam auxiliar de algum modo a resolução desta questão. Entretanto, Putnam afirma que o fato de computadores responderem questões a partir de sua própria estrutura, não teria qualquer relação com a subjetividade da experiência humana. Com o intuito de construir uma argumentação sobre o problema ontológico comparando as abordagens concorrentes com máquinas de Turing. O autor atinge a seguinte conclusão:

A moral, acredito, é bastante clara; não é mais possível acreditar que o problema mente-corpo é um genuíno problema teórico ou que uma "solução" lançaria a menor luz sobre o mundo em que vivemos. [...] Mas se o chamado "problema mente-corpo" nada mais é do que uma realização diferente do mesmo conjunto de questões lógicas e linguísticas, então deve ser tão vazio quanto verbal. [...] Nesse caso, o *insight* leva em consideração que qualquer conclusão que possa ser alcançada no caso do problema mente-corpo teria que ser alcançada, e pelas mesmas razões, no caso da máquina de Turing. (PUTNAM, 1960, p. 32)

A conclusão dada por Putnam considera que a noção de máquina de Turing é uma excelente analogia, constituindo-se em um modelo suficiente para compreensão da relação mente-cérebro. Desse modo, a resolução do problema ontológico da mente seria fundamentada por “um psicoparalelismo que dispensaria qualquer tipo de pressuposição metafísica responsável pela possibilidade de interação entre o físico e o mental” (TEIXEIRA, 1998, p.49). A consideração de um psicoparalelismo entre *software* (mente) e *hardware* (cérebro) pressupõe uma *independência* do primeiro em relação ao segundo, sendo esta a base para o argumento da múltipla realizabilidade.

Na seção seguinte será apresentado o contexto que possibilitou a emergência da Ciência Cognitiva, além de ser examinada a relevância dada pela abordagem funcionalista à modelagem computacional, bem como aos conceitos de representação e múltipla realizabilidade do mental.

1.3 Ciência Cognitiva: Interdisciplinaridade, Paradigmas e Informação

A primeira metade do século XX, antes mesmo da invenção do termo “Inteligência Artificial”, o movimento predominante na busca por um modelo mecanicista da mente era denominado de *cibernética*¹¹. “Os ciberneticistas acreditavam que toda atividade psicológica humana poderia um dia ser estudada por meio de modelos matemáticos. [...] Seu ponto de partida baseava-se na possibilidade de criar circuitos elétricos que pudessem modelar o funcionamento do cérebro” (TEIXEIRA, 1998, p.35). Em meio aos anos cibernéticos, grandes teóricos se sobressaem na fundamentação destas ideias, são eles: Claude Shannon, Warren McCulloch e Walter Pitts.

A contribuição de Shannon para os desenvolvimentos iniciais da cibernética foi sua defesa sobre a ideia de que circuitos elétricos poderiam expressar operações lógicas.

Qualquer circuito é representado por um conjunto de equações, os termos das equações correspondentes aos vários reles e interruptores no circuito. Um cálculo é desenvolvido para manipular essas equações por processos matemáticos simples, a maioria dos quais é semelhante a algoritmos algébricos comuns. Este cálculo mostra-se exatamente análogo ao cálculo de proposições usadas no estudo simbólico da lógica. (SHANNON, 1938, p. 471)

Por outro lado, McCulloch e Pitts propuseram a construção de um modelo de neurônio artificial, que permitisse simular o funcionamento do cérebro. Pressuposto a construção deste modelo, estava a possibilidade de instanciar o raciocínio lógico em um dispositivo artificial.

A lei “tudo ou nada” da atividade nervosa é suficiente para garantir que a atividade de qualquer neurônio possa ser representada como uma proposição. As relações fisiológicas existentes entre as atividades nervosas correspondem, é claro, às relações entre as proposições; e a utilidade da representação depende da identidade dessas relações com as da lógica das proposições. Para cada reação de qualquer neurônio, há uma afirmação correspondente de uma proposição simples. (McCULLOCH; PITTS, 1943, p. 117)

Passados os anos cibernéticos, a interdisciplinaridade, uma característica emergente do conhecimento científico contemporâneo, torna-se mais evidente, principalmente, após a segunda metade do século XX. Os principais fatores foram “o peso esmagador dos problemas sociais e tecnológicos, avanços na pesquisa, novas bolsas de estudos e novas demandas no currículo que colocaram um slogan nas décadas de 1960 e 1970 na agenda acadêmica”

¹¹ Cibernética é o nome dado ao programa de pesquisa precursor da Inteligência Artificial, seu principal pressuposto é que a melhor maneira de se estudar a psicologia humana seria por meio de elaboração de modelos matemáticos, assim como tantas outras áreas das ciências naturais. (TEIXEIRA, 1998)

(KLEIN, 1996, p.1). Outra característica das práticas interdisciplinares é seu conjunto complexo de interação social entre pesquisadores das mais diversas localidades.

Resumidamente, sistemas simples operam de acordo com um único conjunto de regras, mesmo que possuam vários níveis e conexões organizadas em uma hierarquia. Sistemas complicados são variações sobre os temas de sistemas simples. Sistemas complexos, por outro lado, possuem estruturas não hierárquicas. Eles obedecem a múltiplas lógicas conflitantes, utilizam *feedback* positivo e negativo, revelam efeitos sinérgico e podem ter um elemento caótico. [...] Para entender o que está acontecendo, devemos substituir o pensamento reducionista por pensamento não linear, reconhecimento de padrões e analogia. As atividades podem ser interconectadas em uma matriz móvel, repleta de *loops* de *feedback* e relações sinérgicas imprevisíveis em qualquer matriz de contextos aninhados. (KLEIN, 1996, p.20)

É a partir deste contexto que emergem as Ciências Cognitivas, sendo esta um conjunto complexo de interação entre disciplinas como Filosofia, Linguística, Antropologia, Neurociência, Inteligência Artificial e Psicologia. Nascidas em meio a emergência do paradigma naturalista e do desenvolvimento tecnológico informacional, a Ciência Cognitiva considera a cognição, ou seja, os fenômenos mentais, enquanto fenômenos naturais.

As ideias de McCulloch e Pitts deram origem ao que posteriormente ficaria conhecido como *paradigma conexionista*. Este protótipo de neurônio artificial foi progressivamente desenvolvido ao longo da década de 50, até surgirem as críticas em relação a restrita gama de computações que esse modelo poderia efetuar. “Nos primeiros computadores, [...] as instruções ou programa do computador e os dados sobre os quais elas incidiam eram coisas distintas. As instruções tinham que ser [...] parte do *hardware* da máquina” (TEIXEIRA, 1998, p.42). Esta relação de dependência entre *software* e *hardware* geravam limitações ao modelo, até ser, posteriormente, proposto por J. von Neumann que era possível incluir no mesmo plano instruções e dados, possibilitando uma relação de independência entre *software* e *hardware*. “Isto foi uma imensa revolução em termos de como conceber a própria arquitetura dos computadores” (TEIXEIRA, 1998, p.43).

Foi esta revolução que permitiu os pesquisadores vislumbrarem o futuro promissor da *Inteligência Artificial Simbólica*¹². Devido a possibilidade de conceber *software* e *hardware* de maneira independente, permitindo que um mesmo *hardware* pudesse rodar programas

¹² Refere-se à ideia de que seria possível a construção de *softwares* específicos para operar sobre dados ou representações. Basicamente, é a concepção de que a inteligência é o resultado de uma sequência apropriada de representações mentais, ou símbolos. Esta concepção foi responsável pela forte influência da abordagem funcionalista na Filosofia da Mente, a qual concebia a mente como um programa computacional. (TEIXEIRA, 1998)

distintos. Desse modo, o paradigma simbólico, também conhecido como cognitivismo, compreende que “a inteligência resulta do encadeamento adequado de representações mentais – que nada mais seriam do que símbolos. *A mente é um programa computacional*, sua replicação depende de encontrar um programa computacional que permita simulá-la” (TEIXEIRA, 1998, p.43). Concebendo, então, que a inteligência é resultado da capacidade de representação mental. Segundo essa visão, é a capacidade de manipulação de símbolos que distingue o seres humanos dos outros animais, concebida como uma atividade cognitiva superior, a qual permite a resolução de problemas matemáticos e a comunicação através da linguagem.

Daí a ideia que prevaleceu nos anos áureos da Inteligência Artificial (anos 70): *a máquina de Turing é o melhor modelo da atividade mental humana*. Pensar nada mais é do que realizar computações, uma seguida da outra. [...] Pensar é encadear proposições por meio de conectivos lógicos e usar isto de forma a resolver problemas, isto é, de forma algorítmica. (TEIXEIRA, 1998, p.44)

O paradigma simbólico permitiu às Ciências Cognitivas responderem a dois desafios na elaboração de um programa científico para o estudo da mente. O primeiro têm relação com o desafio lançado a todo conhecimento pretende a cientificidade: a formalização. O segundo indica a objetividade requisitada à observação científica, superando então a tradição behaviorista¹³, ao realizar abstrações dos conteúdos mentais, testando-os empiricamente por meio de modelos computacionais.

Entretanto, a prática interdisciplinar das ciências cognitivas permitiram, além da transposição de fronteiras entre disciplinas, mudanças paradigmáticas internas, indo do cognitivismo, “segundo o qual a ‘intencionalidade’ é de tipo simbólico: a mente é vista como um sistema cognitivo que processa informação de maneira sequencial, como uma série de operações de manipulação de símbolos por meio de regras sintáticas, conformes ao formalismo lógico.” (WOLFF, 2012, p.123), ao connexionismo, segundo o qual “os fenômenos mentais podem ser descritos por meio de redes de unidades simples interconectadas. [...] Podemos, assim modelizar operações cognitivas por intermédio de redes de neurônios artificiais inspirados nos neurônios ‘de verdade’”(WOLFF, 2012, p.123).

O desenvolvimento de neurônios artificiais iniciados pelo estudos de Shannon, McCulloch e Pitts mantiveram-se abafados durante toda a década de 70, sendo retomados

¹³ A tradição behaviorista desconsidera quaisquer estados mentais, substituindo a mente, como objeto de estudo da psicologia, pela análise do comportamento.

após a crítica do paradigma cognitivista, levando a uma retomada do conexionismo na Ciência Cognitiva. Seu desenvolvimento só foi retomado no início dos anos 80 devido ao “descontentamento com os modelos simbólicos, a tentativa de reaproximar a abordagem da cognição com a pesquisa em neurociência que começava a ganhar cada vez mais força e o aperfeiçoamento de novas técnicas para treinar redes de múltiplas camadas” (TEIXEIRA, 1998, p.89). No paradigma conexionista, o foco da análise muda da mente, entendida como manipulação de símbolos segundo regras, para o funcionamento cerebral, compreendido, enquanto um dispositivo computacional em paralelo, o objetivo é o processamento de informação ocorrido mediante a um conjunto de relações causais entre unidades que podem ser excitadas ou inibidas.

Abandona-se a ideia de uma mente que executa passos algorítmicos discretos (como uma máquina de Turing) e a suposição de que processos mentais seriam uma justaposição inferencial de raciocínios lógicos. Em vez, o que temos é um conjunto de neurônios artificiais para modelar a cognição; neurônios cujo peso de conexão sináptica pode ser alterado através de estimulação positiva ou negativa da conexão. (TEIXEIRA, 1998, p.84)

Os modelos conexionistas desenvolvidos após a década de 80 apresentavam algumas vantagens em relação ao modelos computacionais do paradigma cognitivista. A primeira delas é referente à sua proximidade com a realidade biológica, concebendo o cérebro biológico como fundamento para sua elaboração. Outra vantagem, foi com relação a sua distribuição do processamento em paralelo, permitindo que a informação seja distribuída por todo o sistema. Tal distribuição é vantajosa na medida em que permite que a informação armazenada possa subsistir, mesmo no caso de mal funcionamento de alguns neurônios. Desse modo, não se faz necessário que o significado das operações esteja localizado em alguma unidade específica, dado que este emerge do estado global do sistema.

Como este estado global emerge de uma rede de unidades mais básicas que os próprios símbolos, o paradigma conexionista é também chamado de “paradigma subsimbólico”. O nível subsimbólico superpõe-se ao biológico, embora esteja mais próximo deste do que a IA simbólica. No paradigma subsimbólico, o significado emerge de complexos padrões de atividades *sem*, entretanto, *residir especificamente em cada um dos componentes da rede* (TEIXEIRA, 1998, p.90, grifo nosso)

Dessa maneira, a utilização dessa modelagem comparativa máquina-humano e humano-máquina, permitiu o estabelecimento de uma rede dinâmica e autogerativa na transposição de fronteiras entre diversas áreas do conhecimento. Segundo Wolff:

O círculo é fechado: para pensar a mente humana, devemos pensá-la a partir do modelo da máquina, que pode, por sua vez, ser pensada a partir do modelo do cérebro. Feito isso, todas as fronteiras são transpostas – e em ambos os sentidos: na ida, da mente ao cérebro (naturalização da intencionalidade), do homem ao animal (via mente-cérebro), da natureza ao artificial (via o modelo computacional) e, na volta, do artificial à natureza (via o modelo conexionista), depois do animal ao homem (via teorias evolucionistas), do cérebro à mente (via localizações cerebrais), do mental ao social (via teorias epidemiológicas da transmissão das ideias) etc. (WOLFF, 2012, p.124)

A origem das Ciências Cognitivas é quase concomitante a apropriação do conceito de informação na filosofia, estabelecendo uma relação intrínseca entre sua emergência e a própria especulação do conceito. Tal concomitância, pode ser explicada pela recorrência nos debates sobre cognição definindo fluxo de pensamento, como um processo informacional.

Umas das primeiras tentativas de formalização do conceito de informação na contemporaneidade foi proposto por Shannon e Weaver, conhecida como Teoria Matemática da Comunicação, a qual concebe a comunicação como um processo de transmissão de mensagens. O modelo teórico proposto seria uma troca unidirecional de mensagens entre uma fonte e um destino, utilizando-se de um canal de comunicação como meio para a distribuição dos sinais, considerando-se a possibilidade de interferência na transmissão da mensagem, denominado ruído. A informação é compreendida por esta teoria como uma entidade independente do conteúdo semântico, sendo um meio de mensurar a autonomia da escolha na seleção de uma mensagem.

O conceito de informação adquiriu um lugar privilegiado em meio as Ciências Cognitivas, principalmente com a emergência do paradigma computacional cognitivista. “A hipótese central da ciência cognitiva é que pensar pode ser melhor compreendido em termos de estruturas representacionais na mente e processos computacionais que operam sobre aquelas estruturas” (THAGARD, 2019, p.5). A noção de estrutura ou conteúdo representacional é importante, pois acarretou o interesse por uma conceitualização do termo informação mais geral do que a proposta de quantidade mensurável de Shannon.

Apesar da ampla gama de abordagens teóricas nas Ciências Cognitivas o papel da informação pode ser observado em muitas delas, além da computacional cognitiva, como por exemplo, na abordagem Bayesiana¹⁴ e no *Deep Learning*¹⁵.

¹⁴ “A abordagem bayesiana assume que a cognição é, aproximadamente, ótima de acordo com a teoria da probabilidade, especialmente o teorema de Bayes, que diz que a probabilidade de uma hipótese, dada sua evidência, é igual ao resultado da multiplicação da probabilidade anterior da hipótese pela probabilidade condicional da evidência dada a hipótese, tudo dividido pela probabilidade da evidência”. (THAGARD, 2019, p.10)

Entretanto, os desenvolvimentos atuais em ciência cognitiva demonstram que devido às analogias entre humano-máquina, o emprego do conceito de informação é utilizado seguindo dois objetivos distintos. Um deles é formal e refere-se ao aprimoramento tecnológico, desenvolvimento de algoritmos robustos capazes de processar quantidades imensas de dados, mineração de dados, organização do conhecimento científico, desenvolvimento no campo da inteligência artificial e *Deep Learning*. O outro é empírico e se refere ao papel biológico da informação, aos meios pelos quais sistemas cognitivos naturais processam a informação disponível no ambiente, o papel do cérebro no processamento de informação e sua relação com o comportamento.

Na seção seguinte, realizo a apresentação do termo Representação, termo evidentemente caro a Ciência Cognitiva, bem como algumas críticas a tal noção, as quais são direcionadas por uma das mais recentes correntes teóricas da Ciência Cognitiva, a abordagem da Cognição incorporada. A apresentação deste termo se faz necessária a presente pesquisa, uma vez que este é um conceito fundamental na compreensão das teorias sobre os fenômenos da Consciência e da Intencionalidade, as quais serão discutidas em maiores detalhes no Capítulo 2.

¹⁵ “O *deep learning* se baseia em ideias do connexionismo e da neurociência teórica, mas usa redes neurais com mais camadas e algoritmos aprimorados, beneficiando-se de computadores mais rápidos e grandes bancos de dados de exemplos. [...] Idéias de *deep learning* estão se espalhando de volta para a neurociência e também começando a influenciar as pesquisas em psicologia cognitiva”. (THAGARD, 2019, p.10)

1.4 A noção de Representação e a Cognição Incorporada e Situada.

Neste capítulo o objetivo geral é apresentar de que modo os estudos contemporâneos sobre a Consciência e a Intencionalidade podem ser combinados com o debate sobre as tecnologias de *Big Data*, a fim de gerar novas compreensões científicas sobre estes três fenômenos. No entanto, antes de nos encaminharmos diretamente para estas questões, uma explicitação da noção de Representação se faz necessária, tendo em vista que este termo é amplamente utilizado pelas teorias orientadoras do debate em questão.

Assim como outros tópicos, a noção de Representação é um tema controverso tanto na filosofia da mente como na ciência cognitiva. Por este motivo, cabe nesta seção uma tentativa de explicitar alguns posicionamentos sobre o tema, bem como algumas soluções propostas.

Em um influente livro de introdução a ciência cognitiva, publicado ainda na década de noventa, Paul Thagard apresenta logo no primeiro capítulo algumas considerações sobre o termo Representação.

A maioria dos cientistas cognitivos concorda que o conhecimento dentro da mente consiste em *representações mentais*. [...] Os cientistas cognitivos propuseram vários tipos de representação mental como regras, conceitos, imagens e analogias. [...] A ciência cognitiva propõe que as pessoas têm *procedimentos* mentais que operam nas representações mentais para produzir pensamentos e ações. Tipos diferentes de representações mentais, como regras e conceitos, promovem tipos diferentes de procedimentos mentais. (THAGARD, 1998, p.16-17)

Entretanto, a polêmica envolvendo a noção de Representação não se limita as críticas elaboradas pelos adeptos do anti-representacionalismo, mas estende-se a conflitos internos entre os adeptos da abordagem representacionalista, sobre qual seria o tipo de representação mental mais eficaz. Sobre os conflitos internos, Thagard (1998) afirma que a adoção de uma abordagem representacional *eclética*, tratando os distintos modos de representação como complementares em vez de um modo competitivo, traria mais vantagem à ciência cognitiva na compreensão da mente e dos procedimentos mentais.

Desse modo, seguindo as considerações do autor, ser adepto de uma abordagem representacionalista é o mesmo que ser concordante com a seguinte hipótese: “o pensamento pode melhor ser entendido em termos de estruturas representacionais na mente e procedimentos computacionais que operam naquelas estruturas” (THAGARD, 1998, p. 21). Denominando de CRUM (*Computational-Representational Understanding of Mind*) o tipo de

abordagem eclética que considera a relevância da hipótese, bem como os vários tipos de representação mental.

No entanto, a fim de avaliar o grau de eficiência na explicação do pensamento, cinco critérios complexos devem ser aplicados as formas de representação mental. Sendo eles: 1) poder representacional; 2) poder computacional; 3) plausibilidade psicológica; 4) plausibilidade neurológica; e 5) aplicação prática. O quadro abaixo lista os cinco critérios em adição a alguns subcritérios avaliativos.

Quadro 1 – Critérios de avaliação de teorias de representação mental

1. Poder de representação	3. Plausibilidade psicológica
2. Poder computacional	4. Plausibilidade neurológica
a. Resolução de problemas	5. Aplicação prática
I. Planejamento	a. Educação
II. Decisão	b. Desenho
III. Explicação	c. Sistemas inteligentes
b. Aprendizado	
c. Linguagem	

Fonte: Paul Thagard (1998)

Entretanto, uma descrição e análise profunda sobre os tipos de representação mental, assim como seus critérios de avaliação, fogem ao objetivo proposto por esta seção, podendo ser verificado, em Andrade (2021, no prelo), uma descrição mais completa. O que nos é interessante é que apesar dos vários desafios direcionados a esta abordagem “direi que nenhum dos desafios oferece uma razão para que a CRUM seja abandonada. Entretanto, muitos deles mostram que a CRUM necessita ser expandida e suplementada, principalmente em questões que a integram a fatores biológicos e sociais” (THAGARD, 1998, p. 136). Fato que parece ter guiado as pesquisas posteriores de Thagard, como observado pelas publicações mais recentes do autor.

Contudo, um esclarecimento sobre a noção de Representação é evidentemente apresentado em seu artigo *How Brains Make Mental Models*, publicado em 2010.

Muitos pesquisadores recentes usaram modelos mentais para explicar aspectos do pensamento, incluindo resolução de problemas, aprendizado indutivo e interação homem-máquina. Centenas de artigos psicológicos foram publicados sobre modelos mentais. No entanto, a natureza dos modelos mentais permaneceu bastante confusa. [...] Mas qual é a natureza exata das representações psicológicas que podem preservar as restrições da maneira exigida? [...] Este artigo faz uma nova abordagem para desenvolver a noção vaga, mas fértil, de modelos mentais, caracterizando-os em termos de processos neurais. Uma abordagem neural vai contra a suposição de modeladores mentais como Johnson-Laird e Craik de que a explicação psicológica pode prosseguir em um nível funcional e computacional abstrato, mas tentarei

mostrar as vantagens de operar tanto no nível neural quanto no psicológico. (THAGARD, 2010, p. 448-449)

Desse modo, ao considerarmos a complexa composição do cérebro, notamos que é constituída por bilhões de neurônios, cada um deles com milhares de conexões com outros neurônios, estabelecendo entre si conexões excitatórias (o disparo de um neurônio aumenta a taxa de disparo de outro neurônio ligado a ele) e inibitórias (o disparo de um neurônio diminui a taxa de disparo de outro neurônio conectado a ele). Denominadas de população neuronal, os conjuntos de neurônios que apresentam ampla taxa de comunicação entre eles.

Além da composição cerebral, deve ser levada em conta, a comunicação dinâmica em paralelo entre as células neuronais e não-neuronais, a qual é responsável por modos de inferência muito além das capacidades apresentadas pelos computadores.

Assim, uma população neuronal é capaz de representar algo por meio de seus padrões de disparo. “Um padrão de ativação no cérebro constitui uma representação de algo quando há uma correlação causal estável entre o disparo de neurônios em uma população e a coisa que é representada, como um objeto ou grupo de objetos no mundo” (THAGARD, 2010, p. 450). A maior parte das evidências que suportam essa afirmação é advinda de técnicas experimentais de varredura cerebral e das teorias sobre como os cérebros funcionam.

No entanto, o tipo de correlação causal entre o padrão de ativação do cérebro e os objetos do mundo deve ser capaz de demonstrar que eles estabelecem a mesma estrutura relacional, tanto estática quanto dinamicamente. “Os modelos mentais estáticos têm uma estrutura espacial semelhante ao que representam, enquanto os modelos mentais dinâmicos têm uma estrutura temporal semelhante” (THAGARD, 2010, p. 451).

Os modelos estáticos são adquiridos por meio de mapas sensoriais topográficos. Dado que o sistema nervoso realiza computações para processar a informação biologicamente relevante, estas computações podem ser mapeadas pela correlação entre o ajuste dos valores de um determinado parâmetro com a taxa de comunicação entre uma população neuronal e uma população vizinha. “Nesse caso, a organização espacial dos neurônios corresponde sistematicamente à organização espacial do mundo, da mesma forma que a localização das grandes cidades em um mapa do Brasil corresponde à localização real dessas cidades” (THAGARD, 2010, p. 451).

Por outro lado, os modelos dinâmicos temporais são bem menos pesquisados, mas o autor explora duas possibilidades. A primeira é relativa a codificação de relações temporais

executado por redes recorrentes, ou seja, são redes de neurônios artificiais onde os neurônios de saída estabelecem uma relação de *feedback* com os neurônios de entrada, produzindo uma rede temporal que pode reter informações. Partindo dessa ideia, a segunda possibilidade, é que um padrão de ativação poderia, então, codificar um estado de coisa por meio de uma coleção de regras *se-então*, por meio de um tipo de representação puramente neural. Assim, “executar um padrão temporal é uma questão de disparar uma sequência de regras, não pela correspondência verbal usual empregada por arquiteturas cognitivas baseadas em regras, [...] mas por operações de rede puramente neural” (THAGARD, 2010, p. 453).

Contudo, muito recentemente, a ciência cognitiva contemporânea baseada sobre algumas abordagem da cognição incorporada direcionou várias críticas a noção clássica de Representação mental. Para ser mais preciso, desde a segunda metade do século XX, autores como J.J. Gibson (1966; 1979) defendem um posicionamento no qual a cognição é entendida como um processo destituído de representações e computações. A abordagem de Gibson, conhecida como *Psicologia ecológica* compreende que:

Os movimentos de um organismo criam um padrão de estimulação em constante mudança, no qual características invariáveis vêm à tona. A detecção desses invariantes, segundo a psicologia ecológica, fornece todas as informações necessárias para a percepção. A percepção da forma de um objeto, por exemplo, torna-se aparente como resultado da detecção dos tipos de transformações no padrão de estímulo que ocorrem ao se aproximar ou mover ao redor do objeto (SHAPIRO; SPAULDING, 2021, p. 2).

Já a cognição incorporada, em um sentido geral, apresenta em sua composição influências, tanto de abordagens descomprometidas com o representacionalismo/computacionalismo, quanto com abordagens connexionistas, sendo, na maioria dos casos, concordantes com a ideia de que “a cognição é incorporada, [...] quando aspectos do corpo do agente além do cérebro desempenham um papel causal ou fisicamente constitutivo significativo no processamento cognitivo” (WILSON, 2015, p.1). Assim, as críticas a noção clássica de Representação mental são direcionadas sobre três princípios da ciência cognitiva tradicional: 1) As representações mentais são autônomas em relação ao sistema perceptual, a ação corporal e aos detalhes operacionais, uma vez que a informação transmitida por uma representação não exhibe características específicas da modalidade. 2) O conhecimento é organizado proposicionalmente, e seu significado é emergente da relação com outros símbolos internos. 3) As representações internas instruem programas motores, essencialmente separados e independentes da cognição.

Uma das posturas mais radicais subscritas entre os adeptos do anti-representacionalismo é a teoria dos sistemas dinâmicos, os quais negam ou até mesmo minimizam a necessidade de um processador representacional centralizado. Sendo defendido que “enquanto uma criatura situada puder sentir seu mundo de modo a permitir que seu corpo seja diretamente influenciado, as descrições simbólicas abstratas podem ser dispensadas” (WILSON, 2017, p. 4). No entanto, a defesa de um posicionamento essencialmente anti-representacionista é envolto em dificuldades, como por exemplo, na explicação do comportamento antecipatório para a satisfação de necessidades, objetivos ou condições futuras.

Além disso, não está claro por que a ciência cognitiva incorporada também não poderia ser simbólica, representacional, abstrata, etc. A perplexidade aqui é ampliada pelo fato de que muitas abordagens auto-denominadas incorporadas à cognição são simbólicas, representacionais, abstratas, etc. A oferta são visões específicas do que são símbolos e representações mentais e como funcionam em nossa economia cognitiva. Normalmente, eles substituem a concepção de codificação proposicional por uma de acordo com a qual os símbolos transmitem uma característica específica da modalidade. (WILSON, 2017, p. 4)

Assim como a citação acima destaca, nem todas as teorias sobre o escopo da cognição incorporada, ou, da teoria dos sistemas dinâmicos, estariam definitivamente comprometidas com o anti-representacionalismo. Apenas como exemplos, podemos citar a abordagem enativa de Francisco Varela e os trabalhos com sistemas dinâmicos de Alicia Juarrero, os quais mantêm-se comprometidos, em algum sentido, com as próprias reformulações da noção de representação.

Retomando a proposta de Paul Thagard sobre o modo como podemos endossar a noção de Representação, ao final do artigo ainda é afirmado que sua abordagem é compatível com abordagens mais moderadas dentro do escopo da cognição incorporada.

Rejeito enfaticamente a tese da corporificação extrema de que pensar é apenas ação corporificada e, portanto, incompatível com abordagens computacionais-representacionais de como o cérebro funciona. Argumento abaixo que mesmo o controle motor requer um alto grau de representação e computação. Muito mais plausível é a tese da corporificação moderada de que a linguagem e o pensamento são inextricavelmente moldados pela ação corporificada. [...] Nessa visão, pensar ainda requer representações e cálculos, mas a natureza particular destes depende em parte do tipo de corpo que as pessoas têm, incluindo suas capacidades sensoriais e motoras. (THAGARD, 2010, p. 455)

O autor ainda afirma que existem duas razões principais para que a tese da corporificação extrema seja rejeitada. A primeira delas se refere aos diversos tipos de raciocínios em que algum nível de representação/computação é necessário, dado que

processos puramente sensório-motores não são suficientes para explicá-los. Como exemplos o autor cita o raciocínio causal, a manifestação de emoções e a própria teorização científica. “Se você tem o poder de representação de um humano ou um robô poderoso, pode construir modelos simplificados, mas imensamente úteis, de eventos passados e futuros, bem como de eventos que seus sentidos não permitem que você observe” (THAGARD, 2010, p. 457).

A segunda razão é devido ao fato de que mesmo nos casos centrais de ação incorporada, com abrangência do controle motor, requer algum nível de representação/computação. Sendo afirmado que mesmo em uma tarefa aparentemente simples, como estender o braço para alcançar um livro, muitas representações e computações estão envolvidas, sendo elencadas cinco evidências que sustentam essa tese: Primeira, os sinais entre cérebro, sensores e músculos são repletos de ruídos e envolvem múltiplos estágios de transformação neural; Segunda, o controle motor é dificultado pelas mudanças constantes do contexto, uma vez que além das mudanças na orientação do corpo, outras mudanças ambientais como intensidade de luz, outros objetos e outros agentes podem modificar o modo de agir; Terceira, há evidências de atrasos inevitáveis entre o planejamento da ação pelo cérebro e sua execução. Quarta, o controle motor não é um processo automático, mas requer muito aprendizado; Quinta, o controle motor não é um processo linear, mas opera pela integração de movimentos não-lineares coordenando vários músculos e articulações com diversos graus de liberdade.

A ação incorporada requer controle motor. O controle motor requer modelos mentais, tanto em um sentido direto quanto inverso, para identificar as relações dinâmicas entre as informações sensoriais e a atividade muscular. Os modelos mentais são representacionais e computacionais. Consequentemente, a ação incorporada requer representação e computação, de modo que não pode fornecer uma alternativa para a visão representacional / computacional da mente. (THAGARD, 2010, p. 457)

Dessa forma, podemos notar como a noção de Representação adotada por Thagard, enquanto um padrão de ativação neuronal, a qual estabelece uma correlação causal com os objetos representados, pode ser útil mesmo considerando alguns posicionamentos recentes da ciência cognitiva contemporânea. Além disso, devemos salientar que, apesar da noção de Representação tradicional da ciência cognitiva, a qual privilegia a representação simbólica, manter-se como alvo das críticas da abordagem da cognição incorporada. O mesmo torna-se bem mais difícil em relação a noção de Representação adotado por Thagard, uma vez que a representação simbólica é empregada de modo secundário, sendo privilegiada a representação

do padrão de disparo neuronal subjacente. Esta concepção de Representação é importante ao debate proposto, pois, uma vez que as técnicas de varredura cerebral são aplicadas com o objetivo de mapear este tipo de padrão, os padrões registrados são armazenados em formato de dados e poderão posteriormente ser utilizados para realizar as análises de *Big Data*.

No Capítulo seguinte, abordarei mais especificamente algumas teorias que tratam sobre os fenômenos da Consciência e da Intencionalidade. Além de buscar demonstrar o debate construído por alguns autores influentes, tanto na Filosofia da mente, quanto na Ciência cognitiva. As abordagens de destaque que compõem o capítulo 2 são: a *Abordagem dos sistemas intencionais* de Dennett, o *Naturalismo biológico* de Searle, a abordagem da *Consciência afetiva* de Thagard e Aubie e a *Teoria neural baseada em marcadores semânticos* de Thagard, Schröder e Stewart.

CAPÍTULO 2

**Os fenômenos Consciência e Intencionalidade:
Conceituação, limites investigativos, detalhes
estruturais e modelagem computacional**

2.1 O conceito de Consciência: Consciência psicológica e consciência fenomênica

No capítulo anterior, vimos que são muitas as abordagens que debatem sobre a natureza dos estados mentais e os meios pelos quais, uma verdadeira ciência da mente deve seguir a fim de desvelar a natureza fundamental do mental. Contudo, visando o alcance dos objetivos da presente pesquisa, ao longo desta seção, será destacado a concepção de *consciência psicológica* e *consciência fenomênica*¹⁶. A razão para este destaque é a distinção elaborada e endossada por alguns autores (CHALMERS, 1996; SEARLE, 1997; NAGEL, 1974) sobre estes dois aspectos, afirmando que uma compreensão científica da Consciência enfrenta níveis de dificuldade distintos ao serem analisados cada um destes aspectos.

O conceito de Consciência é um termo polissêmico, o qual, segundo Chalmers (1996), assume frequentemente os seguintes significados: *Vigília, introspecção, descrição de estados mentais, autoconsciência, atenção, controle voluntário e conhecimento*. Afirmando que estes significados figuram em meio a Ciência Cognitiva destacando-se seus aspectos psicológicos, ou funcionais. Entretanto, é salientado que “Embora esses conceitos tenham um núcleo psicológico, muitos, ou todos eles, estão associados a estados fenomênicos. [...] É claro que existe uma propriedade fenomenal e uma propriedade psicológica na vizinhança de cada um desses conceitos”. (CHALMERS, 1996, p. 27). A crítica a estas acepções do termo adotadas pela Ciência Cognitiva, a qual é majoritariamente funcionalista, é que o *problema difícil da consciência*¹⁷, ou seja, seu aspecto fenomênico é deixado completamente de lado.

Dito isso, iniciaremos retomando uma das teses centrais do funcionalismo, a fim de compreender como a Ciência Cognitiva aborda tais questões. A tese em questão refere-se a possibilidade de replicação de estados mentais em uma ampla gama de sistemas cognitivos, fundamentada pelo argumento da múltipla realizabilidade. O argumento afirma que “um mesmo papel funcional que caracteriza um determinado estado mental pode se instanciar em criaturas com sistemas nervosos completamente diferentes, e nesse caso diremos que eles estão no *mesmo* estado mental” (TEIXEIRA, 1998, p.49). Assim, concebendo a Consciência enquanto um estado mental, em conjunto ao argumento da múltipla realizabilidade,

¹⁶ “Em geral, uma característica fenomenal da mente é caracterizada por como é para um sujeito ter essa característica, enquanto uma característica psicológica é caracterizada por um papel associado na causação e / ou explicação do comportamento”. (CHALMERS, 1996, p. 12)

¹⁷ Segundo Chalmers, o problema difícil da consciência refere-se à capacidade dos organismos de serem possuidores de uma experiência consciente, ou seja, explicar por que todo estímulo percebido está sempre associado a uma experiência consciente. “A parte mais difícil do problema mente-corpo é a pergunta: como um sistema físico pode dar origem à experiência consciente?” (CHALMERS, 1996, p. 25)

poderíamos concluir que ao compreendermos todos os aspectos funcionais da consciência poderíamos, então, replicá-la independentemente do substrato físico? Dado que o argumento “diz que podemos ignorar o *hardware* ao caracterizar o mental, uma vez que o mesmo *software* pode ser executado em vários tipos diferentes de *hardware*: sendo crucial o desempenho funcional do *software*” (THAGARD, 1986, p.302).

Segundo Thagard (1986), mesmo assumindo uma concepção funcional do sistema, esta é uma suposição “completamente ingênua”. A objeção a tese, da independência entre papel funcional (*software*) e substrato físico (*hardware*), é elaborada considerando a computação em paralelo. “Tentarei mostrar que mesmo de uma perspectiva computacional, o *hardware* é importante” (THAGARD, 1986, p.302).

Segundo Thagard (1986), o primeiro ponto a ser considerado é a velocidade do processo computacional. O autor afirma que a computação em paralelo é evidentemente mais rápida do que a computação serial, apresentando melhor desempenho na recuperação de informação em uma rede semântica. Entretanto, a teoria da computação afirma que qualquer máquina paralela pode ser simulada em uma máquina de Turing universal equivalente. Assim, conceitualmente não há nada que o paralelismo poderia acrescentar. Por outro lado, “a metafísica não é uma empresa puramente conceitual. Do ponto de vista matemático, uma máquina de Turing é um modelo perfeitamente adequado para a operação de um computador, mas é um modelo seriamente defeituoso para entender a inteligência” (THAGARD, 1986, p.305).

Desse modo, se compararmos inteligência e processos computacionais em termos de velocidade da operação, consideramos que o tempo adequado da resposta é definido pelas restrições do ambiente a fim de garantir a “sobrevivência” do sistema.

Se é um sistema natural, é limitado pela seleção darwiniana; e se é um sistema artificial, é limitado pelas necessidades de seus projetistas. De qualquer maneira, um sistema excessivamente lento não sobreviverá por muito tempo. A inteligência, então, deve ser vista como relativa ao ambiente em que o comportamento deve ocorrer. (THAGARD, 1986, p.305)

Assim, o argumento funcionalista da múltipla realizabilidade ao estabelecer uma relação de independência entre *software* e *hardware*, negligência o aspecto prático da produção tecnológica, onde muitas vezes, *hardware* e *software* se misturam no aprimoramento do modelo. Outro ponto relevante à análise do argumento da múltipla realizabilidade tem

relação com sistemas compostos por substratos muito distintos, como é o caso de sistemas cognitivos naturais de composição orgânica e sistemas artificiais inorgânicos.

Segundo Thagard (2002), mesmo ao considerarmos que células orgânicas possam realizar computações em seu sentido mais estrito, a influência de sua composição orgânica, regida por processos elétricos e bioquímicos, permite níveis muito superiores de computações quando comparadas a sistemas cognitivos artificiais, os quais operam basicamente mediante processos elétricos.

A negligência da neuroquímica nas ciências cognitivas seria apropriada se as propriedades computacionais do cérebro relevantes para explicar o funcionamento mental fossem de fato elétricas em vez de químicas. Mas há evidências consideráveis de que a complexidade química realmente importa para a computação do cérebro. (THAGARD, 2002, p. 430)

No entanto, comparações entre a computação em paralelo e o funcionamento cerebral são comuns em meio a Ciência Cognitiva, ao serem evidenciados, principalmente após a retomada do paradigma conexionista, o caráter paralelo da comunicação nas redes de neurônios. Fato que levou alguns autores (JOHNSON-LAIRD, 1983; BURKS, 1984) a proporem que a aplicação de um modelo computacional da consciência poderia melhorar o desempenho de redes neurais artificiais. “Embora essa visão tenha muita plausibilidade inicial, argumentarei contra ela, propondo que a consciência possa não ter função computacional essencial ou, na melhor das hipóteses, uma função muito diferente do controle executivo” (THAGARD, 1986, p.311).

Thagard admite ser exagerado atribuir um papel central no controle dos processos cognitivos aos estados consciente, uma vez que não podemos nem, ao menos, decidir conscientemente sobre nossos próprios pensamentos. Nem mesmo controlarmos conscientemente processos fisiológicos de nível inferior, como por exemplo, os batimentos cardíacos. Apesar de podermos, de algum modo, lançar mão de métodos conscientes para afetar esses níveis inferiores indiretamente. Desse modo, para o autor a consciência poderia ser apenas um efeito colateral relativo ao tipo de computação ou substrato utilizado. Entretanto, o autor afirma que assumir uma *natureza epifenomenal da consciência*¹⁸ possa ser uma atitude precipitada, uma vez considerado sua relevância no planejamento de nossas

¹⁸ “Epifenomenalismo é a visão de que os eventos mentais são causados por eventos físicos no cérebro, mas não têm efeitos sobre quaisquer eventos físicos. [...] Na visão epifenomenalista, os eventos mentais não desempenham nenhum papel causal neste processo. [...] A discussão moderna do epifenomenalismo, no entanto, remonta a um contexto do século 19, no qual uma visão dualista dos eventos mentais foi assumida como correta. [...] Em contraste, muitas discussões contemporâneas trabalham dentro de uma suposição de fundo da preferibilidade do monismo materialista” (ROBINSON, 2019, p. 1).

ações, ou, então, seu caráter pedagógico, permitindo-nos comunicar aos outros como duplicar nossa performance. Além de apresentar um papel importante para o estabelecimento de relações sociais, possibilitando a adoção de uma perspectiva empática.

Como conheço meus próprios sentimentos, posso apreciar como os outros parecem estar se sentindo, dando-me muito mais compreensão do comportamento deles. Estudando sistemas naturais e projetando sistemas artificiais, podemos obter uma compreensão muito melhor do papel da consciência na inteligência. Assim, embora tenhamos motivos para duvidar que a consciência seja o mecanismo de controle central da mente, a conclusão de que a consciência é epifenomenal é prematura. (THAGARD, 1986, p.314)

Outro autor influente com muitas considerações sobre a Consciência é John R. Searle, o qual denomina sua abordagem de naturalismo biológico. A adoção do naturalismo biológico e, por conseguinte, suas considerações sobre a Consciência aproximam a teoria de Searle da distinção, apresentada anteriormente por Thagard sobre a composição de sistemas cognitivos naturais e artificiais. No entanto, a sua defesa sobre o aspecto subjetivo e qualitativo da consciência é o ponto de divergência entre a abordagem de Thagard, mas, de outro modo, aproxima-se da crítica desenvolvida por Chalmers. Verificaremos então algumas similaridades e distanciamentos entre as abordagens apresentadas até aqui.

Searle admite que por vezes há muita confusão no emprego do conceito de consciência, sendo confundida com noções de “ciência”, “autoconsciência” e “cognição”. No entanto, o autor admite que se nos afastarmos de definições analíticas sobre o termo e assumir uma definição de *senso comum*, podemos apreender uma definição razoável do termo:

“Consciência” se refere àqueles estados de sensibilidade e ciência que começam normalmente quando acordamos de um sono sem sonho e continuamos até que durmamos novamente, caímos em coma, morremos ou ficamos “inconscientes”. [...] A consciência, assim conceituada, é como um interruptor. [...] A consciência, assim definida, é um fenômeno interno qualitativo de primeira pessoa. Os seres humanos e animais superiores são evidentemente conscientes, mas não sabemos até onde a escala filogenética da consciência se estende. (SEARLE, 1998, p.33)

Segundo a definição de Searle, podemos destacar duas características centrais sobre o que ele entende por consciência. A primeira delas admite uma compreensão mais simples e plausível do termo, neste caso, o termo consciência seria análogo à estados de “vigília”, podendo, inclusive variar em níveis. A segunda caracterização refere-se a consciência, enquanto “um fenômeno interno qualitativo de primeira pessoa” e é exatamente sobre este aspecto que são direcionadas as críticas e objeções.

Searle afirma que a principal fonte de críticas a consciência é devido ao “antiquado pressuposto dualista/materialista de que o caráter “mental” da consciência torna impossível, para ela, ser uma propriedade ‘física’” (SEARLE, 1997, p.135). A fim de evitar as críticas ao aspecto “mental” da consciência, o naturalismo biológico se posiciona afirmando que a abordagem leva em consideração dois aspectos fundamentais e bem estabelecidos da visão de mundo da Ciência contemporânea, sendo elas “a teoria atômica da matéria e a teoria evolutiva da biologia” (SEARLE, 1997, p.127). Desse modo, ao levar em conta estes dois aspectos, o naturalismo biológico afirma que processos de nível inferior ocorridos no cérebro estabelecem relações causais com estados mentais conscientes. Entretanto, muita(s) críticas se direcionam a essa noção de “relação causal”. As críticas são elaboradas da seguinte maneira: Dado que processos cerebrais são fenômenos objetivos e publicamente observáveis, como estes poderiam estabelecer relação causal com fenômenos qualitativos particulares?

A contra argumentação de Searle afirma que estas objeções se fundamentam em uma concepção incorreta de causalidade, pressupondo que “todas as relações causais devem ocorrer entre eventos distintos ordenados sequencialmente no tempo” (SEARLE, 1998, p. 35). Todavia, ao afirmar que processos cerebrais de nível inferior causam estados conscientes, os estados consciente citados, não são compreendidos como entidades distintas do cérebro, mas, apenas, enquanto propriedades do funcionamento cerebral. O mesmo tipo de distinção ocorre em relação ao aspecto qualitativo dos estados conscientes.

Eu particularmente estou hesitante de usar o termo *qualia* e seu singular, *quale*, porque ambos dão impressão de que existem dois fenômenos distintos: consciência e *qualia*. Mas é evidente que todos os fenômenos conscientes são experiências subjetivas qualitativas e portanto são *qualia*. Não há dois tipos diferentes de fenômenos. Há apenas a consciência, que é uma série de estados qualitativos. (SEARLE, 1998, p. 36)

É precisamente sobre este ponto que notamos a primeira distinção entre a abordagem da consciência de teóricos como Chalmers (1996); Searle (1997) e Nagel (1974), os quais sustentam a crítica de que a Ciência Cognitiva incorpora em suas teorias apenas aspectos funcionais da experiência consciente, e a abordagem de Thagard (1986; 1996; 2002; 2008; 2011). Segundo a concepção de Thagard e Aubie (2008) a experiência consciente apresenta uma íntima relação com a emoção e a avaliação cognitiva, sendo este o resultado da

interações de diversas áreas cerebrais, coordenadas pela *memória de trabalho*¹⁹. Tal concepção “esboça um modelo de como as emoções surgem de uma combinação de representação neural, percepção somática, avaliação cognitiva e memória de trabalho”. (THAGARD, AUBIE, 2008, p. 811)

No entanto, Searle se posiciona contra a ideia de que o cérebro é um computador digital, sendo este um pressuposto compartilhado por muitos cientistas cognitivos simbólicos. Segundo essa concepção, a mente consciente seria análoga a um programa de computador. “Tomada dessa forma, a mente é para o cérebro o mesmo que um *software* é para o *hardware*” (SEARLE, 1998, p.36). Searle (1998) denomina essa concepção da mente de “Inteligência Artificial Forte (IA Forte)”, com o intuito de diferenciar de uma concepção mais moderada, que compreende os computadores enquanto ferramentas epistêmicas úteis para realizar simulações mentais, denominada de “Inteligência Artificial Fraca (IA Fraca).

No entanto, apesar do posicionamento contrário à constituição de uma IA Forte pelo naturalismo biológico, seu principal argumento não se constitui na defesa da irreducibilidade da consciência, nem tão pouco, que apenas a matéria orgânica seja capaz de constituir as condições necessárias à emergência da consciência. Segundo Searle (1997) até aquele momento, não haveria nenhum obstáculo teórico que impediria o desenvolvimento de consciência em sistemas constituídos de elementos diferentes, o verdadeiro impecilho seria, na realidade, o pouco conhecimento sobre a neurofisiologia da consciência para, então, permitir sua replicação artificial.

Entretanto, uma coisa que sabemos antes mesmo de começarmos a investigação é que *qualquer sistema capaz de causar consciência tem que ser capaz de reproduzir as capacidades causais do cérebro*. Se, por exemplo, isso for conseguido com circuitos integrados de silício em lugar de neurônios, tem que ser porque a química dos circuitos integrados de silício é capaz de reproduzir as capacidades causais específicas dos neurônios para causar consciência. (SEARLE, 1997, p.137)

As considerações de Searle são relevantes a presente pesquisa em dois aspectos: Primeiro, poderíamos supor que, com o desenvolvimento tecnológico e aplicação de ferramentas de *Big Data*, o conhecimento da estrutura neurofisiológica da consciência poderia, então, ser melhor compreendido? Possibilitando evidenciar, de maneira satisfatória, as relações causais entre processos cerebrais e estados conscientes. E supondo, então, uma

¹⁹ Apesar desta breve citação, uma explicação mais elaborada do termo é apresentada na próxima seção, nas páginas 51 e 52, desta dissertação.

resposta positiva a hipótese anterior. Seguindo este cenário, poderíamos, então, dizer que este seria o primeiro passo para o estabelecimento de um programa de pesquisa no estabelecimento de uma IA Forte?

Contudo, apesar de ser relevante à presente pesquisa, esta divisão sobre a concepção de uma IA no sentido Forte e Fraco, a crítica sobre a impossibilidade dos processos cerebrais ocorrerem por meio de computações, ou, que estes processos apresentem características computacionais muito distintas da noção teórica de computação desenvolvida por Turing, aparenta ser pouco plausível. Segundo Thagard (2002), a capacidade de neurônios e redes neuronais comunicarem-se por meio de computações é algo já bem estabelecido na Ciência Cognitiva, uma vez que se comportam recebendo, integrando e transmitindo sinais elétricos para excitar ou inibir outros neurônios da rede. Atualmente, há evidências de que até mesmo células não-neurônais apresentam capacidades computacionais. Thagard ainda responde ao cetismo sobre tais capacidades da seguinte maneira:

Descrever proteínas e células executando computações é uma ampliação metafórica que viola a noção matematicamente precisa de computação desenvolvida por Turing e outros? De forma alguma, pois há vários resultados matemáticos e experimentais recentes que mostram que o processamento molecular é computacional no sentido mais rigoroso. (THAGARD, 2002, p. 433)

Outro autor que também endossa a capacidade computacional do cérebro é Daniel Dennett. Ou de outro modo, defende a analogia cérebro-computador ao explicitar a capacidade de dispositivos digitais de simularem, com amplo grau de eficiência, dispositivos analógicos.

O computador é, no fundo, um dispositivo digital, mas um dispositivo digital pode simular um dispositivo “analógico” em qualquer grau de continuidade que desejamos; e, em um nível mais alto de agregação, no computador, poderíamos encontrar elementos analógicos que fossem mapeáveis em partes não-digitais do cérebro. (DENNETT, 2006, p. 169)

Em parte, é esta concepção, de contornos funcionalista, que fundamenta o cetismo de Dennett sobre o mistério em torno do aspecto fenomenal da consciência. “A consciência humana é apenas o último mistério sobrevivente. Um mistério é um fenômeno que as pessoas não sabem como pensar sobre – ainda” (DENNETT, 1991, p.21). Partindo disso, Dennett busca apresentar as bases de sua investigação sobre a consciência, de uma perspectiva exclusivamente materialista.

Atualmente, falamos sobre nossas decisões conscientes e hábitos inconscientes, sobre as experiências conscientes que adoramos – mas não estamos certos de

qualquer forma, sobre o que queremos dizer, quando dizemos essas coisas. Embora ainda existam pensadores que pensam que a consciência é algo genuíno e precioso (como amor, ouro), algo que é "óbvio" e muito, muito especial, cresce a suspeita de que isso é uma ilusão. (DENNETT, 1991, p.23)

Dennett (1991), pressupõe que a razão da recusa por parte de muitos pensadores sobre a possibilidade da consciência assumir uma papel ficcional, é devido a posição que ocupa na atribuição de responsabilidade moral, na garantia da livre-escolha, na distinção entre objetos animados e inanimados, na apreciação de sentimentos, sonhos e até mesmo do sofrimento, bem como por ser a fonte de ações intencionais genuínas. Considerando que este posicionamento receoso é uma barreira a ser transposta para esclarecer a compreensão da consciência. “Quando compreendermos a consciência – quando não houver mais mistérios – a consciência será diferente, mas ainda será bela e terá mais espaço do que nunca para admiração.” (DENNETT, 1991, p.25).

Sob este contexto, é afirmado que apesar da íntima relação entre cérebro e mente, ao observarmos os processos cerebrais não estamos aptos à captar os conteúdos mentais. Certamente, padrões neuronais podem ser associados com alguns estados, mas pequenas mudanças no conteúdo do pensamento não apresentam mudanças significativas nos padrões cerebrais. “O problema com eventos cerebrais, [...] é que não importa o quão próximo eles sejam “correspondentes” aos eventos em nosso fluxo de consciência, eles têm uma desvantagem aparentemente fatal: Não há ninguém lá assistindo” (DENNETT, 1991, p.29). Desse modo, a consciência é o que fundamenta a concepção de *self*, o Eu experienciador, a perspectiva de primeira pessoa, sendo, exatamente, a sua distinção com o cérebro que provoca, por um lado a defesa do aspecto fenomenal e anômalo da consciência, e de outro a suspeita de seu caráter ilusório ou superestimado. “Existem problemas notáveis com esta ideia, mas isso não impede que pareça certa: a menos que exista uma mente consciente por trás da ação, não há nenhum agente *real* no comando” (DENNETT, 1991, p.32). Desse modo, o autor afirma que um de seus principais objetivo é “explicar a consciência sem nunca ceder ao canto da sirene do dualismo” (DENNETT, 1991, p.33).

Há suspeita a espreita de que a característica mais atraente da mente é a promessa de ser tão misteriosa que mantém a ciência à distância para sempre. [...] Essa postura fundamentalmente anti-científica do dualismo é, em minha opinião, a sua característica mais desqualificante, e é a razão pela qual neste livro adoto a regra aparentemente dogmática de que o dualismo deve ser evitado *a todo custo*. Não que eu ache que posso provar que o dualismo, em todas as suas formas, é falso ou incoerente, mas que, dado o modo como o dualismo se afunda em mistério, aceitar o dualismo seria o mesmo que desistir. (DENNETT, 1991, p.37)

A conclusão de Dennett é que a consciência não passa de “operações de uma “máquina virtual”, uma espécie de programa de computador evoluído (e em evolução) que molda as atividades do cérebro” (DENNETT, 1991, p.431). Está máquina opera gerando *Rascunhos Múltiplos*²⁰ compondo os processos de fixação de conteúdo, seu funcionamento modular semi-independente, permite controlar a ampla economia cerebral, os *qualia* são concebidos, enquanto disposições do cérebro e a noção de primeira pessoa, ou *significador geral*, “acaba sendo uma abstração valiosa, no entanto, mais a ficção de um teórico do que um observador ou controlador” (DENNETT, 1991, p. 431).

De tal forma, apesar da relevância das críticas argumentativas, muito bem elaboradas por filósofos que defendem uma abordagem sobre o aspecto fenomênico da consciência. Uma abordagem menos ambiciosa e mais branda, sobre o aspecto psicológico e funcional da consciência, possibilita explicitar de modo mais objetivo os mecanismos neurais que *causam* a consciência, como a concepção de *consciência afetiva*²¹ de Thagard e Aubie (2008). Além de ser compatível com um “senso interno de bondade ou maldade moral da própria conduta real ou imaginária” (THAGARD; FINN, 2011, p. 150), por meio do qual combina as avaliações cognitivas com a percepção de estados corporais.

Os autores ainda afirmam que seu modelo da consciência afetiva permite responder a três importantes questões filosóficas. A primeira delas é a objeção de Chalmers (1996), a qual afirma que a Consciência não está aberta a uma explicação puramente neurológica. Essa objeção é conhecida com o argumento do zumbi, ou seja, sendo possível concebermos seres com a mesma estrutura fisiológica, mas completamente destuídos de uma vida mental consciente, poderíamos então afirmar que a estrutura neurofisiológica não é capaz, por si só, de explicar a emergência da consciência. A resposta de Thagard e Aubie (2008) é que neste argumento é suposto, erroneamente, uma distinção entre os sentimentos emocionais e as causas neuronais, sendo interpretada como “uma sugestão de que os sentimentos emocionais resultam da atividade neural, mas não os causam. Mas sustentamos que a consciência

²⁰ O modelo em questão é a maneira com que Dennett busca substituir a concepção da consciência, enquanto um Teatro Cartesiano, ou seja, o ponto de vista de um observador subjetivo centralizado em alguma área do cérebro. O modelo de Rascunhos Múltiplos é concebido como processos em paralelo ocorridos no cérebro, compostos de múltiplas faixas interpretativas e reelaboração contínua de informação captada pelos órgãos do sentido. (DENNETT, 1991)

²¹ A consciência afetiva é uma proposta de análise do termo consciência como um importante aspecto das emoções, a qual seria o resultado emergente da interação de várias áreas cerebrais coordenadas pela memória de trabalho. (THAGARD; AUBIE, 2008)

emocional é *apenas* a atividade neural distribuída por várias áreas do cérebro, não o resultado epifenomenal da atividade neural” (THAGARD; AUBIE, 2008, p. 829)

A segunda objeção faz alusão a possibilidade de existirem seres com um neurofisiologia distinta, como robôs ou extraterrestres, mas com os mesmos estados mentais. “Nossa resposta é que a teoria da consciência emocional proposta aqui não deve se aplicar a todos os seres pensantes possíveis, mas apenas aos humanos e outros animais terrestres semelhantes” (THAGARD; AUBIE, 2008, p. 829).

A terceira objeção apresentada se fundamenta no celebre argumento “*o que é ser como*” proposto por Nagel (1974). Segundo o argumento, a explicação da Consciência em termos de atividade neuronal mecanista é incompleta, pois não diz nada sobre *o que ser como alguém* que experiência um sentimento, como por exemplo, felicidade. “Sob esse ponto de vista, nossa própria experiência pessoal nos diz muito sobre a consciência emocional que nenhuma teoria científica poderia abordar” (THAGARD; AUBIE, 2008, p. 830). Salientando ainda, que a noção de *consciência afetiva* ainda permite sugerir vários aspectos do que seria estar feliz ou triste, como por exemplo, estar feliz é o mesmo que estar em posse de uma emoção com valência positiva, de intensidade, início e cessação variados e que é discriminável de outras emoções, como a tristeza.

Desse modo, na seção seguinte aprofundarei em alguns pontos apenas introduzidos nesta seção, como a importância da composição bioquímica do cérebro na emergência da Consciência, além de apresentar com maiores detalhes estruturais a conceituação de consciência afetiva de Thagard e Aubie.

2.2 Consciência afetiva: Composição, estrutura e processos

Como já apresentado na seção anterior, abordamos previamente de que modo o papel funcional de alguns estados mentais (*software*) estão intimamente relacionado com o substrato físico (*hardware*) que os compões. De maior prioridade para a presente seção é a compreensão de *como as moléculas importam para a computação mental*²² do nosso cérebro.

Nos “modelos computacionais de redes neurais o processamento neuronal normalmente é tratado como um fenômeno elétrico, no qual o disparo de um neurônio afeta o disparo de todos os neurônios conectados a ele por ligações excitatórias e inibitórias” (THAGARD, 2002, p. 429). No entanto, além de disparos elétricos, o funcionamento dos cérebros biológicos contam com dezenas de mensageiros químicos como os neurotransmissores, hormônios e outras moléculas. Estes mensageiros químicos são capazes de tipos de computações qualitativamente superiores quando comparados com os modelos artificiais, sejam eles simbólicos ou connexionistas.

O tipo de computação realizada por células orgânicas, conta com as moléculas de proteínas como protagonistas neste processo. “Podemos pensar em células individuais, como computadores que possuem *inputs* na forma de moléculas que se ligam a proteínas receptoras, *outputs* na forma de moléculas emitidas pelas células e processos internos realizados por reações químicas envolvendo proteínas” (THAGARD, 2002, p. 433). As proteínas funcionam em meio ao processo como interruptores de *liga/desliga*, este tipo de computação molecular é massivamente paralela, uma vez que, dentro das células, diversos receptores podem iniciar simultaneamente várias reações químicas.

São três os tipos de sinalização que uma célula, em geral, pode realizar. Uma célula pode sinalizar a si mesma por um processo conhecido como *sinalização autócrina*, ou sinalizar uma célula adjacente pelo processo de *sinalização parácrina*, assim como, sinalizar células-alvo muito distantes pela produção de hormônios, um processo conhecido como *sinalização endócrina*, além das sinalizações específicas do cérebro conhecidas como *sinalização sináptica* e *neuroendócrina*.

Os neurônios são tipos de células ainda mais desenvolvidas e é este tipo de interação que mantém, uns com os outros, que possibilita a ampla gama de processos cognitivos existentes. O tipo específico de interação dos neurônios é conhecido como sinapse, sendo por

²²

Título do artigo *How Molecules Matter To Mental Computation* de Paul Thagard (2002)

meio deste processo que é possível que um neurônio mantenha comunicação com milhares de outros neurônios. As sinapses podem ocorrer mediante impulsos elétricos entre os neurônios, ou por interação química por meio de neurotransmissores, hormônios e outras moléculas que modulam os efeitos dos neurotransmissores. Contudo, qual é a vantagem adicional que a interação química neuronal traz à computação mental? “Proponho, entretanto, que o conjunto de neurotransmissores faz diferenças qualitativas e quantitativas no processamento mental, afetando tanto seu estilo quanto sua velocidade” (THAGARD, 2002, p. 434).

Na avaliação do desempenho de redes neurais, sejam elas artificiais ou não, são três os critérios a serem observados para determinar-se as operações computacionais: O processamento interno dos neurônios; a topografia da rede; e o efeito temporal.

O processamento interno é avaliado segundo a capacidade dos neurônios de processar muitos *inputs* e a complexidade dos *outputs* possíveis. Sob este critério, modelos conexionistas mais sofisticados apresentam uma boa avaliação do seu processamento interno, respondendo de maneiras diferentes aos vários padrões que chegam a eles e produzindo diferentes padrões de resposta. No entanto, o poder computacional de células neuronais é ainda maior, uma vez considerado os três tipos de sinalizações que as células geralmente são capazes de realizar.

A topografia da rede refere-se ao padrão de conectividade que permite um neurônio afetar o disparo de outro. Em redes neurais artificiais, a topografia da rede é comumente verificada pelo padrão de conectividade estabelecido pela excitação ou inibição dos neurônios. Já as células neuronais podem estabelecer inúmeras vias de conectividade, assim, vias diferentes apresentam funções distintas no sistema. Por exemplo, a integração do movimento pela dopamina e regulação da emoção pela serotonina são processos que quando apresentam defasagem podem causar doença de Parkinson e depressão. Além disso, o padrão de conectividade do cérebro pode ser modificado por neuromoduladores, que são produzidos por outras células ampliando o poder computacional por meio de *ciclos de retropropagação*²³.

O efeito temporal tem relação com a evolução dinâmica da rede ao longo do tempo. As redes artificiais têm uma avaliação de desempenho bem inferior em relação às redes naturais. Nas redes artificiais é muito comum um tipo de ativação síncrona, em que todos os neurônios apresentam atualização de ativação ao mesmo tempo. Já nas redes naturais a ativação tende a ser assíncrona, “os neurônios reais são assíncronos e dependem da história

²³

Termo traduzido do inglês: *Feedback loops*

temporal na forma dos padrões de pico que são inseridos neles” (THAGARD, 2002, p. 435). Este é um dos principais fatores na garantia da dinamicidade das redes neurais do cérebro.

Desse modo, podemos notar como a composição bioquímica do cérebro é de extrema relevância na produção de experiências emocionais genuínas. Essa noção é importante para que possamos compreender melhor a concepção de consciência afetiva de Thagard e Aubie.

Para que possamos conceber a consciência afetiva como um tipo de experiência emocional consciente, devemos ter em mente o modo como os autores compreendem as emoções. “Emoções são processos neuronais que representam uma ampla gama de estados cognitivos e somáticos do organismo” (THAGARD; AUBIE, 2008, p. 811). De tal forma, a experiência consciente emerge em um dado organismo, quando a representação neuronal atinge altas taxas de ativação como parte da memória de trabalho. Assim, as emoções seriam resultantes da combinação dos seguintes componentes: Representações neuronais, percepção somática, avaliação cognitiva e memória de trabalho.

Contudo, mesmo considerando os componentes envolvidos na emergência de estados emocionais, é de nosso conhecimento que as emoções podem assumir as mais diversas formas e graus de expressão, dotadas de experiências qualitativas distintas. Dessa forma, é necessário avaliar a intensidade de sua incidência, seu aspecto positivo ou negativo e o início e cessação da experiência emocional, permitindo-nos explicar a distinção entre os diversos estados emocionais conscientes, sua integração com outros processos mentais, a intensidade do grau de excitação e a alternância entre estados emocionais. Além de explicitar os mecanismos que possibilitariam a emergência de características conscientes.

Um mecanismo é uma estrutura que desempenha uma função em virtude das operações, interações e organização de suas partes componentes. Os candidatos para explicar fenômenos emocionais incluem: mecanismos neurais em que as partes são neurônios e as operações são excitação elétrica e inibição; mecanismos bioquímicos nos quais as partes são moléculas e as operações são reações químicas organizadas em vias funcionais; e mecanismos sociais em que as partes são pessoas e as operações são interações sociais. (THAGARD; AUBIE, 2008, p.812)

Expostos os mecanismos, voltaremos nossa atenção na compreensão dos fenômenos *diferenciação, integração, intensidade, valência e alternância*. A *diferenciação* é o fenômeno, no qual os organismos são capazes de diferenciar um estado emocional dentre tantos outros. Alguns estados emocionais são mais básicos e considerados universais, sendo compartilhados até mesmo com animais não-humanos como, por exemplo, medo, raiva e felicidade. Assim,

“uma teoria da consciência afetiva deve ser capaz de explicar como cada uma dessas experiências é gerada por operações neuronais” (THAGARD; AUBIE, 2008, p. 812).

As experiências emocionais conscientes não ocorrem de modo isolado, mas são *integradas* a outros processos mentais como, por exemplo, percepção, memória, julgamento e inferência. Assim, uma teoria da consciência afetiva deve explicar como a combinação de processos mentais podem ocasionar em respostas emocionais, além de explicitar como interpretações distintas de um mesmo evento podem levar a diferentes reações emocionais.

É igualmente importante, que um teoria da consciência afetiva possa explicar as diferentes *intensidades* de experiências emocionais, assim como sua *valência*. É comum ao nosso conhecimento que alguns estados emocionais apresentam valências positivas como, por exemplo, felicidade, alegria e satisfação, ou negativas como tristeza, raiva e melancolia. Além da valência, podemos notar que os exemplos apresentam graus distintos de intensidade. Afinal a valência positiva de um estado de satisfação é bem menor em relação a intensidade de estados de felicidade ou alegria. Assim, é necessário que sejam identificadas as bases neuronais que resultam nas diferentes valências.

Outro ponto importante é que as experiências emocionais não são constantes, elas tendem a *alternarem-se* devido a mudança no foco da atenção de um objeto ou situação para outro. Além disso, a alternância de uma experiência emocional para outra pode ser ocasionada pela reinterpretação de um mesmo objeto ou situação, podendo ocorrer ainda em longos intervalos de tempo como, por exemplo, a modificação de uma visão negativa para positiva devido a aderência a um processo terapêutico.

Antes de apreciarmos os componentes responsáveis pela emergência das emoções, devemos elucidar a íntima relação entre estados emocionais e experiências conscientes. Segundo Thargard e Aubie (2008) a literatura sobre emoções e Consciência são repletas de aspectos controversos sobre ambos assuntos. Especificamente sobre as emoções, a disputa é posta da seguinte forma: Seriam as emoções avaliações cognitivas sobre as dificuldades e facilidades no alcance dos objetivos, ou então, seriam percepções somáticas de estados corporais? Sendo apresentada a seguinte resposta a questão: “Nossa opinião é que as teorias somáticas e cognitivas da emoção são de fato compatíveis, sendo cada uma parte da geração de emoções e, portanto, da consciência emocional” (THAGARD; AUBIE, 2008, p. 813). Além desse posicionamento sobre as emoções, os autores ainda endossam, em sua teoria da consciência afetiva, o seguinte posicionamento sobre a consciência:

De acordo com Damasio (1999), a consciência é um "sentido interno" que está envolvido com vigília, atenção e emoção. Ele distingue entre consciência central e consciência ampliada, que envolve um senso de identidade. A consciência central requer apenas uma imagem, que é um padrão mental em qualquer uma das modalidades sensoriais, e um objeto como uma pessoa ou outra entidade. A consciência ampliada requer memória que possibilite um self autobiográfico. Sentir uma emoção "consiste em ter imagens mentais decorrentes dos padrões neurais que representam as mudanças no corpo e no cérebro que constituem a emoção" (DAMÁSIO, 1999, apud THAGARD; AUBIE, 2008, p. 813)

Tendo compreendido os fenômenos envolvidos na experiência emocional consciente, bem como o modo como os autores compreendem a relação entre estados emocionais e experiência consciente. Passaremos agora a apreciar os componentes envolvidos na emergência das emoções.

O primeiro componente descrito, presente no processo da experiência consciente emocional, é a *Representação neuronal*²⁴. Este talvez seja um dos mais importantes termos para compreendermos a teoria da consciência afetiva de Thagard e Aubie (2008), dado que é por este componente que o organismo estabelece sua comunicação com o mundo. Segundo os autores, “precisamos de uma teoria da representação neuronal que seja suficiente para explicar como o cérebro pode representar o mundo, os estados corporais e suas próprias representações” (THAGARD; AUBIE, 2008, p. 814). Para entendermos como neurônios podem representar características do mundo, podemos pensar em uma população de neurônios, ou um grupo de neurônios interconectados, os quais representam tais características codificando-as em padrões de atividade neuronal com base nos objetos representados. Dizemos então, que a representação neuronal do mundo é eficiente, quando podemos estabelecer dependências estatísticas causais entre os disparos neuronais e a resposta dos sentidos (*sensores externos*) aos objetos.

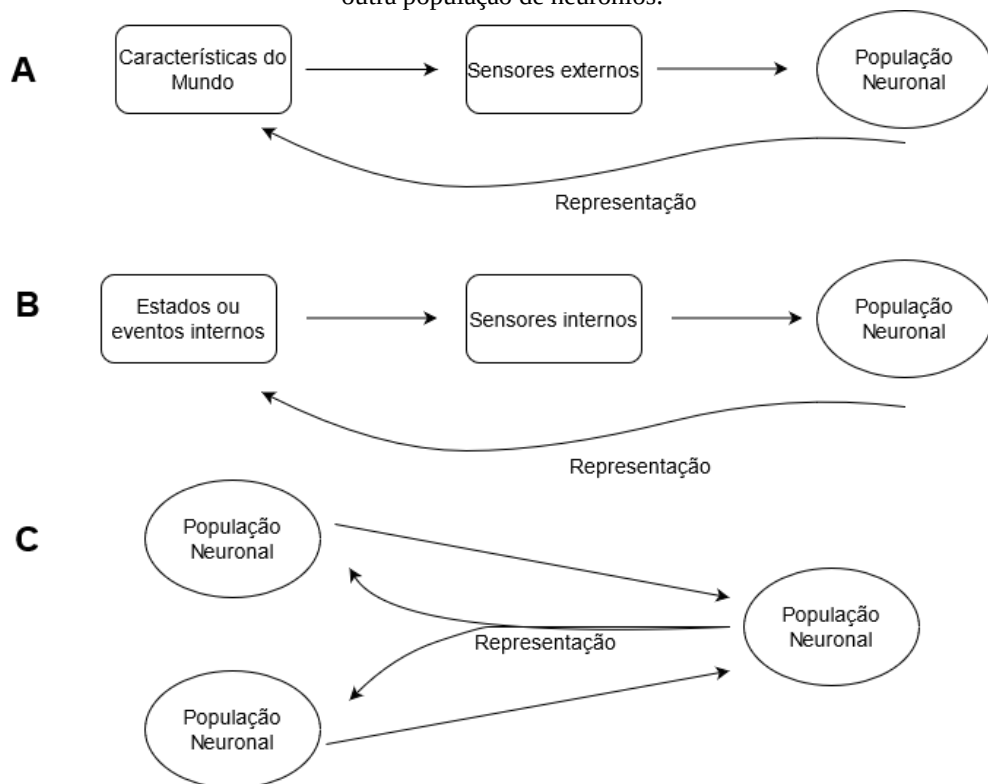
Além da representação do mundo, os neurônios precisam ser capazes de representar os próprios estados e eventos internos do organismo. Assim, como a representação do mundo necessita de sensores externos (órgãos do sentido) para captar suas características, na representação dos estados internos são os sensores internos como as conexões neurais existente em alguns órgão vitais como, por exemplo, coração e pulmão, bem como a concentração de hormônios presente na corrente sanguínea, que sinalizam a população de

²⁴ Cabe salientar que o termo representação é um termo que trás consigo uma carga de grandes controversas em Filosofia da mente. No entanto, no capítulo seguinte separarei uma seção exclusiva para tratamos sobre este termo, que pode ser tão controverso quanto o termo Consciência.

neurônios sobre as características que serão representadas. Desse modo, semelhante a representação do mundo, um grupo de neurônios pode representar os estados e eventos internos por meio de dependência estatística causal entre o padrão de disparo neuronal e a ocorrência de eventos e estados internos.

Além dos dois tipos de representações descritas, uma rede neuronal deve ser capaz de produzir representações de representações, ou seja, ser capaz de representar padrões de disparos de outra população ou grupo de neurônios. Sendo este o processo que torna possível a produção de pensamentos de ordem superior, considerados pelos filósofos como um aspecto importante da Consciência. Assim, quando “uma população neural representa outras, [...] não é apenas porque o disparo de neurônios nas populações de entrada causa disparos na população de saída. O cérebro está cheio de conexões de *feedback* pelas quais grupos neurais enviam sinais de volta aos grupos que os enviaram” (THAGARD; AUBIE, 2008, p. 815). Logo, a correlação estabelecida entre uma população de neurônios que representa outra população é o resultado da ocorrência de influência causal em ambas as direções.

Figura 1 – Representação pela população de neurônios de (A) aspectos do mundo, (B) aspectos do corpo e (C) outra população de neurônios.



Fonte: Adaptado de Thagard e Aubie (2008, p. 814).

O segundo componente descrito, refere-se as áreas cerebrais ativas em meio aos processos de tomada de decisão e percepção somática. Segundo Thagard e Aubie (2008), são sete as principais áreas envolvidas nestes processos, sendo elas: a amígdala, córtex orbitofrontal, córtex cingulado anterior, córtex pré-frontal dorsolateral, estriado ventral, neurônios dopaminérgicos do mesencéfalo e neurônios serotonérgicos centralizados no núcleo dorsal da rafe do tronco cerebral. Cada uma destas áreas tem seu papel reconhecido em meio aos processos de tomada de decisão e percepção somática, através de uma ampla gama de estudos experimentais com humanos e outros animais.

A amígdala recebe informações de sensores externos e internos e modula a intensidade das emoções positivas e negativas, especialmente o medo. O córtex orbitofrontal desempenha um papel central na avaliação da valência positiva e negativa dos estímulos e coopera com o sistema dopaminérgico do mesencéfalo e o sistema serotonérgico para calcular os ganhos e perdas potenciais de ações potenciais. O córtex cingulado anterior está envolvido na detecção de conflitos entre o comportamento atual e os resultados desejados. O córtex pré-frontal dorsolateral contribui para a representação, planejamento e seleção de comportamentos relacionados ao objetivo. (THAGARD; AUBIE, 2008, p. 815)

O terceiro componente descrito é processo pelo qual o cérebro pode valer-se das informações sobre a avaliação de seu estado atual geral, seus elementos perceptuais, sejam eles externos ou somáticos, e de todo o conhecimento acumulado. Segundo os autores, este é o processo pelo qual o cérebro realiza avaliações cognitivas. Contudo, este tipo de processo requer inferências ainda mais complexas do que a representação por *retroalimentação*²⁵ de entradas sensoriais. Os autores afirmam que um bom meio para compreender este processo é compreendê-la como um tipo de *satisfação paralela da restrição*²⁶, sendo esta uma técnica comumente utilizada na modelagem computacional de redes neurais artificiais. Entretanto, é salientado que “para maior plausibilidade biológica, seria melhor se a avaliação cognitiva fosse realizada por representações distribuídas nas quais elementos únicos são representados pela atividade em muitos neurônios e nas quais neurônios individuais participam de muitas representações” (THAGARD; AUBIE, 2008, p. 815).

Por fim, o último componente descrito, que possibilita a emergência da consciência afetiva, é a memória de trabalho. A relevância em considerar a memória de trabalho como o componente responsável pela coordenação da experiência consciente emocional esta ligada ao

²⁵ Traduzido do inglês *feedforward*. (THAGARD; AUBIE, 2008)

²⁶ A satisfação paralela da repressão é um processo em que muitas restrições são consideradas simultaneamente, gerando a melhor maneira de resolver o problema segundo os objetivos do sistema. Elaborar um calendário de disciplinas, assim como a captação de imagem por cada um de nossos olhos, são exemplos de satisfação paralela da repressão. (ANDRADE, 2021, no prelo).

seu caráter serial, muito distinto das atividades cerebrais, as quais ocorrem de modo simultâneo, paralelo e assíncrono. A memória de trabalho é responsável pelo “armazenamento de curto prazo de diferentes tipos de informação em diferentes áreas do cérebro e processos executivos de atenção seletiva e gerenciamento de tarefas que envolvem o córtex cíngulo anterior e o córtex pré-frontal dorsolateral” (THAGARD; AUBIE, 2008, p. 815). Desse modo, a coordenação realizada pela memória de trabalho dos outros componentes é o que concede a nossa experiência consciente seu caráter serial, o qual é muito distinto das atividades cerebrais.

Uma vez tendo sido apresentados os componentes necessários a emergência da consciência afetiva, podemos então verificar como este mecanismos interagem na causação dos fenômenos descritos anteriormente.

Em relação a *valência*, estudos apontam que a valência positiva dos estados emocionais está associada ao aumento de ativação do sistema dopaminérgico e do córtex pré-frontal esquerdo, enquanto as valências negativas associam-se a ativação de neurônios serotoninérgicos da rafe dorsal e do córtex pré-frontal direito. Com adição da emoção de repulsa que se correlaciona com a atividade da ínsula, a qual possui grupos de neurônios que representam estados viscerais. Os céticos quanto a relação causal entre os mecanismos e a emergência destes estados emocionais poderiam argumentar que este tipo de correlação não garante, por si só, relações causais. No entanto, a verificação da relação causal pode ser facilmente verificada na manipulação de drogas psicotrópicas que alteram a bioquímica do cérebro, resultando na modificação da valência emocional, como no caso da ingestão de antidepressivos.

Semelhante ao tipo de correlação causal entre valência e áreas cerebrais, a *intensidade* dos estados emocionais pode ser verificada pelo aumento da taxa de atividade nas populações neuronais das áreas indicadas a uma valência específica. “Por exemplo, as anfetaminas aumentam o disparo neural nos circuitos de dopamina e, portanto, aumentam a intensidade de algumas experiências emocionais positivas” (THAGARD; AUBIE, 2008, p. 819). Entretanto, a testagem do grau de previsibilidade da relação, aumento da taxa de disparos *versus* aumento da intensidade, é difícil devido a questões éticas na pesquisa em humanos usando gravações de células únicas e a limitações nas técnicas de varredura do cérebro.

A eclosão de experiência emocionais tem como característica o curto período de tempo de atividade, diferente do humor que pode durar horas ou dias. Assim, para explicar a

alternância entre experiências emocionais, podemos recorrer a modificação do estímulo externo, mas os estímulos externos não produzem novas emoções por si mesmo, então o que mais seria necessário? Para responder a esta questão, recorreremos ao componente memória de trabalho. É chamada de memória de trabalho a parte mais ativa da memória de longo prazo, a qual também pode ser gerada indiretamente, pela ativação de conteúdos utilizando-se de processos cognitivos como associação e inferência. Assim, podemos explicar como as emoções deixam e retomam os conteúdos da memória de trabalho por processos como estimulação, recorrência, decadência e inibição.

A memória de trabalho tem dois aspectos cruciais: ativação recorrente e decadência. Conexões recorrentes (também chamadas de reentrada) são aquelas que permitem que as populações neurais se estimulem, de modo que o conteúdo da memória de trabalho tende a permanecer ativo. No entanto, a memória de trabalho também está sujeita à decair, de modo que, se não houver estimulação contínua de seus conteúdos ativos pela percepção e pela memória, essas representações tenderão a desaparecer da memória de trabalho. Outro mecanismo na memória de trabalho é a inibição, em que a ativação de alguns elementos tende a suprimir a ativação de outros, de modo que os estímulos competem pela representação consciente. (THAGARD; AUBIE, 2008, p. 819)

Para que possamos explicar como são possíveis os fenômenos da *diferenciação* e *integração* na experiência emocional, além de todos os componentes descritos, é necessário que o organismo seja capaz de realizar avaliações cognitivas por meio do processo de satisfação paralela da repressão. Assim, a avaliação cognitiva que permite discriminar precisamente uma emoção dentre outras é “um modelo de satisfação de restrição que cobre mais emoções, que não são simplesmente percepções de estados corporais, mas requerem inferências sobre como a situação geral de uma pessoa está relacionada a estímulos externos e estados internos” (THAGARD; AUBIE, 2008, p. 820). Este mesmo processo também ajuda a compreender o surgimento de emoções sociais mais complexas, como vergonha, culpa e orgulho. “Assim, o processo de avaliação cognitiva pode ocorrer em paralelo com o processo de representação dos estados corporais internos, produzindo a diferenciação de uma ampla gama de experiências emocionais” (THAGARD; AUBIE, 2008, p. 821).

O que pode ser concluído então é que a Consciência está longe de ser uma faculdade mental especial, mas pode ser muito melhor compreendida como um processo neuronal que combina representação, avaliação cognitiva, percepção somática e memória de trabalho. Na seção seguinte apresentarei o problema em torno do conceito de Intencionalidade.

2.3 O conceito de Intencionalidade: Intencionalidade intrínseca e derivada

Nesta seção, abordaremos o problema da Intencionalidade, a qual, por vezes, é desenvolvido como uma ramificação do problema da Consciência, uma vez que ao se ter consciência – ao menos em um entendimento do termo – é necessário que se tenha consciência sobre algo. Logo, o conteúdo de estados conscientes são comumente direcionados a objetos ou estados de coisa do mundo, assim, dizemos que estes conteúdos apresentam Intencionalidade. A grosso modo, podemos então dizer que a Intencionalidade é a capacidade das mentes ou estados mentais de referirem-se a algo no mundo além de si mesmos.

Segundo a definição de Intencionalidade da *Enciclopédia de Filosofia de Stanford*:

Intencionalidade é o poder das mentes e dos estados mentais de tratar, representar ou direcionar à coisas, propriedades e estados de coisas. Dizer que os estados mentais de um indivíduo têm intencionalidade é dizer que são representações mentais ou que têm conteúdos. Além disso, na medida em que um falante profere palavras de alguma linguagem natural ou desenha imagens ou símbolos de uma linguagem formal com o propósito de transmitir a outros o conteúdo de seus estados mentais, esses artefatos usados por um falante também têm conteúdo ou intencionalidade. 'Intencionalidade' é uma palavra de filósofos: desde que foi introduzida na filosofia por Franz Brentano no último quarto do século XIX, tendo sido usada para se referir aos quebra-cabeças da representação, todos os quais estão na interface entre a filosofia de mente e a filosofia da linguagem. (JACOB, 2019, p. 1)

Desse modo, Intencionalidade, grafada com 'c', além de ser uma característica difusa de vários estados mentais como, por exemplo, crenças, expectativas, julgamentos e intenções²⁷, é uma característica das mentes que pode ser de ordem superior, quando o conteúdo descrito se refere a estados psicológicos de outras pessoas.

A mesma, não pode ser confundida com Intensionalidade, grafada com 's', a qual é um termo da lógica que refere a capacidade da linguagem de fazer discriminações, selecionando de modo particular um elemento específico. Por exemplo, caso eu queira me referir a um triângulo, no qual todas as suas arestas e ângulos sejam iguais, posso me referir ao mesmo objeto de dois modos distintos. Posso denominar este triângulo de *triângulo equilátero* ou de *triângulo equiangular*. Desse modo, em ambos os casos eu seleciono a mesma *extensão*²⁸, mas com duas *intensões*²⁹ diferentes.

²⁷ Neste ponto, o termo *intenções* é descrito em seu sentido ordinário: o que se pretende fazer; o qual deve ser distinguido do sentido de Intencionalidade.

²⁸ “É a coisa ou conjunto de coisas a que se refere o termo”. (DENNETT, 1997, p.41)

²⁹ “É a maneira particular pela qual esta coisa ou conjunto de coisas é selecionada ou determinada”. (DENNETT, 1997, p. 42)

Tendo posto alguns pormenores sobre alguns falsos cognatos do termo Intencionalidade, assim como uma breve definição do termo, retomaremos a íntima relação estabelecida entre Consciência e Intencionalidade. Segundo Siewert (2017, p. 2) “consciência e a intencionalidade podem parecer permear grande parte, ou toda a vida mental - talvez elas de alguma forma expliquem o que é ter uma mente; de qualquer forma, esses parecem ser aspectos importantes e amplos dela”. Esta relação estreita entre Consciência e Intencionalidade, assim como sua aparente presença pervasiva em meio aos processos mentais é tema de reflexão por filósofos e psicólogos há séculos, mas que não pode ser confundida com o tipo de representação linguística que utilizamos na fala e na escrita. Fato que pode ser verificado pelas reflexões de um influente autor da filosofia e da psicologia, William James:

E nunca se perguntou o leitor a si mesmo que espécie de fato mental é a sua INTENÇÃO DE DIZER ALGUMA COISA antes que a tenha dito? É uma intenção inteiramente definida, distinta de todas as outras intenções, portanto, um estado de consciência absolutamente distinto; e, contudo, quanto dele consiste de imagens sensoriais definidas ou de palavras ou de coisas? Alguma coisa apenas! Esperemos, e as palavras e as coisas vêm à mente; lá não está mais a intenção antecipatória, a adivinhação. Mas à medida que vão chegando as palavras que a substituem, ela as vai acolhendo sucessivamente e chamando-as exatas, se concordam com ela, e rejeitando-as e chamando-as erradas se não concordam. A intenção de DIZER ISTO E AQUILO é o único nome que ela pode receber. Pode-se admitir que uma boa terça parte de nossa vida psíquica consiste nessas visões rápida e premonitória perspectiva dos esquemas de pensamento ainda não articulados. (JAMES, 1892, p. 164)

Desse modo, podemos notar que os processos mentais não se restringem a linguagem, dado que, se este fosse o caso, não haveria sentido em se percorrer por um processo tão laborioso de encontrar as palavras exatas as intenções de fala. Mas, de outro modo, o que fundamenta esse tipo de intenção pré-linguística ou não-linguística anterior ao emprego da linguagem, seria um idioma mental? Uma espécie de *mentals*³⁰ que possa traduzir imagens sensoriais, pensamentos abstratos e expressões linguísticas? Bom, estas são as principais questões que fundamentam um dos mais relevantes debates em relação a Intencionalidade, o qual faz referência as sub-categorias *Intencionalidade intrínseca* e *derivada*.

Contudo, a fim de exemplificar o debate, devemos retomar a crítica de Searle a IA Forte, uma vez que suas considerações fazem referência a *intencionalidade intrínseca*³¹ aos

³⁰ É a linguagem mental utilizada na computação das informações na mente, a qual o pensamento seria efetivado numa linguagem universal.

³¹ O termo intencionalidade intrínseca é um conceito utilizado por Searle para se referir ao tipo de intencionalidade genuína dos seres conscientes. Entretanto, cabe salientar que nem todos os estados, nem mesmo todos os estados mentais conscientes são dotados de intencionalidade intrínseca, mas se refere a uma qualidade garantida pela existência de consciência no sistema.

estados conscientes. O jogo da imitação, proposto por Allan Turing na década de 50, constitui-se de um método para avaliar se uma máquina seria capaz de pensar, por meio de um jogo de troca de mensagens entre seres humanos e máquinas. “Em essência, [...] se o comportamento de uma máquina for indistinguível daquele exibido por um ser humano, não há razão para não atribuir a essa máquina a capacidade de pensar” (TEIXEIRA, 1998, p.33). Dessa maneira, o que é criticado por Searle seria a consideração do comportamento, enquanto condição suficiente para garantir a capacidade de pensar. Para isso, Searle propõe um experimento de pensamento, que ficou conhecido posteriormente como argumento do quarto chinês.

Imagine que você execute as etapas de um programa elaborado para responder as perguntas em um idioma que você não compreende. Eu não entendo chinês, então imagino que estou trancado em um sala cheia de caixas com símbolos chineses (a base de dados), recebo uma pequena quantidade de símbolos chineses (perguntas em chinês), e, então, procuro em um manual (o programa) o que deveria fazer. Realizo algumas operações com os símbolos de acordo com as regras (i.e., eu executo as etapas do programa) e entrego uma pequena quantidade de símbolos (respostas às perguntas) aos que se encontram fora do quarto. Eu sou um computador executando um programa para responder perguntas em chinês, mas ao mesmo tempo não compreendo uma palavra de chinês. [...] Se não entendo chinês nem ao executar um programa de computador para se entender chinês, tampouco qualquer outro computador digital entende, pois nenhum computador digital tem algo que eu não tenha. (SEARLE, 1998, p. 38)

Dessa maneira, o experimento acima procura demonstrar que um programa de computador executa tarefas seguindo, exclusivamente, regras sintáticas, enquanto a mente humana opera por meio de combinações entre sintaxe e semântica atribuindo significado as proposições. “Não importa quão bem o sistema possa imitar o comportamento de um indivíduo que realmente compreende, nem quão complexas as manipulações dos símbolos são. Você não pode extrair a semântica só a partir dos processos sintáticos” (SEARLE, 1998, p.39). Está afirmação de Searle, apesar de bem desenvolvida por meio do argumento em questão, cabe ser reconsiderada à luz da era dos *Big Data*, dado que na era dos *petabytes*, há pessoas como Chris Anderson (2008) afirmando que “Nenhuma análise semântica ou causal é necessária. É por isso que o Google pode traduzir idiomas sem realmente “conhecê-los””. Contudo, cabe destacar que apesar desta pretensa afirmação de Anderson, o mesmo não apresenta os detalhes de como este processo ocorre, tão pouco, uma justificativa para a eliminação de análises semânticas ou causais.

Por hora, o que nos é relevante é compreender como Searle relaciona conteúdo semântico e intencionalidade. O que o argumento do quarto chinês mostra, é que “assim que

colocamos algo no sistema que realmente tem intencionalidade (um homem), e o programamos com o programa formal, você pode ver que o programa formal não carrega intencionalidade adicional” (SEARLE, 1980, p.11-12). Dessa maneira, se a manipulação formal de símbolos é compreendida como uma manipulação sintática, podemos dizer que a compreensão semântica e a intencionalidade são equivalentes?

Apesar da relação próxima entre semântica e intencionalidade, a resposta é não. A intencionalidade, como compreendida por Searle, é a capacidade de seres vivos de relacionarem os seus conteúdos mentais com os objetos e estados de coisas no mundo. Toda *intenção* é dirigida à alguma coisa; por exemplo, se tenho sede, este estado interno é dirigido à alguma coisa, no caso água; o mesmo acontece com os estados intencionais de crença, desejo, medo, esperança e etc. A semântica está diretamente relacionada com o *significado*, e sua emergência ocorre ao transformarmos nossos estados intencionais em código linguístico mediante a manipulação de um conjunto de signos. Desse modo, a intencionalidade seria, segundo Searle, a condição necessária para que um conjunto de representações linguísticas adquira uma dimensão semântica. No entanto, estas não seriam equivalentes.

Muitas outras espécies, além do ser humano são dotadas de percepção sensorial e ações intencionais; e diversas espécies, sobretudo os primatas, têm crenças, desejos e intenções, mas só algumas espécies, talvez apenas os humanos, são dotados de uma forma biologicamente peculiar de Intencionalidade, a qual associamos com a linguagem e significado. [...] Uma consequência natural da abordagem biológica defendida neste livro é considerar o significado, no sentido na qual falantes significam algo por sua expressão, como um desenvolvimento especial de formas mais primitivas de intencionalidade. (SEARLE, 1997, p.160)

Contudo, apesar da conceituação de intencionalidade apresentada por Searle. Outros teóricos, como Daniel Dennett, são discordantes dessa posição. Dennett elabora sua teoria seguindo, o que ele denomina de *abordagem dos sistemas intencionais*, seu objetivo é destituir o aspecto interno/subjetivo atribuído à consciência e a intencionalidade, concebendo estes conceitos, enquanto entidade teóricas atribuídas à sistemas cognitivos, com a finalidade de explicação e previsão do comportamento.

Searle elabora sua teoria da mente, partindo dos estados conscientes, enquanto um critério distintivo entre sistemas cognitivos naturais e artificiais. Dennett, por outro lado, tem como objetivo estabelecer um método instrumental que permita explicar e prever o

comportamento, denominado pelo autor de *postura intencional*³². “A postura intencional é a estratégia de interpretar o comportamento de uma entidade (pessoa, animal, artefato, qualquer coisa) tratando-a *como se* fosse um agente racional que governa suas “escolhas” de “ação” por uma “consideração” de suas “crenças” e “desejos”” (DENNETT, 1997, p.32). Desse modo, Dennett, ao contrário de Searle, busca transpor as fronteiras estabelecidas entre as diversas formas de sistemas cognitivos.

O uso impróprio da postura intencional pode confundir seriamente o pesquisador desprevidado, mas, entendida de modo apropriado, ela pode fornecer uma perspectiva sólida e frutífera em várias áreas diferentes, mostrando a unidade subjacente dos fenômenos e dirigindo nossa atenção para as experiências cruciais que precisam ser feitas. (DENNETT, 1997, p.32)

A adoção de uma postura intencional proposta por Dennett é interessante, pois nos permite retomar a questão do fundamento de nossas intenções pré-linguísticas. O autor afirma que postular uma língua do pensamento, ou seja, um *mentalês* é uma resposta inútil, não devido à impossibilidade de sua existência, mas em virtude da circularidade entre o entendimento desse idioma e seu emprego como fundamento pré-linguístico. “É inútil à pergunta que fizemos, pois ela simplesmente a adia. [...] Este ponto de vista é inútil como resposta à nossa pergunta fundamental, porém, porque depende do próprio entendimento do que se supõe que ela deva explicar, e portanto entra em um círculo vicioso” (DENNETT, 1997, p. 52-53).

Assim, como então poderíamos responder a questão? Para responder a questão, Dennett recorre a uma estratégia estranha a princípio, mas que pode ser melhor compreendida ao nos lembramos da importância dos nossos componentes biológicos.

A solução para o problema da nossa intencionalidade é direta. Nós simplesmente concordamos que os artefatos representacionais (como descrições escritas e esboços) possuem intencionalidade derivada, em virtude do papel que desempenham nas atividades de seus criadores. [...] E como seu cérebro chegou a atingir uma organização de tais estados espantosos com esses poderes espantosos? Use o mesmo método: o cérebro é um artefato e obtém sua intencionalidade, seja qual for a intencionalidade que suas partes possuam, a partir do papel que desempenha na economia em funcionamento do sistema maior de que faz parte – ou, em outras palavras, das intenções de seu criador, a Mãe Natureza (também conhecida por processo de evolução por seleção natural). (DENNETT, 1997, p. 53)

³² A postura intencional é um método de antropomorfização de entidades, atribuindo intencionalidade derivada com a finalidade de explicar e prever o comportamento. Entretanto, cabe salientar que, tendo em vista que a filosofia de Dennett considera a mente enquanto uma teoria útil na explicação e previsão do comportamento, todas as formas existentes de intencionalidade são derivadas.

Dennett ainda afirma, que o poder de convencimento do argumento do quarto chinês é devido a habilidade do autor em fazer com que “as pessoas, na realidade, não imaginem o caso nos detalhes que lhe são necessários” (DENNETT, 1991, p.436). Desse modo, Dennett personifica e solicita que imaginemos o Quarto Chinês (QC), como descrito por Searle (ou seja, um ser humano assumindo o papel de um programa de computador), participando do Teste de Turing. O diálogo entre o juiz do teste e o QC é suposto discorrer sobre uma anedota étnica, envolvendo um irlandês, um gênio e uma lâmpada mágica. O intuito deste exemplo, é demonstrar que a suposição de Searle de que o QC passaria no teste de Turing é, de fato, simplificadora demais. Assim, a inclusão de detalhes serve como fundamento ao estabelecimento desse possível diálogo, e a complexidade necessária a um programa, que pudesse responder tais questões a fim de ter sucesso ao confundir o juiz.

Ao contextualizar o argumento do quarto chinês, Dennett estabelece o cenário adequado para formular a sua crítica, de que, se o quarto chinês operasse sob o contexto real do teste, muitas outras coisas deveriam ser consideradas. Neste contexto, um programa que realmente pudesse operar obtendo sucesso real no teste, necessitaria de um alto grau de flexibilidade e sofisticação, várias camadas de programação e de conhecimento global, meta-conhecimentos sobre suas respostas e motivações, bem como das respostas e motivações de seu interlocutor. “Searle não nega que programas possam ter toda essa estrutura, é claro. Ele simplesmente nos desencoraja de realizarmos isso” (DENNETT, 1991, p.438). Assim, Dennett afirma que a plausibilidade do argumento de Searle, depende da concordância com “a premissa suprimida: Certamente *mais do mesmo*, não importa quanto mais, nunca resultaria em um entendimento genuíno” (DENNETT, 1991, p.438).

Concluindo que ao considerarmos a complexidade que as máquinas possam vir à alcançar, não haveria qualquer razão, a não ser a alegação da existência de uma alma imaterial ou propriedades de natureza física anômalas, que nos impediriam em *princípio* de supor que uma máquina possa ter um entendimento genuíno, ou até mesmo, que essa possa ser uma característica irrelevante e empiricamente vazia.

Complexidade importa. Caso contrário, haveria um argumento muito mais curto contra a IA forte: “Ei, olhe para esta calculadora de mão. Ela não entende chinês, e qualquer computador concebível é apenas uma calculadora de mão gigante, então nenhum computador pode entender chinês. .Q.E.D.”. Quando levamos em conta a complexidade, como devemos, precisamos realmente incluí-la - e não apenas fingir levar em consideração. Isso é difícil, mas até o momento, qualquer intuição que temos sobre o que “obviamente” não está presente, não são confiáveis. Como o caso de Frank Jackson de Mary, a cientista das cores, o experimento mental de Searle só

produz uma convicção forte e clara quando deixamos de seguir as instruções. Essas bombas de intuição estão com defeito; elas não melhoram, mas enganam nossa imaginação. (DENNETT, 1991, p.441)

Apesar do tópico sobre a possibilidade de construção de uma inteligência artificial dotada de singularidade não ser recorrente. Creio que com o desenvolvimento da internet, a popularização das tecnologias digitais, o avanço exponencial da capacidade de processamento de informação e o desenvolvimento do processo de *datificação* são fatores impulsores a emergência dos *Big Data*. Assim, este pode se apresentar como um contexto interessante para a retomada sobre o debate da consciência e da intencionalidade em dois aspectos.

Primeiro, ao supor a consciência segundo Searle e o maquinário tecnológico atual, enquanto ferramentas epistêmicas, podemos estar em vista de compreender de maneira detalhada a estrutura neurofisiológica dos processos conscientes, como proposto a duas décadas atrás. Dessa forma, o desenvolvimento das técnicas de processamento de quantias exorbitantes de dados, podem contribuir para um amplo desenvolvimento científico em áreas como a Neurociência e a Psicologia. Entretanto, ao concebermos a consciência e a intencionalidade, enquanto entidades teóricas a fim de prever o comportamento, como defendido por Dennett. Poderíamos conceder ainda mais confiança as ferramentas de *Big Data* na predição do comportamento humano, uma vez considerado a complexidade de sua computação? E em caso positivo, poderíamos lançar mão de um questionamento adicional: Uma vez que na era dos grandes dados o comportamento humano pode ser previsto com um grande taxa de acerto, e conteúdos informacionis podem ser direcionados a fim de controlar a ação individual e coletiva, como poderíamos garantir a autonomia dos indivíduos segundo este contexto? Essas são questões que serão desenvolvidas no próximo capítulo.

Na Seção seguinte, apresentarei uma terceira abordagem, sendo esta a proposta selecionada como a mais compatível com os objetivos desta dissertação, sobre o fenômeno da Intencionalidade, a qual o concebe enquanto um processo neuronal que integra representações de estados de coisas, ações e avaliação emocional. A abordagem em questão foi desenvolvida por Paul Thagard em conjunto com Tobias Schröder e Terrence C. Stewart, descrita como *Uma Teoria Neural Baseada em Marcadores Semânticos*³³.

³³ Traduzido do título do artigo: *Intention, Emotion, and Action: A Neural Theory Based on Semantic Pointers*. (SCHRÖDER; STEWART; THAGARD, 2014)

2.4 Intencionalidade, emoções e ação: uma teoria neural baseada em marcadores semânticos

Como visto na seção anterior, uma vez considerado o caráter pervasivo da Intencionalidade, torna-se evidente sua relevância em meio às disciplinas como filosofia, psicologia, inteligência artificial e direito. Em filosofia, o debate se configura por meio de questões como: qual a natureza dos estados intencionais? Qual sua relação com outros estados mentais? A Intencionalidade constitui relação causal com as ações ou apenas as justificam? Qual a sua relação com ações futuras e com ações em andamento? Qual a diferença entre ações intencionais e não-intencionais? Por que algumas pessoas falham em agir segundo suas intenções? E qual sua relação com a responsabilidade moral e livre-arbítrio?

Na psicologia, as questões assumem um aspecto pragmático, questionando-se sobre sua relação causal com comportamento. É também um tema muito debatido por psicólogos as chamadas “lacunas entre intenção-ação”, as quais têm relação com as falhas em agir segundo as próprias intenções. A fim de auxiliar pessoas com esse tipo de queixa, psicólogos desenvolveram técnicas terapêuticas chamadas de “intenções de implementação”, sendo estas, conjuntos de regras que ligam características percebidas com situações específicas e compromissos assumidos anteriormente, facilitando uma escolha comportamental assertiva.

A neurociência cognitiva atenta-se aos mecanismos fisiológicos que sustentam a emergência de estados intencionais. Por meio da aplicação de ferramentas de neuroimagem com primatas não-humanos, sugerindo evidências de que a área frontal e parietal do cérebro, assim como o controle sensoriomotor, são regiões importantes dos fenômenos intencionais.

Em inteligência artificial, a Intencionalidade é um aspecto importante na construção de computadores que possam ser considerados agentes inteligentes. Sendo, por vezes, identificada como uma espécie de “desejo” que mantém o agente comprometido com o plano a ser seguido.

No direito, a Intencionalidade é concebida como uma importante característica da atribuição de responsabilidade moral. É muito comum observarmos em julgamentos, que ações evidentemente intencionais são julgadas com penas mais severas do que ações não-intencionais, as quais ocorreram devido à negligência ou imprudência do agente.

Assim, notamos que este fenômeno mental tão importante requer que três aspectos principais sejam elucidados: 1) Quais os mecanismos e de que modo relacionam-se a fim de

possibilitar a emergência de intenções; 2) De que modo, intenções podem causar ações e o que ocorre quando uma intenção não é capaz de causar a ação desejada?; e 3) Qual a relação entre intenções e ações automáticas e deliberadas? Estes três questionamentos perpassam todo o debate sobre a Intencionalidade em todas as esferas disciplinares descritas.

Para tanto, a fim de apresentar uma teoria capaz de responder aos questionamentos sobre os fenômenos intencionais, apresento a teoria neural da intenção de Schröder, Stewart e Thagard (2014). Segundo os autores, intenções podem ser concebidas como um tipo de *marcador semântico*³⁴, o qual pode ser considerado como “padrões de atividade de disparo neural cuja estrutura é uma consequência de operações de compressão de informação implementadas em conexões neurais” (SCHÖRDER; STEWART; THAGARD, 2014, p. 854). Um modo mais simples de conceber a ideia de marcadores semânticos é por meio da analogia com arquivos compactados em formato JPEG ou MP3, os quais de maneira similar não mantêm em sua estrutura toda a informação disponível.

Sendo, portanto, uma espécie de representação de outras representações, os marcadores semânticos têm a capacidade de serem decompostos nas estruturas representacionais subjacentes, permitindo ao sistema controlar o fluxo de informação de diferentes modalidades. Este processo de descompressão é de suma importância para compreender como as intenções estabelecem uma relação causal com o comportamento, dado que é por este processo que representações simbólicas desencadeiam representações motoras de baixo nível, as quais são responsáveis pelo controle do comportamento. O processo ocorre por meio da conexão entre uma semântica superficial, compostas pela relação simbólica que estabelece com o mundo, e uma semântica profunda, constituída pelas relações entre informação perceptiva, motora e emocional.

Desse modo, os autores se posicionam de forma a concordar com a ideia de “que as intenções são fenômenos cognitivos de alto nível que modelam configurações de representações de nível inferior em modalidades múltiplas. Quando ligados, eles podem causar ação por meio do roteamento de informações semânticas para o sistema motor” (SCHÖRDER; STEWART; THAGARD, 2014, p. 855). Outro aspecto importante é que as ligações por marcadores semânticos são recursivas, sendo por conta deste aspecto que sua

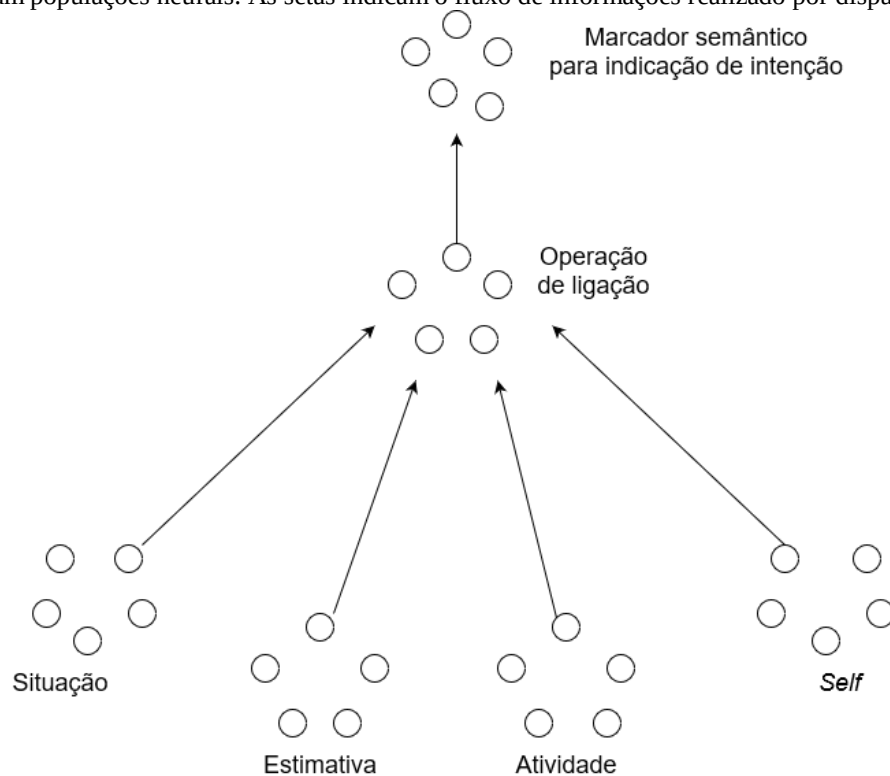
³⁴ É um termo cunhado por Chris Eliasmith, o qual por ser caracterizado funcionalmente por “representações compactadas que apontam para um conteúdo semântico mais completo e podem ser compostas em estruturas representacionais complexas para apoiar a cognição” (ELIASMITH, 2013, p. 94)

aplicação permite reconciliar representações conexionistas distribuídas com a perspectivas hierárquicas baseadas em regras.

Intenções são marcadores semânticos, ou seja, padrões de atividade de disparos que unem representações neurais de situações e sua avaliação junto com ações e às vezes o *self*. Todos esses componentes são marcadores semânticos, ou seja, padrões de atividade de disparos por conta própria. A operação de ligação depende de transições de padrão neural embutidas nos pesos de conexão entre as respectivas populações de neurônios. (SCHÖRDER; STEWART; THAGARD, 2014, p. 855)

Assim, para que possamos melhor visualizar como este processo funciona, o diagrama abaixo adaptado de Schröder, Stewart e Thagard (2014) permite-nos visualizar como as intenções formadas por representações vinculadas à situação, avaliação, ação e o *self* dinamizam o fluxo de informação, sendo então selecionada e comprimidas pelas operações de ligação, gerando marcadores semânticos. Cabe ainda salientar que, apesar de intenções serem consideradas marcadores semânticos, nem todos os marcadores semânticos são intenções.

Figura 2 - Intenções formadas por representações vinculadas à situação, avaliação, ação e *self*. Os conjuntos de círculos indicam populações neurais. As setas indicam o fluxo de informações realizado por disparo neural.



Fonte: adaptado de Schröder, Stewart e Thagard (2014, p. 855)

Tendo sido apresentado como os autores compreendem as intenções enquanto uma espécie de marcador semântico, passamos agora a buscar responder as questões elaboradas.

Visando elucidar os componentes envolvidos e sua relação na emergência de estados intencionais, descreverei como processos cerebrais são capazes de captar informações relevantes de situações pela percepção, como são realizadas as avaliações necessárias, o planejamento das ações possíveis e as normas subjetivas do *self*.

Para que um grupo de neurônios possa gerar uma representação de uma situação, algumas características do ambiente devem ser processadas pelas regiões cerebrais envolvidas na percepção, ou seja, pelas áreas cerebrais destinadas ao processamento de informação visual, olfatória, auditiva e táteis. Sendo essas representações mais básicas restrições à formação de intenções. Além das características físicas do ambiente, na avaliação de situações devem ser captadas as informações simbólicas do ambiente, seguindo o tipo de conexão descrita acima entre semântica superficial e profunda.

Além da captação das características físicas do ambiente pela percepção, assim como os aspectos simbólicos, por meio do conhecimento compartilhado culturalmente sobre identidades e instituições sociais, parte da avaliação de uma situação perpassa os estados emocionais relativos a uma situação específica. Segundo os autores:

Os humanos avaliam constantemente as situações com o sistema emocional do cérebro e acreditamos que essas avaliações sejam um importante bloco de construção das intenções. As áreas do cérebro com papel proeminente no processamento da avaliação emocional incluem (mas não estão limitadas a) amígdala, ínsula, córtex pré-frontal ventromedial e núcleo accumbens. (SCHÖRDER; STEWART; THAGARD, 2014, p. 856)

Desse modo, para a avaliação de uma situação específica utiliza-se das representações de conceitos simbólicos em avaliações afetivas que refletem o conhecimento cultural comum. Esse tipo de significado afetivo, assume um papel importante no planejamento das ações, uma vez que vale-se de pistas sútis do ambiente para alinhar ações automáticas sem percepção consciente. Contudo, é comum um tipo de ação deliberada, onde as pessoas ignoram associações emocionais automáticas quando estas entram em conflito com objetivos e valores conscientemente endossados. Sendo esta o tipo de desassociação que permitem explicar as “lacunas entre intenção-ação”.

As intenções também necessitam que as próprias ações pretendidas sejam representadas. Esse tipo de representação não é de natureza necessariamente linguísticas, antes disso, elas devem assumir a forma de padrões de ativação em meio as áreas do cérebros destinadas ao processamento de instruções motoras. Mediante isso, evidências neurocientíficas têm sustentado a hipótese de um “vocabulário de ação” no córtex pré-motor,

consistindo em representações abstratas de programas motores relacionado aos objetivos. Dessa forma, “a analogia com marcadores semânticos como modelos compactados de representações sensório-motoras mais profundas é direta e é apenas mais um degrau na hierarquia do sistema cognitivo para uma representação simbólica de ações com a linguagem” (SCHÖRDER; STEWART; THAGARD, 2014, p. 857).

Em um tipo específico de intenção, onde as pessoas planejam refletindo sobre as suas próprias ações, se faz necessária uma avaliação das representações do próprio *self*. “Auto-representações são marcadores semânticos que resultam da ligação de informações auto-relacionadas em várias modalidades, de caracterizações verbais abstratas, [...] aos significados emocionais associados a representações cinestésicas” (SCHÖRDER; STEWART; THAGARD, 2014, p. 857). Assim, as auto-representações são constituídas por auto-conceitos, memórias afetivas e experiências de agência sensoriomotoras. Os auto-conceitos são como rótulos linguísticos que se utilizam de categorias culturalmente construídas, cristalizadas pela linguagem, para dar sentido a si mesmo e as próprias experiências sociais. Por meio de auto-representação de memórias afetivas, as pessoas experienciam uma coerência contínua de seus estados afetivos. O senso de agência presente no cerne das auto-representações é resultante de “ligações intencionais” das informações motoras *aferentes*³⁵ com entrada perceptual *eferente*³⁶.

Conjecturamos que esse processo fornece a base para a representação do *self*. O resultado da vinculação *eferente* – *aferente* é um marcador semântico que pode ser armazenado na memória e posteriormente recuperado e ele próprio vinculado recursivamente em um marcador semântico de nível superior diferente. Assim, as experiências sensório-motoras anteriores de agência formam a base para a posterior inclusão do *self* em intenções complexas. (SCHÖRDER; STEWART; THAGARD, 2014, p. 858)

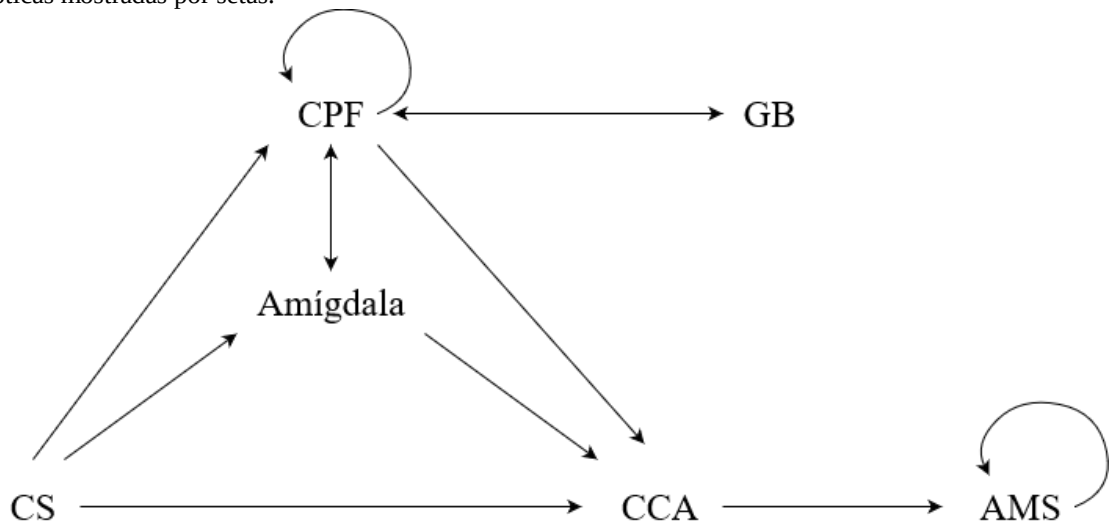
Passamos agora a analisar como este processo ocorre segundo uma perspectiva neurológica e psicologicamente plausível. Para tanto, os autores utilizam de um modelo connexionista, no qual as unidades representam populações neuronais de seis áreas cerebrais: córtex sensorial (CS), córtex pré-frontal (CPF), gânglios da base (GB), amígdala, córtex

³⁵ “Os neurônios sensitivos têm a função de levar as informações ocorridas no meio externo ao sistema nervoso central (SNC). São *aferentes*, pois captam os estímulos externos e os conduzem ao SNC. Em suas extremidades periféricas, possuem estruturas muito elaboradas denominadas de receptores. Os receptores transformam os vários tipos de estímulos (físicos ou químicos) em impulsos nervosos antes de conduzi-los ao SNC” (DAMILANO *et al*, 2005, p.9).

³⁶ O neurônio motor tem como função conduzir o estímulo nervoso ao órgão efetuator: músculo ou glândula. É *eferente*, pois conduz o impulso nervoso do SNC ou do sistema nervoso periférico (SNP) ao órgão efetuator. (DAMILANO *et al*, 2005, p.9)

cingulado anterior (CCA) e a área motora suplementar (AMS). O fluxo de informação entre as áreas correspondentes pode ser melhor visualizado na figura 3.

Figura 3 – Componentes funcionais do modelo de intenção, consistindo em seis grupos de neurônios e conexões sinápticas mostradas por setas.



Fonte: Adaptado de Schröder, Stewart e Thagard (2014, p. 858)

Analisando a figura, notamos que a entrada de dados é inteiramente realizada pelo córtex sensorial, onde os disparos de diferentes padrões de atividades resultam em padrões distintos de estimulação do modelo. A saída de dados é realizada pela área motora suplementar, na qual a presença de diferentes padrões de disparos representa ações distintas realizadas pelo modelo.

Para que o modelo execute tarefas complexas, precisamos manipular esses padrões internamente. Para fazer isso, tratamos cada padrão de disparo como um marcador semântico diferente, permitindo-nos definir cálculos para combinar e extrair informações desses padrões. Em particular, nosso modelo se baseia em transições de padrão neural: criar conexões sinápticas entre dois grupos de neurônios de modo que, se um determinado padrão fizer parte da atividade no primeiro grupo, o segundo grupo de neurônios será conduzido a algum outro padrão de atividade. (SCHÖRDER; STEWART; THAGARD, 2014, p. 859)

No CPF e na AMS, além das conexões com outras áreas há também conexões de *feedback*, as quais permitem que essas áreas mantenham o padrão de atividade produzido. Esse tipo de conexão fornece a essas áreas uma espécie de memória, permitindo que um padrão de atividade mantenha-se constante mesmo na ausência da entrada que o acionou. Além de possibilitar uma transição gradual do padrão de atividade, mudando lentamente o padrão atual até que este corresponda com o padrão desejado. Esse tipo de conexão é importante, pois permite armazenar um marcador semântico arbitrário ao longo do tempo.

A conexão entre o CCA e a AMS é mediada pelo padrão de atividade contido na amígdala. O padrão de atividade na região da amígdala atribui uma *valência* sobre a ação atual. Assim, quanto maior o valor da *valência* maior a chance da área motora suplementar armazenar o padrão de atividade presente no córtex cingulado anterior. “Essa preferência permite que o modelo execute ações rapidamente se forem consideradas muito boas, e até mesmo decida não fazer uma ação se perceber que ela teria um valor baixo” (SCHÖRDER; STEWART; THAGARD, 2014, p. 859).

Por fim, o GB possibilita a escolha de uma ação dentre a uma lista de ações possíveis. “Esse modelo segue um processo semelhante de ter regras que mapeiam um padrão em outro padrão, mas força apenas uma regra a ser ativa por vez” (SCHÖRDER; STEWART; THAGARD, 2014, p. 860). Funcionando, portanto, como uma espécie de funil na seleção da ação que deverá ser desempenhada.

Toda a descrição realizada até o presente momento, responde a questão de número 1), elaborada no início desta sessão. Além de nos colocar em uma posição favorável que permita responder as questões de número 2) e 3).

Com base no modelo apresentado, podemos notar que a associação da Intencionalidade a uma espécie de marcador semântico permite explicar como é estabelecida a relação causal com o comportamento. Uma vez que estes marcadores resultam da ligação de diferentes representações, estes possibilitam encaminhar informações para as áreas motoras do cérebro, responsáveis pela realização das ações. Desse modo, diferente das crenças ou desejos, as intenções abrangem uma espécie de comprometimento distinto, exigindo o estabelecimento de ligações com as representações de situações pelo córtex sensorial e pré-frontal e avaliações emocionais realizadas pela amígdala, as quais atribuem uma valência, determinando o tipo de ligação estabelecida com a área motora na coordenação de ações.

Com base na teoria de que todos os conceitos são baseados em significados afetivos culturalmente compartilhados, argumentamos que os efeitos de planejamento comportamental ocorre porque os conceitos de planejamento automaticamente evocam avaliações específicas nas redes afetivas do cérebro, que, por sua vez, ativam representações de ações emocionalmente congruentes. (SCHÖRDER; STEWART; THAGARD, 2014, p. 869)

Além disso, a inclusão da ligação entre o córtex pré-frontal e os gânglios da base, permite ao modelo um caminho deliberativo adicional. Assim, os marcadores semânticos da intenção operam no córtex pré-frontal, vinculando as representações subjacentes de modo que podem interromper uma tendência de ação impulsiva, substituindo uma avaliação emocional

inicial. “A competição dinâmica do controle automático e deliberativo da ação no cérebro é atualmente a explicação psicológica mais amplamente aceita para o frequente fracasso das intenções em produzir ações” (SCHÖRDER; STEWART; THAGARD, 2014, p. 870). Assim, para compreendermos melhor como intenções podem falhar na produção de ações, tomaremos como exemplo o fenômeno psicológico da procrastinação.

O fenômeno da procrastinação ocorre quando um agente tem o dever de realizar alguma tarefa pouco prazerosa e, mesmo mantendo a intenção de fazê-la, acaba realizando outras tarefas mais prazerosas com pouca ou nenhuma ligação com a tarefa inicial. Agora analisando este fenômeno em termos analíticos, o que ocorre de fato, é que o caráter aversivo da tarefa, em conjunto com uma dificuldade de substituir o afeto negativo por aspectos positivos (comumente associados à recompensa de longo prazo), resulta em adiamento da tarefa inicial (que deveria ser realizada) e busca por outra atividade que possa gerar afetos positivos em curto prazo. Para que as intenções possam ser efetivas em situações como esta “o afeto negativo inicial associado à tarefa precisa ser substituído por avaliações mais positivas das consequências de longo prazo de lidar com a tarefa, e isso requer esforço e capacidade cognitivos” (SCHÖRDER; STEWART; THAGARD, 2014, p. 870). De modo a modificar o aspecto automático de uma ação em uma ação deliberada.

Desse modo, podemos notar que a teoria proposta por Schörder, Stewart e Thagard (2014) associa a noção de Intencionalidade com marcadores semânticos, os quais possibilitam a explicação de vários fenômenos intencionais. Além disso, a noção de marcadores semânticos possibilita substituir a necessidade de uma linguagem mental, o *mentalês* descrito na seção anterior, servindo de fundamento as intenções pré-linguísticas, sem que para isso ocorra petição de princípio. Para isso, a teoria proposta baseia-se em evidências empíricas de outras disciplinas, como a neurofisiologia e a inteligência artificial, além de confirmar e adaptar teorias psicológicas e filosóficas sobre o fenômeno intencional. Por fim, a referência ao papel desempenhado pela emoção e avaliação cognitiva, permitem que os marcadores semânticos conectem a Intencionalidade com a Consciência, como proposto por filósofos e psicólogos.

No capítulo seguinte, será apresentada uma breve explicação da noção de Representação envolvida nas teorias apresentadas, bem como os riscos envolvidos na aplicação dos *Big Data* e como o debate sobre a Consciência e Intencionalidade pode auxiliá-los a refletir sobre eles. Além de discutir como a emergência do fenômeno contemporâneo dos *Big Data* poderia auxiliar no desenvolvimento da investigação destes fenômenos.

CAPÍTULO 3

***Big Data*, Consciência e Intencionalidade: os riscos envolvidos e os desenvolvimentos alcançados**

3.1 A promessa de revolução científica dos *Big Data*

O termo *Big Data* refere-se a uma tecnologia contemporânea no campo da ciência de dados. Sua emergência teve como principais fatores às novas formas de produzir, armazenar e analisar dados associada ao desenvolvimento de técnicas computacionais, algorítmicas, estatísticas e matemáticas para extrair o máximo conhecimento destes. Outro fator importante foi o movimento *Open Data*³⁷, o qual encorajou o compartilhamento e interligações heterogêneas de grandes estruturas digitais. Suas primeiras aparições ocorreram ainda ao final da década de 90, mas o debate sobre o tema ganhou maior repercussão após a primeira década do terceiro milênio.

Apesar da dificuldade em atribuir uma definição precisa do termo, autores como Schönberger e Cukier (2014) afirmam que “os *big data* referem-se a coisas que podem ser feitas em grande escala que não poderiam ser feitas em uma escala menor, para extrair novas ideias ou criar novas formas de valor”. Desse modo, “a mudança de escala levou a uma mudança de estado. A mudança quantitativa levou a uma mudança qualitativa” (SCHÖNBERGER; CUKIER, 2014, p.6.).

Os autores enfatizam que, assim como a construção de instrumentos ópticos permitiram a humanidade explorar o universo para além do sistema solar, bem como descobrir a existência de um micro-universo habitado por microorganismos, as novas técnicas de coleta e análise de dados permitirão uma revolução científica que estamos, ainda, apenas começando a apreciar. A maior modificação que a emergência dos *Big Data* implicaria nas práticas científicas, segundo os autores, seria a ênfase em correlações ao invés de causalidade. Munidos de uma quantidade descomunal de dados digitais, os algoritmos podem correlacionar esses dados possibilitando a emergência de informações relevantes. “Os dados podem revelar segredos para aqueles com humildade, vontade e as ferramentas para ouvi-los” (SCHÖNBERGER; CUKIER, 2014, p.5)

Desse modo, é destacado que “a verdadeira revolução não está nas máquinas que calculam dados, mas nos próprios dados e em como os usamos” (SCHÖNBERGER; CUKIER, 2014, p.7). Sendo afirmado pelos autores que “a mera existência de computadores não faria com que o *big data* acontecesse” (SCHÖNBERGER; CUKIER, 2014, p.15).

³⁷

É uma tendência política emergente que impulsiona o compartilhamento de dados governamentais e científicos. (LEONELLI, 2020).

De tal modo, o que possibilita a grande revolução causada pela emergência dos *Big Data* é possibilidade de correlação entre uma grande quantidade de dados, viabilizando a emergência de padrões informacionais, indisponíveis em análises de quantidades menores. A existência deste imenso volume de dados na contemporaneidade é devido aos avanços no processo de *datificação*³⁸, o qual deve ser distinguido do processo de *digitalização*.

A diferença entre datificação e digitalização, pode ser melhor compreendida por meio do seguinte exemplo: Segundo Schönberger e Cukier (2014), suponha que nosso objetivo seja *digitalizar* todos os textos de uma biblioteca. Podemos utilizar *scanners* convertendo páginas de livros em imagens digitais, sendo estas armazenadas em um banco de dados de um computador digital. Este procedimento traz consigo algumas vantagens, como por exemplo economia de espaço físico, maior segurança e conservação dos documentos. Entretanto, o processo de digitalização não permite que realizemos análises textuais, uma vez que os documentos digitalizados são meras cópias digitais, ou seja, faltam-lhes o conteúdo indexical que permita análises mecânicas. Desse modo, “recupera-lo, [...] exigiria saber qual livro tinha as informações desejadas ou fazer muita leitura para encontrar a parte certa. Não seria possível pesquisar no texto por palavras específicas ou analisá-lo, porque o texto não havia sido *datificado*” (SCHÖNBERGER; CUKIER, 2014, p.83, grifo nosso).

Assim, para que pudéssemos criar um sistema de busca, realizar análises textuais ou mesmo descobrir quais palavras eram mais frequentes ao longo dos anos, o processo de digitalização necessitaria de um *software* para reconhecimento óptico de caracteres, reconhecendo individualmente cada letra, palavra, sentença e frase. “A *datificação* torna o texto indexável e, portanto, pesquisável. Permitindo um fluxo interminável de análise textual. [...] Transformar palavras em dados desencadeia vários usos. Os dados podem ser usados por humanos para leitura e por máquinas para análise. ” (SCHÖNBERGER; CUKIER, 2014, p.85, grifo nosso). Em síntese, o processo de digitalização permite que transformemos um conteúdo analógico em digital, ao passo que a *datificação* permite que o conteúdo digitalizado seja informativo tanto para os humanos quanto para as máquinas, permitindo análises automatizadas.

³⁸ O processo de datificação (*datification*) é concebido como o procedimento de transformar ações individuais e coletivas em dados *online* metrizados, permitindo o rastreamento e controle em tempo real, além da realização de análises preditivas. (SCHÖNBERGER; CUKIER, 2014)

Apesar do exemplo acima, na atualidade o processo de datificação não se limita a esse tipo de dado. “Por causa dos *smartphones* e da tecnologia computacional barata, nunca foi tão fácil encontrar dados sobre os atos essenciais da vida” (SCHÖNBERGER; CUKIER, 2014 p. 95). No período contemporâneo, além da datificação usual de variáveis como tempo, espaço, peso, área e volume é possível transformar em dados aspectos ainda mais pessoais, como nossas interações, experiências e humor. Um exemplo desse tipo de análise, pode ser verificada em meio aos estudos liderados pelo pesquisador Michal Kosinski³⁹, neles os pesquisadores demonstram a eficácia na predição de traços comportamentais com base na análise de *Likes* da plataforma *Facebook*, além de apresentarem como estas análises podem ser aplicadas na persuasão psicológica em massa.

Segundo Schönberger e Cukier (2014) as plataformas de redes sociais não oferecem apenas meios de nos mantermos conectados com nossos amigos e familiares, mas permitem que empresas por trás dessas plataformas transformem elementos de nossa vida cotidiana em dados, sendo posteriormente correlacionados para gerarem informações. Além do conteúdo da troca de mensagens pelos usuários das redes sociais, nestas mensagens estão contidos muitos outros dados importantes, este tipo de dado é denominado de *metadado*, ou seja, “informação sobre informação” (SCHÖNBERGER; CUKIER, 2014, p. 93). Exemplos de metadados contidos nas mensagens são os dados referentes a geo-localização do usuário, o nome do usuário, as pessoas que o seguem, a hora do envio da mensagem, o idioma utilizado, entre outros. Desse maneira, o processo de datificação “não é apenas transformar atitudes e sentimentos em uma forma analisável, mas também o comportamento humano” (SCHÖNBERGER; CUKIER, 2014, p.93).

Depois que o mundo é datificado, os usos potenciais das informações são basicamente limitados pela genialidade de alguém. [...] Hoje temos as ferramentas (estatísticas e algoritmos) e o equipamento necessário (processadores digitais e armazenamento) para executar tarefas semelhantes muito mais rapidamente, em escala e em muitos contextos diferentes. [...] A datificação representa um enriquecimento essencial na compreensão humana. Com a ajuda dos *Big Data*, não consideraremos mais o mundo como uma sequência de acontecimentos que explicamos como fenômenos naturais ou sociais, mas como um universo composto essencialmente de informações. (SCHÖNBERGER; CUKIER, 2014, p. 96)

39

Os artigos de Kosinski que tratam dos assuntos citados são *Private traits and attributes are predictable from digital records of human behavior* e *Psychological targeting as an effective approach to digital mass persuasion*.

Este entusiasmo com avanços tecnológico, nos remete a atitude semelhante de Marc Weiser que na década de 90 apontou as vantagens da *computação ubíqua*⁴⁰, a qual têm um papel importante na emergência dos *Big Data*. Segundo Araujo (2003), o termo computação ubíqua é muitas vezes concebido como um sinônimo dos termos computação pervasiva e computação móvel, entretanto, estes três termos são conceitualmente distintos. A computação ubíqua é um fenômeno emergente da convergência entre a computação pervasiva e computação móvel.

A computação móvel refere-se a capacidade de deslocamento físico dos dispositivos computacionais. Um exemplo contemporâneo é a disseminação abundante de *smartphones*, o quais permitem ao usuário acesso à *internet*, capacidade computacional e conectividade independente de sua localização. A computação pervasiva refere-se a capacidade dos computadores de ajustarem as informações do ambiente com a finalidade de aumentarem seu desempenho computacional. Como exemplo, podemos citar a capacidade atual dos *smartphones* de reconhecerem e se conectarem a redes de *wi-fi* de forma automática, bem como a capacidade de autorizarem baixar certos arquivos, apenas ao estabelecerem conexão à uma rede com velocidade compatível. Dessa forma, “a computação ubíqua beneficia-se dos avanços da computação móvel e da computação pervasiva. A computação ubíqua surge então da necessidade de se integrar mobilidade com a funcionalidade da computação pervasiva” (ARAUJO, 2003, p.50).

Uma característica importante da computação ubíqua é que, segundo Weiser (1991), “elementos especializados de *hardware* e *software*, conectados por cabos, ondas de rádio e infravermelho, serão tão ubíquos que ninguém notará sua presença”. A computação ubíqua é um fenômeno, onde a presença e conectividade dos computadores, está tão arraigada ao ambiente que os usuários passam a não se darem conta de sua presença.

Segundo Weiser (1991) para que uma tecnologia possa alcançar seu mais alto grau de utilidade e benefício à vida humana, esta deve se misturar ao seu cotidiano sem que seja notada. O autor exemplifica sua asserção, afirmando que a linguagem escrita poderia ser a primeira tecnologia da informação. Dessa maneira, capturamos cotidianamente informação de representações simbólicas, sem ao menos nos darmos conta da engenhosidade desse

⁴⁰ O termo computação ubíqua se refere a uma relação simbiótica entre sistemas informacionais e seres humano. De modo, que os dispositivos artificiais inteligentes espalhados pelo ambiente, não sejam mais percebido pelos seres humanos devido à familiaridade com a sua presença.

processo. De mesmo modo, o autor enfatiza que a computação deve atingir esse mesmo patamar, tornando-se uma tecnologia que não seja mais percebida aos usuários.

Tal desaparecimento é uma consequência fundamental não da tecnologia, mas da psicologia humana. Sempre que as pessoas aprendem algo suficientemente bem, deixam de ter ciência disso. [...] Todos dizem, em essência, que somente quando as coisas desaparecem dessa maneira somos liberados para usá-las sem pensar e, assim, focar além delas em novos objetivos. (WEISER, 1991, p.94).

O artigo de Weiser escrito no início da década de 90 é apenas um vislumbre, bem informado, sobre os futuros avanços da tecnologia. No entanto, na contemporaneidade podemos admitir que a tecnologia já atinge amplo grau de ubíquidade nas sociedades industriais, sendo este um fator importante para o advento da era dos grandes dados. “Uma característica fundamental desse novo cenário, [...] é a natureza diversa de seus agentes. Essa diversidade dá lugar a um sistema complexo invisível que envolve interações entre seres vivos e uma infinidade de dispositivos artificiais” (QUILICI-GONZALEZ *et al.*, 2014, p.166).

A popularização das Tecnologias Informacionais de Comunicação (TICs) em conjunto com a implementação de câmeras de vídeo e sensores eletrônicos, constituíram um ambiente de contínua interação humano-máquina, permitindo a produção de quantias exorbitantes de dados, os quais puderam ser correlacionados com o intuito de detecção de padrões sobre os mais diversos aspectos da vida humana. Assim, podemos notar a existência de uma forte relação entre a ubiquidade computacional e a emergência do fenômeno *Big Data*.

Até o momento as maiores aplicações das técnicas de *Big Data* são direcionadas ao comércio, negócios e, até mesmo, na política visando a criação de novas formas de valores, direcionamento da ação de consumo dos usuários e reforçamento de crenças ideológicas.

Contudo, “outros projetos buscam criar novas ferramentas e métodos para extrair conhecimento de grandes conjuntos de dados para acelerar o progresso em pesquisas científicas e de engenharia e criar uma força de trabalho de *Big Data*” (PORTMESS; TOWER, 2015, p.2, grifo nosso). Assim, dado que o fenômeno do *Big Data* é um dos mais novos entusiasmos no campo da tecnologia e da inteligência artificial, ao passo que consideramos o histórico desta área no campo da ciência cognitiva, bem como os debates sobre a mente e sua relação com o corpo, poderíamos supor que com o avanço e desenvolvimento das técnicas de *Big Data*, estes poderiam auxiliar a compreendermos melhor a consciência e a intencionalidade? E ainda, qual seriam seus papéis em meio a uma sociedade que tem as suas decisões influenciadas por tais tecnologias?

A relevância destas questões, leva em conta, o modo divergente de abordar tais fenômenos tanto pela Filosofia da mente, quanto pela Ciência Cognitiva, assim como o entusiasmo envolto na aplicação destas novas ferramentas tecnológicas, as quais já tem apresentado resultados importantes em diversas áreas do conhecimento. Além disso, considero que a analogia humano-máquina, iniciada na Filosofia da mente e consolidada pela Ciência cognitiva, mantém-se presente em meio à proposta de revolução científica pelos *Big Data*. Assim, na próxima sessão, serão destacados em maiores detalhes os riscos direcionados a liberdade de escolha na era dos grandes dados, bem como um apontamento de como estes riscos se relacionam com os fenômenos da Consciência e da Intencionalidade.

3.2 Os riscos da era dos *Big Data*: a ditadura dos dados

No primeiro capítulo, vimos como a proposta do dualismo substancial, elaborado no século XVII, suscitou a retomada do debate na Filosofia da mente no século XX, possibilitando o desenvolvimento da Ciência Cognitiva e da noção de representação. No segundo capítulo, analisamos alguns modos de compreender os fenômenos da Consciência e da Intencionalidade, bem como a apresentação de algumas formas de abordar estes fenômenos utilizando-se de ferramentas tecnológicas contemporâneas. Agora, cabe lembrar, que apesar do grande auxílio que o desenvolvimento científico e tecnológico pode oferecer, muito frequentemente, há diversos riscos inerentes a sua aplicação e desenvolvimento. Assim, buscando elaborar uma análise responsável sobre a possibilidade de implementação de técnicas de *Big Data*, apresentarei os principais riscos envolvidos na sua implementação, além de demonstrar como o estudo da Consciência e da Intencionalidade pode auxiliar na reflexão de alguns desses riscos.

Segundo Schönberger e Cukier (2014) a ubiquidade da tecnologia no período contemporâneo, bem como a ampla taxa de compartilhamento de informação digital e *datificada*, impõe sérios riscos a privacidade dos usuários, principalmente, se estes dados forem manipulados para gerar novas formas de informação útil com a aplicação de técnicas de *Big Data*. No entanto, este não é o único risco envolvendo o uso destas técnicas, mas incluem-se a controversa previsibilidade criminal, baseada em probabilidades, e uma espécie de ditadura dos dados.

Como explicaremos, essa transformação não apenas torna a proteção da privacidade muito mais difícil, mas também apresenta uma ameaça inteiramente nova: penalidades baseadas em propensões. Essa é a possibilidade de usar previsões de *Big Data* sobre as pessoas para julgá-las e puni-las antes mesmo de agirem. Fazer isso nega ideias de equidade, justiça e livre arbítrio. Além da privacidade e da propensão, existe um terceiro perigo. Corremos o risco de ser vítimas de uma ditadura de dados, pela qual fetichizamos a informação, o resultado de nossas análises, e acabamos fazendo mal uso disso. (SCHÖNBERGER; CUKIER, 2014, p. 151)

Desse modo, uma vez conscientes do contexto histórico que possibilitou a emergência de uma era dos grandes dados, bem como de alguns aspectos sobre os fenômenos mentais da Consciência e da Intencionalidade. Podemos, então, seguir avaliando os riscos envolvidos na implementação das técnicas de *Big Data*, utilizando de tais informações para guiar a nossa avaliação e conectá-la aos fenômenos descritos.

O principal risco envolvendo as técnicas de *Big Data* faz referência às dificuldades imposta à manutenção da privacidade dos usuários da rede, apresentando-se como uma grande barreira à manutenção da liberdade de escolha ou livre-arbítrio dos indivíduos. A conexão, entre este risco e os fenômenos propostos, é mais geral neste caso específico, mas cabe de igual modo uma explicação dos problemas envolvidos.

A questão da privacidade é que, na era dos grandes dados, a maior parte dos dados coletados inclui informações pessoais dos usuários. Além disso, mesmo os dados que não apresentam informações pessoais podem ser correlacionados a fim de deduzir detalhes íntimos sobre a vida pessoal dos indivíduos. “A questão principal, no entanto, não é se os *Big Data* aumentam os riscos à privacidade (pois aumentam), mas se isso modifica o caráter do risco” (SCHÖNBERGER; CUKIER, 2014, p. 153). Entretanto, cabe salientar que, na era dos *Big Data*, os dados coletados não tem seu valor atribuído aos propósitos primários da coleta, mas sim em seu uso secundário. Desse modo, o caráter do risco é distinto da mera exposição das informações pessoais do usuário, repousando na dificuldade imposta a um consentimento consciente pelo usuário, uma vez que os propósitos secundários não estariam definidos no momento da coleta.

Essa mudança enfraquece o papel central atribuído aos indivíduos nas leis de privacidade atuais. Hoje eles são informados no momento da coleta quais informações estão sendo coletadas e com que finalidade; então eles têm a oportunidade de concordar, para que a coleta possa começar. [...] Surpreendentemente, na era dos *Big Data*, a maioria dos usos secundários inovadores não foram imaginados quando os dados foram coletados pela primeira vez. Como as empresas podem fornecer notificações para um propósito que ainda não existe? Como os indivíduos podem dar consentimento informado a um desconhecido? (SCHÖNBERGER; CUKIER, 2014, p. 153)

A proposta de informar os usuários no momento da coleta sobre os possíveis usos secundários dos dados é apenas uma tentativa desesperada, mas ineficaz, de conseguir o consentimento dos usuários. Segundo Thagard e Finn (2011), um modo de conceber o fenômeno da Consciência seria compreendê-la como uma espécie de intuição moral. “Mas o que torna as intuições morais diferentes de outros tipos não morais de reações emocionais? [...] A diferença crucial deve envolver a avaliação cognitiva, com intuições morais ativando um conjunto mais restrito de objetivos do que a consciência emocional geral” (THAGARD; FINN, 2011, p. 153).

Desse modo, como é possível conceber que a avaliação cognitiva dos usuários seria genuína, se nem mesmo os objetivos secundários da coleta são claros. “No contexto de *big*

data, o conceito testado e confiável de aviso e consentimento é muitas vezes muito restritivo para descobrir o valor latente dos dados ou muito vazio para proteger a privacidade do indivíduo” (SCHÖNBERGER; CUKIER, 2014, p. 154). No melhor dos casos, o que de fato este tipo de notificação alcançaria seria um tipo de “consentimento no escuro”.

Além disso, outra questão importante é que mesmo quando os dados são apagados de um banco de dados, eles podem deixar rastros. E mesmo as reconhecidas técnicas de anonimização parece não funcionar na maioria dos casos. A anonimização se refere ao procedimento de retirar dos bancos de dados as informações pessoais dos usuários como nome, endereço, número do cartão de crédito, do seguro social entre outros. Contudo, “isso funciona em um mundo de pequenos dados. Mas, na era dos *Big Data*, o aumento da quantidade e variedade da informação, facilita a reidentificação” (SCHÖNBERGER; CUKIER, 2014, p. 154). Isso ocorre devido ao amplo grau de conectividade e ubiquidade das redes, possibilitando que um indivíduo gere e compartilhe diversos tipos de dados sobre vários aspectos de sua vida particular, os quais podem ser correlacionados a fim reidentificar o indivíduo.

Não é por acaso que, muito recentemente, a Lei Geral de Proteção de Dados (LGPD), instaurada no Brasil, dispõe em seu Art. 6º, inciso I, que as atividades de tratamento de dados pessoais deverão observar, entre outros, o princípio da finalidade. “A realização do tratamento (dos dados) para propósitos legítimos, específicos, explícitos e informados ao titular, sem possibilidade de tratamento posterior de forma incompatível com essas finalidades” (BRASIL, 2018/2019). Além de dispor em seu Art. 16º sobre a eliminação dos dados após o alcançado o término do tratamento dos dados pessoais.

Art. 16. Os dados pessoais serão eliminados após o término de seu tratamento, no âmbito e nos limites técnicos das atividades, autorizada a conservação para as seguintes finalidades: I - cumprimento de obrigação legal ou regulatória pelo controlador; II - estudo por órgão de pesquisa, garantida, sempre que possível, a anonimização dos dados pessoais; III - transferência a terceiro, desde que respeitados os requisitos de tratamento de dados dispostos nesta Lei; ou IV - uso exclusivo do controlador, vedado seu acesso por terceiro, e desde que anonimizados os dados. (BRASIL, 2018/2019)

Dessa forma, podemos notar como a era dos grandes dados impõe um risco severo a privacidade, de um modo talvez irreversível. “Na era dos *Big Data*, as três estratégias principais há muito usadas para garantir a privacidade – notificação e consentimento individual, apagar dados e anonimização dos dados – perderam muito sua eficácia” (SCHÖNBERGER; CUKIER, 2014, p. 156). Uma proposta que poderia auxiliar na

manutenção da privacidade seria a conscientização dos riscos inerentes ao uso de técnicas de *Big Data*, como faço nesta seção, mas de modo muito mais amplo e abrangente. Entretanto, essa proposta serve apenas para remediar a situação, mas é provável que uma solução para esta situação necessite da colaboração e reflexão de um grupo maior de pessoas.

Outro risco importante que os *Big Data* oferecem à sociedade contemporânea concerne a sua capacidade preditiva. A questão aqui é a seguinte: com os *Big Data*, assim como podemos lançar mão de algoritmos poderosos para correlacionar diversos dados produzidos pelo usuário e o reidentificar, é pressuposto que o mesmo processo possa ser empregado para analisar os dados e mapear padrões de comportamentos do usuário, e aplicá-los para prever o comportamento. Esse tipo de capacidade preditiva teria aplicações diversas como na área da saúde, segurança e até mesmo criminal.

Essa punição baseada em previsões parece uma melhoria em relação às práticas que já aceitamos. A prevenção de comportamentos prejudiciais à saúde, perigosos ou de risco é a base da sociedade moderna. Tornamos o fumo mais difícil para prevenir o câncer de pulmão; exigimos o uso de cinto de segurança para evitar fatalidades em acidentes de carro; não permitimos que as pessoas embarquem em aviões com armas de fogo para evitar sequestros. Essas medidas preventivas restringem nossa liberdade, mas muitos as veem como um pequeno preço a pagar em troca de evitar danos muito mais graves. (SCHÖNBERGER; CUKIER, 2014, p. 160)

No entanto, a aplicação da predição de *Big Data* no contexto penal trás a tona questões mais complexas que estas outras áreas. Para ilustrar, os autores Schönberger e Cukier (2014), utilizam de uma situação semelhante apresentada pelo filme *Minority Report*. No filme, o policial John Anderton trabalha em uma divisão especial da polícia, conhecida como *Precrime*, a qual, com base nas premonições de três videntes, indicam os suspeitos (ou culpados?) e os prendem antes que cometam o crime. Apesar de ser um filme de ficção, os autores supõe que um cenário muito semelhante poderia ser construído pela aplicação de ferramentas preditivas de *Big Data*. “Um número crescente de lugares nos Estados Unidos [...] emprega “policiamento preditivo”: usando análise de *Big Data* para selecionar quais ruas, grupos e indivíduos devem ser submetidos a um escrutínio extra” (SCHÖNBERGER; CUKIER, 2014, p. 158).

Agora suponha que um algoritmo desse tipo seja conectado a outros bancos de dados, os quais além de permitirem análises em tempo real, indicariam o indivíduo e o momento exato da ocorrência do crime. A princípio pode parecer uma alternativa atraente ao combate a situações ilícitas, afinal as atividades criminosas tenderiam a diminuir, bem como o número

de vítimas. Entretanto, culpar alguém com base em eventos futuros e análises de *Big Data* é um caminho arriscado a se seguir. “Isso parece aceitável se o objetivo for simplesmente evitar ações indesejadas. Mas torna-se muito perigoso se usarmos previsões de *Big Data* para decidir se alguém é culpado e deve ser punido por um comportamento que ainda não aconteceu” (SCHÖNBERGER; CUKIER, 2014, p. 161).

A aplicação de predição probabilística para julgar ou condenar alguém pela realização de um crime tem implicações sérias tanto do ponto de vista ético, quanto metodológico. Do ponto de vista ético “acusar uma pessoa de algum possível comportamento futuro é negar o próprio fundamento da justiça: [...] É um princípio fundamental de nossa sociedade que a responsabilidade dos indivíduos está ligada à escolha individual de ação” (SCHÖNBERGER; CUKIER, 2014, p. 161). Assim, podemos notar como uma compreensão dos fenômenos da Consciência e da Intencionalidade são importantes para dimensionar a gravidade da situação, uma vez que esses dois fenômenos são de grande relevância na explicação de como os indivíduos direcionam o curso de sua ação e, por conseguinte, a atribuição de responsabilidade pelos próprios atos. “O problema da consciência é fundamental para a filosofia da mente e, mais geralmente, para a metafísica, epistemologia e ética. [...] A ética e a estética também são afetadas pela determinação da natureza da consciência” (THAGARD, 2018, p. 383).

Desse modo, considerando as teorias de Thagard sobre Consciência e Intencionalidade, percebemos o quão complexo é a interação neuronal na constituição de um *self*, o qual desempenha um papel importante no direcionamento de ações intencionais, sendo esta uma característica importante para a atribuição de responsabilidade penal.

Consequentemente, a ciência pode continuar a desenvolver teorias neuropsicológicas de decisão e ação. Uma teoria recente modela intenções e emoções como marcadores semânticos e descreve ações como resultantes de processos neurais em várias áreas do cérebro, incluindo o tálamo, gânglios da base, amígdala, córtex pré-frontal e córtex motor (Schröder et al. 2014). No entanto, essa teoria ainda não foi intimamente integrada com a teoria do marcador semântico da consciência usando modelos computacionais unificados. Uma vez que boas teorias de ação, decisão, escolha e consciência estejam estabelecidas, será possível abordar a questão filosófica fundamental sobre se as pessoas poderiam ter agido de outra forma se assim tivessem escolhido. (THAGARD, 2018, p. 388)

Consequentemente, mesmo considerando um algoritmo equipado com uma espécie de modelo integrado que pudesse avaliar os possíveis cursos de ação. O que parece improvável por si só, dado a complexidade dos eventos internos do indivíduo. Contudo, mesmo nesse

caso, seria igualmente antiético punir alguém por uma ação não cometida, pois a possibilidade *agir de outra forma* seria diretamente violada.

De modo semelhante, os problemas metodológicos envolvidos iriam desde a complexidade exigida no desenvolvimento de um algoritmo desse porte, ao que poderia vir a ser, o primeiro passo para uma era da ditadura dos dados. Segundo Schönberger e Cukier (2014) “os *Big Data* exprimem um problema muito antigo: confiar nos números quando eles são muito mais falíveis do que pensamos” (SCHÖNBERGER; CUKIER, 2014, p. 163). A ideia geral em uma ditadura dos dados seria a confiança indiscriminada nas predições e análises de *Big Data*, tomando o mundo como uma massa de informação desorganizada, a qual, caso venha a ser delineada, denotada, demarcada e quantificada, seria passível de ser domadas por mãos humanas e obedientes a nossa vontade.

Entretanto, apesar dos dados serem úteis na compreensão de fenômenos e no planejamento de soluções para diversos problemas, devemos ter em mente, que eles apresentam graus distintos de qualidade, são passíveis de serem enviesados, podem ser mal analisados e mal utilizados, além de ser possível que falhem em capturar o que fora pretendido quantificar.

Somos mais suscetíveis do que pensamos à "ditadura dos dados" - ou seja, a permitir que os dados nos governem de maneiras que podem fazer tanto mal quanto bem. A ameaça é que vamos nos deixar ser estupidamente limitados pelo resultado de nossas análises, mesmo quando temos motivos razoáveis para suspeitar que algo está errado. Ou que ficaremos obcecados em coletar fatos e números por causa dos dados. Ou que atribuiremos aos dados um grau de verdade que eles não merecem. (SCHÖNBERGER; CUKIER, 2014, p. 166)

Após averiguarmos alguns riscos inerentes às aplicações das ferramentas de *Big Data*. Na seção seguinte, apresentarei algumas dificuldades técnicas e culturais envolvidas na aplicação neurobiológica dos *Big Data*, bem como algumas propostas contemporâneas, como a *governança responsável dos dados*, a qual seria aplicada visando diminuir os riscos envolvidos, além de garantir a privacidade e segurança no uso de dados tão sensíveis, como são os dados registrados no campo da Neurociência.

3.3 *Big ‘Brain’ Data*: A aplicação dos *Big Data* na investigação da Consciência e Intencionalidade

Apesar dos riscos envolvidos na implementação das técnicas de *Big Data*, é verdadeiramente tentador pressupor as vantagens de sua aplicação “no mundo da neurobiologia, onde os *Big Data* são, realmente, epicamente grandes” (LANDHUIS, 2017, p. 559). Para que possamos ter uma ideia, segundo Landhuis (2017), mesmo no mapeamento de células nervosas de uma criatura simples como uma mosca, foi necessário uma década inteira para mapear os 60.000 neurônios e cada uma de suas sinapses, gerando uma taxa estimada de 1 *gigabyte* de informação por célula. Seguindo esta estimativa, é esperado que um mapeamento completo dos 86 bilhões de neurônios do cérebro humano levaria cerca de 17 milhões de anos. “Apesar dos avanços na infraestrutura de computação e transmissão de dados, os neurocientistas continuam a lutar com sua versão da revolução do “*big data*” que varreu o campo da genômica décadas atrás” (LANDHUIS, 2017, p. 559). Além da questão da quantidade de informação no mapeamento do cérebro humano, há também problemas relativos à variedade nas escalas de medição.

Os cientistas podem mapear o cérebro em vários níveis. O HCP (*Projeto Conectoma Humano*) busca mapear a conectividade do cérebro em escala macroscópica, usando imagens de ressonância magnética (MRI). Alguns laboratórios estão mapeando trilhas neurais em um nível microscópico, enquanto outros, como o de Chiang, rastreiam cada sinapse e ramo neural com precisão em nanoescala. Outros ainda estão trabalhando para sobrepor padrões de expressão gênica, medições eletrofisiológicas ou outros dados funcionais nesses mapas. As abordagens usam métodos diferentes - mas todas criam *Big Data* (concebendo os ‘*Big Data*’ em números’). (LANDHUIS, 2017, p. 560).

Além da quantidade e da variedade, outra questão que tem preocupado os neurocientistas são questões relativas à privacidade, como já salientado antes. “Além disso, os dados da neurociência, além de sua absoluta complexidade e multimodalidade, são inerentemente sensíveis e pessoais, pois possuem relevância filosófica no que diz respeito à natureza da humanidade e são concomitantemente de natureza pessoal” (FOTHERGILL *et al*, 2019, p. 2). Dessa forma, os autores propõem que a produção e manipulação dessa espécie de dados deve ser guiada por uma *governança responsável dos dados*⁴¹.

⁴¹ “Consideramos a “governança responsável dos dados” seja inclusiva abarcando preocupações relacionadas às questões éticas, legais e sociais que podem surgir em relação a qualquer uso dos dados em questão” (FOTHERGILL *et al.*, 2019, p. 2).

Como visto anteriormente, os principais problemas envolvidos na aplicação dos *Big Data* são relativos às dificuldades de se conseguir consentimento informado sobre o uso dos dados, a dificuldade de manter a anonimização e a possibilidade de reidentificação dos usuários. Contudo, Fothergill *et al.* (2019, p. 3) propõe que “A governança de *Big Data* pode, portanto, apresentar uma oportunidade de abordar de forma prática questões éticas e tentar mitigar alguns riscos, evitando a perda de benefícios sociais potenciais gerados pelo trabalho em questão”. Desse modo, os autores afirmam que um modo de oferecer um caminho a seguir em direção a uma governança responsável dos dados, seria aplicar ao uso da tecnologia de *Big Data* os princípios da PIR (*Pesquisa e Inovação Responsável*)⁴². Entretanto, a massiva aplicação da tecnologia de *Big Data* em contextos mercantil e monetário dificulta a implementação de tais princípios, uma vez que “esses exemplos também exigem implementação de cima para baixo e pressupõem percepções éticas e posições morais homogêneas nas organizações” (FOTHERGILL, 2019, p. 3).

A ética dos dados abertos ou ciência aberta vs. privacidade e proteção de dados e problemas relacionados à coleta de dados, compartilhamento de dados, anonimização e consentimento informado na pesquisa em saúde humana são locais significativos e óbvios para o debate na pesquisa de *Big Data*. Essas discussões geralmente apresentam perspectivas conflitantes sobre questões de direitos humanos, entre outras questões. O compartilhamento de dados é reconhecido como um imperativo geral porque melhora a replicabilidade e a transparência de maneira econômica, aprimora a prática de pesquisa e pode revelar novas questões de pesquisa. Visões sobre a ética do que fazer com isso, no entanto, não são compartilhadas e tendem a abraçar perspectivas deontológicas ou consequentialistas. (FOTHERGILL *et al.*, 2019, p. 4)

Nesse sentido, o modo como alguns autores (POLDRACK; GORGOLEWSKI, 2014; SALERNO *et al.*, 2017) se posicionam, sobre as questões ética envolvidas, é afirmar que em meio ao contexto dos sistemas de saúde, os indivíduos que são beneficiados por esse sistema têm o dever recíproco de não se por, e até mesmo facilitar, a coleta de seus dados. Dado que é afirmado que impedir a coleta seria uma forma de barrar o conhecimento científico. “Como um exemplo de uma estrutura ética para lidar com *Big Data*, eles se referem à *Estrutura para Compartilhamento Responsável de Dados Genômicos e Relacionados à Saúde*” (FOTHERGILL *et al.*, 2019, p. 4). Entretanto, segundo Fothergill *et al.* (2019), esse tipo de análise não leva em consideração as desigualdades sociais presentes no acesso aos sistemas de

42

Traduzido do inglês RRI (Responsible Research and Innovation) (FOTHERGILL *et al.*, 2019).

saúde, bem como aprovação dos governos sobre a possibilidade de uso destes dados por terceiros no futuro.

Em outras palavras, os seres humanos são vistos por alguns como responsáveis por doar seus dados ou permitir que eles sejam coletados (embora isso deva ser ponderado em relação aos riscos de danos que isso possa causar no futuro); os pesquisadores são percebidos como tendo o dever de compartilhar dados (embora sejam necessários mecanismos melhores para alocar seus créditos); e há necessidade de engajamento na questão do consentimento informado para seres humanos. Essas posições incorporam a ética, mas não consideram todo o espectro de questões éticas em torno dos *Big Data* ou dos fatores sociais do trabalho; cada situação operará em um contexto jurídico e cultural diferente; e não podemos prever os impactos dos *Big Data*, principalmente em relação aos grupos de risco. (FOTHERGILL *et al*, 2019, p. 4)

Além desses fatores, o uso indiscriminado da tecnologia de *Big Data* pode acentuar ainda mais desigualdades sociais, uma vez que alguns grupos humanos nunca serão representados, seja por recusa de compartilhamento de seus dados ou exclusão tácita devido a *status* social e renda. Com isso, pode ser que um conjunto de dados verdadeiramente representativo seja inalcançável, mesmo com a obtenção de conjuntos maiores de dados. Assim, “para que uma abordagem de governança de *Big Data* da neurociência seja bem-sucedida, esses dilemas e tensões devem ser reconhecidos e tratados de forma inclusiva e pragmática” (FOTHERGILL *et al*, 2019, p. 4).

De tal modo, para que questões éticas sejam aplicadas de modo apropriado, os múltiplos fatores e estágios do ciclo de vida dos dados devem ser considerados e empregados em cada uma das etapas da pesquisa.

Considerando a multimodalidade e a diversidade disciplinar na pesquisa em neurociência, usamos um conjunto de estágios contextuais sobrepostos, que podem ser iterativos e recursivos, mas são apresentados aqui em um formato linear para que as questões éticas sejam mais efetivamente mapeadas. Usamos “*Coleta de dados*” para significar a criação de dados e metadados; “*Processamento de dados*” para se referir ao registro, integração, agregação e análise de dados; “*Curadoria de dados*” significa o armazenamento, manutenção e segurança de dados; “*Aplicação de dados*” significa os resultados de qualquer aspecto do processamento de dados; “*Compartilhamento de dados*” para o fornecimento de acesso aos dados; e “*Exclusão de dados*” refere-se à destruição de dados. (FOTHERGILL *et al*, 2019, p. 5)

Tendo sido esclarecido alguns aspectos importantes à aplicação de uma *governança responsável dos dados*, os autores definem, que esta forma de governança seria um modo de abordar a aplicação da tecnologia de *Big Data* na neurociência considerando as questões éticas envolvidas em cada uma das etapas do ciclo de vida dos dados, abordando-as de modo

a reconhecer a fluidez e pluralidade de definições, as diversas perspectivas e contextos de um grande projeto de neurociência e evitando as possíveis armadilhas do imperialismo ético.

Em resumo, a noção de governança dos dados pode ser concebida como uma estratégia de gestão geral sobre a disponibilidade, usabilidade, integridade, qualidade e segurança dos dados, com o objetivo de aprimorar ao máximo o potencial uso dos dados, seguindo sempre em conformidade com questões éticas envolvidas em contextos organizacionais específicos. Enquanto, a PIR (*Pesquisa e Inovação Responsável*) é uma abordagem com princípios para guiar pesquisas e inovações, a fim de unir e responder às partes interessadas e suas expectativas. Desse modo, a aplicação da PIR à pesquisa e inovação exige que os novos projetos sejam desenvolvidos considerando as consequências futuras, a implementação de mecanismos integrados a fim de garantir e promover a reflexibilidade e o alcance das partes interessadas, estabelecendo com elas uma ação conjunta, ao passo que responde aos seus interesses e age em conformidade com estes. “Essas características aumentam o apelo do uso de PIR em projetos de pesquisa grandes e complexos, em oposição a abordagens mais rígidas como avaliação da tecnologia” (FOTHERGILL *et al*, 2019, p. 7). Desse modo, a união destes dois aspectos tem o potencial de garantir a aplicação da *governança responsável dos dados*, adotando assim uma postura de reflexão ética constante, a qual permita mitigar os riscos à privacidade dos usuários, ocasionados pela aplicação de análises de *Big Data* na neurociência.

Com uma estrutura diferente, "governança responsável dos dados" é necessária nos casos em que as implicações éticas da coleta, processamento, uso e compartilhamento de dados são difíceis de prever, mas provavelmente de importância significativa. Este é certamente o caso dos *Big Data* da neurociência, onde a governança de dados provavelmente se beneficiará de estruturas que expõem a discussão ao escrutínio, que facilitam o envolvimento de uma variedade de vozes e que incorporam incentivos para que criadores e usuários de dados alcancem as outras partes interessadas que provavelmente serão afetadas por seu trabalho. (FOTHERGILL *et al*, 2019, p. 7)

Até este ponto, podemos notar que, de fato, os *Big Data* são realmente uma tecnologia promissora na compreensão de muitos enigmas existentes, não apenas em relação a Consciência e a Intencionalidade, mas sobre muitos outros aspectos. No entanto, devido à intrínseca pessoalidade dos dados produzidos, uma reflexão ética constante que permita direcionar questões específicas para cada contexto é de extrema importância, podendo ser alcançada pelo emprego de uma *governança responsável dos dados*. Contudo, além das questões éticas envolvidas, existem diversos problemas técnicos e mudanças de ordem

cultural que precisam ser desenvolvidas antes que os *Big Data* possam começar a apresentar resultados seguros.

Segundo Sejnowski *et al* (2014, p. 1), a maior modificação de ordem cultural, ocasionada pela implementação da tecnologia de *Big Data* na neurociência, seria à modificação da produção “de muitos esforços "verticais" isolados, aplicando técnicas únicas à problemas individuais em uma única espécie, para esforços mais "horizontais", que integram dados coletados usando uma ampla gama de técnicas, problemas e espécies”. Entretanto, segundo os autores, a modificação do cenário científico da neurociência aplicada na pesquisa de *Big Data*, impõe a resolução de, ao menos, cinco questões de ordem técnica e cultural.

A primeira questão, tem relação com a variedade de métodos e técnicas utilizadas na coleta de dados neurocientíficos. “Diferentes técnicas diferem também em conceitos e vocabulários, em suposições de fundo e normas experimentais” (SEJNOWSKI *et al*, 2014, p. 2). Dessa forma, para que os distintos métodos e técnicas possam ser integrados e analisados de modo automatizado é necessário a aplicação de meios de padronização destes dados. “A padronização de dados coletados com uma única tecnologia não é trivial, tornando ainda mais difíceis as relações causais significativas entre conjuntos de dados obtidos com tecnologias muito diferentes” (SEJNOWSKI *et al*, 2014, p. 2).

A segunda questão, faz referência a ampla diversidade de modelos animais usados na neurociência, bem como a conexão entre os dados neurofisiológicos com os dados comportamentais. A diversidade de modelos animais, os quais abrangem desde insetos até a classe dos vertebrados, “tem suas virtudes distintas, e novos esforços para integrar informações entre espécies e tecnologias podem render muito. Mas isso exigirá uma apreciação aprofundada da neurobiologia comparativa e evolutiva” (SEJNOWSKI *et al*, 2014, p. 2). Além disso, a correlação de dados neurofisiológicos com dados comportamentais, tradicionalmente, é realizada restringindo-se o alcance e a complexidade de medições comportamentais, a fim de facilitar a coleta e interpretação de dados correlacionados aos neurônios gerenciáveis. “Os *Big Data* estão possibilitando o registro do mesmo conjunto de neurônios enquanto o sujeito se envolve em um conjunto muito mais rico de comportamentos. [...] O desafio é descobrir as relações causais entre os *big data* neuronais e os *big data* comportamentais” (SEJNOWSKI *et al*, 2014, p. 2).

A terceira questão, refere-se a transição cultural necessária ao modo como os dados são compartilhados entre os laboratório de neurociência. Tradicionalmente, o modo de

compartilhamento e a integração dos resultados é realizada individualmente pelos próprios laboratórios e seus colaboradores. Já em um contexto de *Big Data*, além da automatização das análises e correlações entre espécies e outras medições, como por exemplo, o comportamento, “serão necessárias equipes de pesquisadores coordenados com especialização complementar” (SEJNOWSKI *et al*, 2014, p. 3).

A quarta questão, tem relação com a proporção entre o aumento da quantidade de dados e o aumento da complexidade, sendo este um fator, que torna o processo de análise e extração de conclusões ainda mais difícil. Contudo, em um contexto de *Big Data*, as dificuldades existentes na manipulação de conjuntos muito extensos de dados podem ser mitigadas pela aplicação de algoritmos de mapeamento de padrões. Entretanto, “Este não é um processo fácil e requer um nível de planejamento e controle de qualidade que vai além da maioria dos experimentos exploratórios realizados na maioria dos laboratórios. Aqui, novamente, uma modesta mudança cultural pode causar um grande impacto” (SEJNOWSKI *et al*, 2014, p. 3).

E por fim, a quinta questão, refere-se ao papel da teoria em meio aos grandes conjuntos de dados e as análises automatizadas. Apesar da possibilidade de aprofundamento do conhecimento devido as tecnologias de *Big Data*, essas análises, na grande maioria das vezes, apresentam mais descrições sobre padrões estatísticos do que a construção de novos modelos teóricos. Desse modo, se faz necessário a emergências de “uma nova geração de pesquisadores treinados computacionalmente que estão cientes da riqueza dos dados e podem se valer do conhecimento de muitos laboratórios, [...] e imaginativos o suficiente para gerar dados interessantes e ideias testáveis em larga escala” (SEJNOWSKI *et al*, 2014, p. 4).

Ao longo desta seção, verificamos como a tecnologia de *Big Data* no campo da neurociência pode auxiliar no desvelamento de muitos enigmas, com mais agilidade e economia de recursos. Dentre os possíveis desvelamentos alcançados, está o aprofundamento do nosso conhecimento sobre a Consciência e a Intencionalidade. Entretanto, apesar do potencial auxílio gerado por essa tecnologia, é evidente que muitas dificuldades precisam ser superadas para que seja possível uma aplicação segura e responsável, abrangendo questões de ordem ética, social, cultural e técnicas. Na seção seguinte, conheceremos, um pouco mais a fundo, alguns projetos em curso que lançam mão de ferramentas de *Big Data* para o desenvolvimento do campo da Neurociência, além de evidenciar como a aplicação dos conhecimentos da Neuroética podem ser de grande auxílio para o sucesso destes projetos.

3.4 Os *Big Data* da Neurociência e sua relação com a Neuroética

Como anteriormente salientado, os *Big Data* passaram a ganhar maior repercussão e o termo passou a ser empregado com maior frequência por volta de 2010. Contudo, cabe distinguir alguns de seus contextos de aplicação, sendo importante ressaltar, nesta seção, sua aplicação no campo da Neurociência, a qual apresenta algumas características particulares em comparação a sua aplicação voltada aos negócios e à ciência de um modo geral.

Em um artigo publicado em setembro de 2013 na revista *Nature Reviews*, cinco proeminentes cientistas explicam alguns objetivos e algumas questões críticas envolvidas em quatro projetos de pesquisas em neurociência com aplicações de ferramentas de *Big Data*.

Neste artigo, os neurocientistas Christof Koch, Eric R. Kandel, Henry Markram, Paul M. Matthews e Rafael Yuste comentam sobre quatro grandes projetos em Neurociência, até então, recém iniciados. Sendo eles o *Human Brain Project*, o qual visa desenvolver simulações computacionais em grande escala do cérebro; o *Brain Activity Map*, tendo o objetivo de estabelecer um conectoma funcional de todo o cérebro; o projeto *MindScope*, sendo este um projeto de 10 anos com o objetivo de compreender o córtex visual de ratos; e o projeto *BRAIN Initiative*, o qual após ser anunciado, pelo, até então, presidente norte-americano Barack Obama, em abril de 2013, destacou o compromisso político dos EUA com a Neurociência trazendo a esperança de novas colaborações interdisciplinares, desenvolvimento de novas tecnologias e avanços médicos no campo da Neurociência.

O cientista Erick R. Kandel afirma neste artigo, que o principal objetivo destes projetos altamente ambiciosos é “obter um melhor entendimento das bases anatômicas, moleculares e de circuito para as operações lógicas realizadas pelo cérebro humano” (KANDEL *et al*, 2013, p. 659). Além disso, afirma que apesar de cada um destes projetos ainda serem iniciais na compreensão do cérebro humano, é de grande relevância a utilização de animais não-humanos em meio ao processo de experimentação, por dois motivos:

Primeiro, a aplicação de experimentos tanto em animais invertebrados simples, como vermes e moscas, bem como em vertebrados, possibilita que sejam dados os primeiros passos na compreensão de como o cérebro processa as informações sensoriais e como os movimentos são iniciados por este. Assim, possibilitando aos cientistas a começar a lidar com a organização geral da estrutura e função cerebral. O segundo motivo salientado por Kandel, refere-se aos desenvolvimentos tecnológicos, moleculares e computacionais da época, os

quais permitiriam, mediante a testagem animal, um registro mais eficiente e preciso de diversas células nervosas ao mesmo tempo, o mapeamento de conexões entre neurônios, a correlação entre estrutura e atividade neural e a manipulação da atividade neural, permitindo averiguar o papel causal dos elementos neuronais na produção do comportamento.

Logo nesta primeira consideração podemos notar que questões relacionadas a compreensão do cérebro e sua relação com a mente, dificilmente podem ser desenvolvidas sem que sejam permeadas por questões de ordem ética, principalmente, considerando questões relativas à Consciência, como veremos mais adiante ainda nesta seção.

No mesmo artigo, o neurocientista Henry Makram, um dos elaboradores do *Human Brain Project*, afirma que projetos como este, que visam simular o cérebro humano e suas doenças são de extrema relevância na compreensão dos “princípios fundamentais que regem a estrutura e função do cérebro” e nos auxiliarão a “reconstruir de forma preditiva o cérebro a partir de fragmentos de dados experimentais” (KANDEL *et al*, 2013, p. 659). Segundo Makram, a lacuna entre os laboratórios de neurociência e a clínica, bem como o limitado conhecimento sobre o cérebro no desenvolvimento de novas tecnologias computacionais são dificuldades que podem ser mitigadas com o sucesso destes projetos de pesquisa.

Segundo Makram, essas novas tentativas de reconstrução computacional do cérebro humano teriam, portanto, o potencial de promover um novo foco para a neurociência, a clínica e o desenvolvimento tecnológico. “Isso nos ajudará a “limpar” relatórios conflitantes e nos ensinará como aplicar o conhecimento dos estudos com animais para compreender o cérebro humano” (KANDEL *et al*, 2013, p. 659).

No entanto, com pouco menos de dois anos desde o início do projeto, em Janeiro de 2013, o projeto de Makram já havia sido alvo de diversas críticas pela comunidade de neurocientistas. Em um artigo escrito por Stefan Theil, publicado na revista *Scientific American*, é apontado que, em Julho de 2014, um documento foi elaborado contendo mais de 800 assinaturas de diversos cientistas criticando o projeto de Makram. As críticas ao *Human Brain Project* apontam desde problemas relativos à gestão centralizada de Makram, a pouca transparência em meio à tomada de decisão e até mesmo a falta de direcionamento científico do projeto, uma vez que o projeto não abordava nenhuma questão de pesquisa específica, teste ou hipótese científica de como o cérebro funciona.

Parte do furor em torno do projeto de Makram, teve como razão o exorbitante quantia desembolsada pela Comissão Européia, algo em torno de 1 bilhão de euros. Segundo os

críticos esse dinheiro poderia ter servido para financiar outros projetos melhor direcionados cientificamente. Após pouco menos de um ano, em Março de 2015, o comitê de mediação do projeto concordou com os críticos e iniciou a reformulação do projeto, concordando em estabelecer um número menor de atividades devidamente priorizadas. Assim, podemos notar que apesar de Makram afirmar que seu projeto teria o potencial de “limpar relatórios conflitantes”, o que acabou ocorrendo de fato é que os “os esforços do HBP e a reação intensa a eles expuseram divisões importantes em como os neurocientistas pensam sobre o cérebro e como ele deve ser estudado” (YOUNG, 2019).

Todos esses fatores levaram o projeto *Human Brain Project* a reestruturar sua proposta inicial, sendo reformulado como um projeto de *software* para curadoria dos dados existentes sobre o cérebro, o fornecimento de ferramentas para pesquisa de dados e desenvolvimento de simuladores, possibilitando a outros pesquisadores construir seus próprios modelos. “O HBP, lançado com a intenção de unificar a comunidade neurocientífica da Europa, em vez disso, trouxe rachaduras à superfície” (MAHFOUD, 2021, p. 324).

Entretanto, apesar das dificuldades enfrentadas e a controversa gerada pelo *Human Brain Project*, a expectativa na aplicação da tecnologia de *Big Data* no desenvolvimento da Neurociência não foi extinta. Outro grande projeto citado anteriormente, o projeto norte-americano *BRAIN Initiative*, apresentou resultados bem mais promissores.

No início, o projeto *Brain Initiative* também gerou desconfianças na comunidade científica dos EUA ao suspeitarem que o projeto pudesse desviar o financiamento de outros projetos de Neurociência, ou que seus objetivos pudessem ser nebulosos e inatingíveis. Porém, o que levou o *BRAIN Initiative* a produzir resultados mais promissores que seu concorrente europeu foi a proposta colaborativa assumida após a desconfiança inicial da comunidade científica. Segundo Theil (2015), “em vez de prosseguir com portas fechadas e análises confidenciais, o NIMH⁴³ reagiu às críticas, suspendendo a iniciativa e envolvendo a comunidade neurocientífica”. A atitude tomada pelo Instituto Nacional de Saúde Mental dos EUA resultou, após um ano de deliberações sobre o projeto, em um ambicioso e verdadeiro programa interdisciplinar no desenvolvimento de novas ferramentas tecnológicas, promovendo novos modos de monitorar, medir e estimular o cérebro.

A principal diferença entre o HBP da Europa e o *BRAIN Initiative* dos EUA é que o último não depende de uma única visão científica. Em vez disso, muitas equipes

43

National Institute of Mental Health

competiram por doações e conduziram a inovação em direções diferentes e não planejadas. [...] A tomada de decisão mais competitiva e transparente do *BRAIN Initiative* está muito distante da caixa preta política em Bruxelas que produziu o HBP. (THEIL, 2015, p. 25)

Além dos resultados científicos e tecnológicos, a atitude colaborativa que permeou o projeto *BRAIN Initiative*, possibilitou direcionar os rumos tomados pelo projeto segundo uma perspectiva ética mais elaborada. Segundo Eberwine e Kahn (2020, p. 136), ao alcançar a metade do prazo do projeto, o diretor do *National Institute of Health* solicitou uma revisão do progresso obtido e a criação de um conjunto de recomendações complementares às direções iniciais. A revisão científica do projeto foi elaborada por um Grupo de trabalho (WG2⁴⁴) e um Subgrupo de Neuroética (BNS⁴⁵), selecionado para desenvolver um Roteiro de Neuroética salientando as principais preocupações éticas envolvendo os avanços previstos.

A relevância da incorporação de questões dessa natureza em um projeto como esse, é que “como o cálculo de custo/benefício da pesquisa em neurociência muda conforme o campo avança rapidamente, é importante que as considerações neuroéticas da pesquisa sejam um componente integral do esforço de pesquisa” (EBERWINE; KAHN, 2020, p. 136).

Desse modo, em um artigo publicado na revista *AJOB Neuroscience*, os Copresidentes do Subgrupo de Neuroética do *BRAIN Initiative* apresentaram um breve resumo das principais áreas de interesse sobre neuroética para a comunidade do projeto *BRAIN*.

As recomendações abrangem muitas questões de neuroética e de pesquisa do *BRAIN*, incluindo as considerações sobre a privacidade, o uso de sistemas com modelos de primatas não humanos, a temática sobre modulação e aprimoramento neural, participação do sujeito na pesquisa do *BRAIN* e equidade na pesquisa em neurociência. Além disso, o grupo recomendou um projeto transformador, cujo objetivo será o detalhamento dos mecanismos científicos e dos fundamentos éticos da consciência, sendo esta uma das questões mais assustadoras que afetam as percepções que temos de nós mesmos. (EBERWINE; KAHN, 2020, p. 135)

Ainda segundo os autores, outro ponto de relevância na consideração de questões de ordem ética tem relação com o avanço científico como um todo. Pois, uma vez que a pesquisa científica avança, questões de ordem ética relacionadas a este progresso emergem e necessitam ser debatidas, a fim de que a ciência possa proceder de modo adequado.

Com relação a privacidade, assim como na aplicação geral das técnicas de *Big Data*, a manipulação e uso dos dados podem ampliar o risco de exposição de características íntimas

⁴⁴ Work Group 2

⁴⁵ Brain Neuroethics Sub-group

dos sujeitos. No caso dos *Big Data* da Neurociência, existe ainda o debate sobre quais dados ou conjunto de dados seriam classificados como possíveis identificadores dos sujeitos de pesquisa. “Embora o avanço da ciência [...] exija que a maioria dos dados seja de acesso aberto, as preocupações com a privacidade, [...] podem sugerir que alguns conjuntos de dados não devem estar disponíveis publicamente e apenas acessíveis por meio de um rigoroso processo de triagem” (EBERWINE; KAHN, 2020, p. 137).

Um aspecto ético mais específico da aplicação de *Big Data* na Neurociência é relativo a modulação e aprimoramento neural. Segundo os autores, os avanços científicos no campo da Neurociência incluem várias abordagens que atuam por meio da modulação ou controle do circuito cerebral, podendo ser aplicadas com intuito terapêutico ou no aprimoramento de capacidades cognitivas. Nesses casos, o debate Neuroético deve estabelecer os critérios, em linhas definidas, do que seria o reestabelecimento das funções cerebrais “normais” e o que seria aprimoramento puro. Além disso, em casos de aprimoramento cognitivo, é ainda necessário levar em conta quais seriam os meios aceitáveis, como por exemplo, a educação e a cafeína, quais seriam controversos como o uso de dispositivos e cirurgias. “Os aspectos conceituais dessas questões merecem atenção de pesquisa séria, assim como questões empíricas sobre os efeitos das intervenções na cognição, o uso de técnicas de aprimoramento cognitivo existentes e a percepção pública dos potenciais usos futuros” (EBERWINE; KAHN, 2020, p. 137).

As outras questões éticas abordadas no artigo, salientam a intersecção entre a Neurociência e o debate sobre as bases neurais da consciência. A primeira delas é referente ao uso de modelos animais. O uso de animais em meio à pesquisa neurocientífica é importante, pois permite imitar sintomas e características das doenças neurológicas humanas. Além disso, o uso de primatas nestas pesquisas permite com que algumas intervenções sejam analisadas empiricamente em seres com capacidades cognitivas, genéticas e fisiológicas semelhantes ao humanos.

No entanto, segundo os autores, “essas mesmas características levantam questões éticas importantes sobre sofrimento, consciência e gerenciamento de objetivos de pesquisa importantes, juntamente com responsabilidade e administração no uso de um recurso de pesquisa valioso e escasso” (EBERWINE; KAHN, 2020, p. 137).

Outra questão relevante que apresenta uma interface com o debate sobre a Consciência, diz respeito a equidade nas pesquisas em Neurociência. Como de costume, os

projetos de pesquisa com aplicação de *Big Data* comumente envolvem diversos grupos de pesquisa, os quais frequentemente são formados por indivíduos de diversas culturas e nações. Dessa forma, o princípio de equidade na pesquisa de Neurociência ressalta que diversas populações de sujeitos possam compartilhar e desenvolver suas pesquisas com oportunidades equivalentes, independente de sua cultura ou nação. “As noções de self, identidade e consciência são da responsabilidade de todos. [...] Desse modo, todos podem se beneficiar da ciência em rápida evolução e dos insights associados ao funcionamento normal e às alterações associadas às doenças” (EBERWINE; KAHN, 2020, p. 137).

Por fim, os autores ainda apresentam uma reflexão importante sobre como um grande projeto como o *BRAIN Initiative*, o qual envolve aplicação de *Big Data* no desenvolvimento da Neurociência, poderia contribuir para o desenvolvimento de uma compreensão científica da Consciência.

Uma das questões mais importantes com que os neuroeticistas são frequentemente solicitados a lidar são as ciências fundamentais, bem como as questões clínicas que se relacionam com o conceito de consciência e as questões associadas de identidade e agência. A consciência é uma das características que nos permite ter autoconsciência e vai à essência de nossa humanidade. No entanto, a definição de consciência é complexa e pode ser vista de forma diferente por cientistas, filósofos, médicos e neuroeticistas. É difícil fornecer uma base neuroética para a consciência quando mesmo a definição de consciência não é acordada. [...] Para facilitar essas discussões, o BNS propôs um projeto de pesquisa transformador para fornecer uma compreensão da consciência que ajudará neuroeticistas, neurocientistas e clínicos em suas deliberações sobre a consciência. [...] Com um objetivo comum bem definido, espera-se que um entendimento comum da base biológica da consciência esteja disponível. Tal resultado seria inestimável para avaliar as implicações neuroéticas da ciência que impactam a consciência. (EBERWINE; KAHN, 2020, p. 138)

Essa reflexão nos permite concluir que a interdisciplinaridade, o dialogo entre diferentes áreas e o trabalho colaborativo associado ao desenvolvimento tecnológico possam vir a ser os caminhos para o desenvolvimento de uma compreensão científica da Consciência e, por conseguinte, da Intencionalidade. Permite-nos ainda supor que, talvez, o papel deste dois fenômenos em meio a era dos grandes dados, esteja na realidade subscrito na intersecção entre a compreensão científica e a reflexão filosófica. E ainda, que uma definição estrita sobre estes fenômenos e suas bases neurológicas não devem ser o objetivo único com um fim em si mesmo, mas deve permitir a investigação por meio da adoção um debate democrático, plural e dinâmico, compartilhando ideias e visões de mundo, buscando o aprimoramento científico e tecnológico adequado as necessidades humanas e ao ecossistema terrestre.

CONSIDERAÇÕES FINAIS

Considerando a linha argumentativa traçada em toda esta dissertação, bem como os objetivos desta pesquisa, cabe agora delinear algumas considerações finais sobre os temas abordados. Contudo, antes de tratar diretamente sobre os objetivos propostos, algumas ressalvas devem ser feitas em relação à ideia pervasiva em torno de todos estes temas, sendo ela: a analogia entre humano-máquina e sua interação.

Ao retomarmos a citação de Weizenbaum, sobre o modo como ideias complexas podem ser deslocadas de seu contexto de pesquisa, podemos notar que o mesmo tipo de deslocamento ocorre em meio à promessa de revolução científica oferecida pelos *Big Data*. Ao ponto de surgirem propostas afirmando sobre o “fim da teoria”, como alegado por Anderson (2008) “Este é um mundo onde grandes quantidades de dados e matemática aplicada substituem todas as outras ferramentas que possam ser utilizadas. Fora com todas as teorias do comportamento humano, da linguística à sociologia. Esqueça a taxonomia, ontologia e psicologia”. Sendo então justificado que “Quem sabe por que as pessoas fazem o que fazem? A questão é que elas fazem isso e podemos rastreá-los e medi-los com uma fidelidade sem precedentes” (ANDERSON, 2008, p. 2). No entanto, como visto antes, essa forma de aplicar a tecnologia de *Big Data* ignora amplamente as questões éticas envolvidas e o papel dos cientistas na interpretação dos dados.

Por outro lado, a metáfora descrita por Weizenbaum não é apenas análoga a proposta revolucionária dos *Big Data*, mas está subscrita nesta, sendo apenas uma nova caracterização contemporânea do mesmo argumento. Grosso modo, penso que a ideia subjacente possa ser posta da seguinte forma: “os computadores podem imitar muitos processos racionais, de modo mais rápido e preciso que os humanos, logo, considerando suas potencialidades computacionais combinadas a avalanche de dados da era dos *petabytes*, devemos deixá-los analisar e sermos confiantes em suas análises, para que os dados falem por si mesmos”.

Desse modo, por considerar que a promessa de revolução científica dos *Big Data* apresenta uma forte relação com esta metáfora, que foram selecionados os fenômenos da Consciência e Intencionalidade para lidar com esta questão.

Como visto anteriormente, os fenômenos em questão foram os mais resistentes a uma abordagem funcionalista, sendo, por vezes, alegado seu caráter físico anômalo em relação a outras propriedades físicas.

Entretanto, uma vez que o primeiro objetivo desta dissertação consiste na adoção de uma abordagem monista/naturalista, que permita lidar com estes fenômenos sem que sua relevância seja mitigada, ou que sua atribuição seja amplamente extrapolada, a abordagem apresentada por Paul Thagard (1986; 1998; 2002; 2008; 2010; 2011; 2018; 2019) serviu muito bem a este objetivo.

Os principais pontos positivos da abordagem de Thagard para esta dissertação são: 1) a possibilidade de compatibilizar em muitos aspectos (mas, não todos), os pontos de vista de duas abordagens diametrais no trato da Consciência e da Intencionalidade (Searle e Dennett); 2) a conceitualização da Consciência enquanto consciência afetiva, a qual permite abordar a íntima relação entre Consciência e emoção, e, por conseguinte, a Intencionalidade, sem ocasionar em sua atribuição indiscriminada, trazendo à tona a relevância dos processos físicos e bioquímicos presentes no cérebro; e 3) a objetividade imbuída no conceito de consciência afetiva, tornando Consciência e Intencionalidade fenômenos tratáveis cientificamente e possibilitando o estreitamento da lacuna entre consciência fenomenológica e psicológica.

Como visto, ao longo do Capítulo 2, os posicionamentos de Searle e Dennett são divergentes em muitos aspectos sobre a Consciência e a Intencionalidade. Searle, com base em seu naturalismo biológico, é defensor da ontologia de primeira pessoa presente na atividade consciente, enquanto Dennett, por outro lado, postula a abordagem dos sistemas intencionais como uma ferramenta teórica na explicação e previsão do comportamento de entidades, levando-o a supor o caráter ilusório da Consciência.

Desse modo, o conceito de consciência afetiva de Thagard possibilita a integração de alguns aspectos presentes nas duas teorias. A noção de consciência afetiva, por evidenciar a íntima relação entre Consciência e emoção, leva em consideração não apenas a comunicação por disparos elétricos entre os neurônios, mas são igualmente relevantes seus processos químicos, sendo compatível com a premissa principal do naturalismo biológico. Além disso, devido à modelagem computacional, esta abordagem permite abordar cientificamente este fenômeno, levando em consideração a complexidade dos mecanismos subjacentes, aprimorando a explicação e a previsibilidade do comportamento.

No entanto, na abordagem de Thagard, a ontologia de primeira pessoa não é eliminada como na abordagem de Dennett, mas é argumentado que talvez ela não possa ser completamente respondida. “Uma teoria da consciência emocional não precisa explicar totalmente *o que é se sentir* feliz ou triste; [...] essa questão é apenas parcialmente

respondível. Mas a teoria deve ser capaz de explicar aspectos onipresentes da experiência consciente” (THAGARD, 2008, p. 812). De tal forma, a abordagem em questão, apesar de intentar estreitar a lacuna entre Consciência fenomenológica e psicológica, não nega por completo sua existência.

Em relação ao debate sobre Intencionalidade intrínseca e derivada, notamos que Searle, ao fundamentar sua teoria sobre a ontologia de primeira pessoa da consciência, estabelece uma distinção entre a Intencionalidade derivada, presente nos artefatos, da Intencionalidade intrínseca, a qual é o tipo de Intencionalidade originadora do conteúdo intencional dos artefatos. De outro modo, Dennett afirma que o único tipo de Intencionalidade passível de ser observada é a Intencionalidade derivada, contida tanto nos artefatos, quanto na mente de seus elaboradores. Neste sentido, a seleção natural é a principal responsável por coordenar a numerosas Intencionalidades primitivas das micropartes orgânicas do corpo, sendo, portanto, tão derivada quanto qualquer outro tipo de Intencionalidade.

Novamente, podemos perceber que a abordagem da competição de marcadores semânticos de Thagard, possibilita integrar aspectos de ambas teorias. Como visto, a definição de Intencionalidade enquanto uma espécie de marcador semântico, ou seja, uma representação compactada de outros tipos de representações vinculadas à situação, avaliação, ação e o *self*, possibilita explicar, não apenas como os artefatos carregam as intenções de seus elaboradores, mas como as intenções podem causar, ou falhar em causar, ações.

Desse modo, uma vez que as intenções (marcadores semânticos) são formadas valendo-se das representações vinculadas ao *self* e às emoções, e uma vez considerada a relevância da estrutura biológica do cérebro, é possível supor sua compatibilidade com a noção de Intencionalidade intrínseca de Searle. Por outro lado, para explicar as “lacunas entre intenção-ação” é salientado o caráter competitivo dos marcadores semânticos, aproximando-os da explicação pela seleção natural de Dennett, apesar de mais detalhada em termos de mecanismos, e mais restritiva aos agentes subscritos.

Além da compatibilidade com as outras duas propostas apresentadas, uma vantagem adicional da proposta de Thagard, diz respeito a possibilidade de compatibilizar o conceito de consciência afetiva com a intuição moral, destacando algumas questões éticas relevantes em meio a uma compreensão científica da Consciência. Outro ponto interessante tem referência à vasta experiência do pesquisador com a modelagem computacional. Thagard em seu artigo sobre a consciência afetiva, escrito junto a Aubie, propõe uma conjunção de alguns de seus

modelos computacionais (os modelos ANDREA, GAGE e NECO) a fim de formar um único e integrado modelo neurologicamente plausível da consciência afetiva, o modelo em questão é conhecido como EMOCON. Sendo afirmado que este poderia auxiliar em grande parte na testagem de algumas hipóteses relativas a sua aceitação da Consciência enquanto consciência afetiva. Tal proposta, poderia ainda, ser configurada enquanto um projeto de pesquisa em Neurociência com aplicação de *Big Data*, semelhante ao *Human Brain Project*, mas com objetivos muito mais esclarecidos.

A tarefa agora é combinar os muitos componentes e mecanismos neurais discutidos na última seção em um mecanismo integrado capaz de explicar uma ampla gama de fenômenos da consciência emocional. [...] Ainda não programamos uma simulação tão grande e complicada, mas vamos extrapolar das partes que estão funcionando em modelos mais simples para oferecer explicações da experiência emocional. (THAGARD; AUBIE, 2008, p. 817)

Portanto, por restringir aos organismos a manifestação genuína de Consciência e Intencionalidade, bem como por permitir estreitar a lacuna entre a concepção de Consciência fenomenológica e psicológica, por meio da adoção de uma abordagem evidentemente interdisciplinar, é que a abordagem defendida por Paul Thagard se torna interessante a presente pesquisa. Uma vez que esta possibilita uma investigação positiva das duas hipóteses levantadas.

Desse modo, dado que a elaboração de respostas a essas duas questões constituem-se como objetivos da presente pesquisa, apresentarei uma discussão sobre o grau de satisfação das respostas apresentadas.

Com relação a primeira questão, afirmo que as discussões travadas ao longo do terceiro capítulo são plausíveis e consistentes, uma vez que estudos sobre a Consciência e Intencionalidade tem potencial suficiente para fundamentar as críticas em relação às questões éticas da aplicação da tecnologia de *Big Data*, tanto no que diz respeito aos riscos a privacidade, liberdade e autonomia dos usuários das redes.

A relação entre questões éticas sobre a privacidade dos usuário e o estudo da Consciência conectam-se devido às dificuldades em se conseguir um consentimento informado aos usos secundários dos dados. Como vimos, na era dos grandes dados, a maior parte do uso dos dados são secundários, e sua aplicação, muitas vezes, não é definida no momento da coleta, tornando-se difícil informar aos usuários quais possíveis usos secundários que seus dados poderiam vir a ter e, conseqüentemente, que o usuário possa ter consciência sobre o que estará sendo autorizado no momento da coleta..

A proposta de aplicação de análises preditivas de *Big Data* em contextos penais, também pode ser melhor refletida com base nos estudos da Consciência e Intencionalidade. Como visto, a previsão de comportamentos criminosos, com base em análise probabilística, e aplicação de sentenças com bases nessas previsões ignoram amplamente o papel da Consciência e da Intencionalidade e, por conseguinte, violam questões éticas relacionadas ao nosso senso de justiça, responsabilidade moral e livre-arbítrio.

Além disso, se considerarmos a possibilidade de manipulação psicológica e o rastreamento de perfis com o intuito de direcionamento de conteúdo específico, já nos posicionamos a fim de observar o grande risco que os *Big Data*, em seu sentido geral, oferecem a privacidade, liberdade e autonomia dos usuários. Sendo ainda mais intensos, ao considerarmos as implicações de sua aplicação no campo da Neurociência. Uma vez que neste contexto, além de nos depararmos com dados ainda mais sensíveis, devemos nos atentar ao debate ético envolvendo as questões sobre o uso de sistemas com modelos de primatas não humanos, a temática sobre modulação e aprimoramento neural, a participação do sujeito na pesquisa e a equidade na pesquisa em neurociência, como salientado pelos codenadores do subgrupo de Neuroética do *BRAIN Initiative*.

Por fim, cabe salientar que o debate da Consciência e da Intencionalidade é extenso e com muitos posicionamentos divergentes, alguns dos quais jamais permitiriam que a aplicação da tecnologia de *Big Data* auxiliassem em algum sentido a exploração desse fenômeno. Entretanto, assumir um posicionamento negativo desse tipo não é interessante a presente pesquisa. De outro modo, as abordagens de Thagard sobre a consciência afetiva e os marcadores semânticos da intencionalidade são passíveis de serem aprofundadas em um contexto de *Big Data*, sendo então compatíveis ao objetivo proposto.

Contudo, como visto nas seções 3.3 e 3.4, a estimativa é que o campo da neurociência, sendo uma área que produz dados massivamente, pode ser muito privilegiada com a aplicação da tecnologia de *Big Data*. Entretanto, a sensibilidade e pessoalidade dos dados produzidos, a variedade de métodos, vocabulários, modelos animais e tipos de dados, além de diversas mudanças de ordem cultural necessárias a uma atuação eficiente no compartilhamento destes dados, são impecílios que precisam ser melhor delineados antes que os *Big Data* possam começar apresentar resultados satisfatórios.

De qualquer modo, estas são dificuldades laboriosas, mas corrigíveis. E uma vez que consideremos o ganho potencial que a aplicação de mapeamento automatizado,

compartilhamento entre diversos laboratórios de pesquisas e possibilidade de modelagem computacional com riquíssimos bancos de dados, não há razão para pressupor o fracasso da aplicação da tecnologia de *Big Data* no campo da neurociência.

No entanto, com o intuito de ser coerente com os pressupostos direcionadores desta pesquisa, penso que a reflexão sobre as questões de ordem ética e neuroética são prioridades no que diz respeito a aplicação de *Big Data* no campo da neurociência, tendo em vista os possíveis riscos que a pessoalidade deste dados carregam na manutenção da livre-escolha por parte dos indivíduos, além de outros pontos ligados intimamente a questão da Consciência, como destacado na seção 3.4. Logo, penso ser plausível afirmar que os *Big Data* tem um potencial imenso no aprofundamento de nossa compreensão de vários aspectos da estrutura neurofisiológica, como a conexão entre o comportamento e emergência das emoções, além da própria Consciência e Intencionalidade. Mas, por outro lado, muitos riscos devem ser corrigidos e muito trabalho ainda deve ser feito para que a aplicação responsável desta tecnologia seja alcançada com um amplo grau de sucesso.

REFERÊNCIAS BIBLIOGRÁFICAS

ARAÚJO, R. B. **Computação Ubíqua: Princípios, Tecnologias e Desafios**. *XXI Simpósio Brasileiro de Redes de Computadores*, 2003-Departamento de Computação – Universidade Federal de S. Carlos (UFSCar) - São Carlos – SP – Brasil.

CHALMERS, David J. **The Conscious Mind: in search of a fundamental theory**. Oxford: Oxford University Press, 1996. 1-414 p. ISBN 0-19-510553-2.

CHURCHLAND, Paul M. **Matéria e consciência: Uma introdução contemporânea à Filosofia da mente**. São Paulo: Editora UNESP, 2004. 286 p. ISBN 85-7139-519-5.

_____. Paul M. Some Reductive Strategies in Cognitive Neurobiology. *In: SILVERS, Stuart. Representatation: Readings in the Philosophy of Mental Representation*. [S. l.]: Philosophical Studies Series, 1989. v. 40, cap. 12, p. 279-309. ISBN 978-94-009-2649-3. Disponível em: https://link.springer.com/chapter/10.1007/978-94-009-2649-3_12#citeas. Acesso em: 1 jan. 2020.

CORDIOLI, Aristides Volpato. **Manual diagnóstico e estatístico de transtornos mentais: DSM-5**. 5. ed. rev. Porto Alegre: Artmed, 2014. 948 p. ISBN 978-85-8271-089-0.

DAMASIO, A. R. **The feeling of what happens: Body and emotion in the making of consciousness**. New York: Harcourt Brace, 1999.

DAMILANO, José Luiz Padilha *et al.* Intention, Emotion, and Action: Primeiras noções e conceitos básicos. *In: DAMILANO, José Luiz Padilha. Fundamentos neuropsicológicos da aprendizagem: 3º semestre*. 1. ed. Santa Maria: Gráfica e Editora Pallotti, 2005. cap. 1, p. 09-11. Disponível em: <https://repositorio.ufsm.br/bitstream/handle>. Acesso em: 10 nov. 2020.

DENNETT, Daniel C. **Brainstorms: Ensaio Filosófico sobre a Mente e a Psicologia**. São Paulo: Editora UNESP, 2006. 431 p. ISBN 85-7139-645-0.

_____. **Consciousness Explained**. New York, USA: Penguin Books, 1991. 492 p. ISBN 0-14-012867-0.

_____. **Tipos de mentes: Rumo a uma compreensão da consciência**. Rio de Janeiro: Editora Rocco, 1997. 166 p. ISBN 85-325-0747-6.

DESCARTES, René. **Discurso do método**. 3. ed. rev. e aum. São Paulo: Martins Fontes, 2007. 123 p. ISBN 978-85-336-2363-7.

_____. **Meditações Metafísicas**. 1. ed. rev. São Paulo: Martins Fontes, 2000. 155 p. ISBN 85-336-1204-4.

ELIASMITH, Chris. **How to Build a Brain: A Neural Architecture for Biological Cognition**. Oxford University: Oxford University Press, 2015. ISBN 9780190262129.

KLEIN, Julie Thompson. **Crossing Boundaries: Knowledge, Disciplinarity, and Interdisciplinarity**. 2. ed. USA: The University Press of Virginia, 1996. 281 p. ISBN 0-8139-1679-8.

JAMES, William. **Psychology: Briefer course**. New York, Henry Holt & Co., 1892.

Manual diagnóstico e estatístico de transtornos mentais [recurso eletrônico]: **DSM-5**/ [American Psychiatric Association ; tradução: Maria Inês Corrêa Nascimento...et al.] ; revisão técnica: Aristides Volpato Cordioli...[et al.].—5.ed.— Dados eletrônicos.—Porto Alegre: Artmed, 2014.

PUTNAM, Hilary. **Representation and Reality**. 5. ed. Cambridge, UK: MIT Press, 1996. 136 p. ISBN 0-262-66074-1.

_____. Minds and Machines. In: HOOK, Sidney. **Dimensions of Minds**. New York, USA: New York University Press, 1960. cap. 2, p. 20-33. Disponível em: <https://philarchive.org/archive/PUTMAMv1>. Acesso em: 14 jan. 2020.

ROSENTHAL, David M. **Materialism and the mind-body problem**. 2. ed. Indianapolis, USA: Hackett Publishing Company, 2000. 315 p. ISBN 0-87220-479-0.

RYLE, Gilbert. **The Concept of Mind**. New York, USA: Penguin Books, 2000. 316 p. ISBN 978-0-14-118217-9.

SCHÖNBERGER, Viktor Mayer-; CUKIER, Kenneth. **BIG DATA: A Revolution That Will Transform How We Live, Work, and Think**. New York, USA: Houghton Mifflin Harcourt Publishing Company, 2014. 252 p. ISBN 978-0-544-22775-0.

SEARLE, John R. **Intentionality: An essay in the Philosophy of mind**. Cambridge, UK: Cambridge University Press, 1997. 278 p. ISBN 0-521-27302-1.

_____. **Mind: A brief introduction**. New York, USA: Oxford University Press, 2004. 224 p. ISBN 0-19-515734-6.

_____. **O mistério da Consciência: E discussões com Daniel C. Dennett e David J. Chalmers**. São Paulo: Paz e Terra, 1998. 239 p. ISBN 85-219-0305-7.

_____. **A Redescoberta da Mente**. 1. ed. rev. São Paulo: Martins Fontes, 1997. 379 p. ISBN 85-336-0619-2.

TEIXEIRA, João de Fernandes. **Filosofia do cérebro**. 1. ed. São Paulo: Paulus, 2012. 93 p. ISBN 978-85-349-3332-2.

TEIXEIRA, João de Fernandes. **Mentes e Máquinas: Uma introdução à ciência cognitiva**. Porto Alegre: Artes Médicas, 1998. 179 p. ISBN 85-7307-329-2.

THAGARD, Paul. **MENTE: Introdução à Ciência Cognitiva**. Porto Alegre: ArtMed, 1998. 210 p. ISBN 85-7307-460-4.

WEIZENBAUM, Joseph. **Computer Power and Human Reason**: From Judgement to Calculation. New York, USA: Penguin Books, 1985. 300 p. ISBN 0-14-022535-8.

WOLFF, Francis. **Nossa Humanidade**: De Aristóteles às Neurociências. São Paulo: Editora UNESP, 2012. 335 p. ISBN 978-85-393-0370-0.

ARTIGOS

ABRANTES, Paulo. Naturalismo em filosofia da mente. **In**: Ferreira, A.; Gonzalez, M. E. Q.; Coelho, J. C. (Org.). Encontro com as Ciências Cognitivas. São Paulo: Cultura Acadêmica, 2004. v. 4, p. 5-37.

ANDRADE, Leonardo F. C. de. Cognição, Emoção e Consciência: Segundo a abordagem de Paul Thagard. **Simbio-Logias**: Teorias da Consciência, Botucatu, UNESP, 2021, *no prelo*.

ANDERSON, Chris. The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. **Wired Magazine**, [S. l.], p. 1-3, 23 jun. 2008. DOI 10.1098/rsta.2011.0325. Disponível em: http://www.wired.com/science/discoveries/magazine/16-07/pb_theory. Acesso em: 4 fev. 2020.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados Pessoais (LGPD). [S. l.], 14 ago. 2018. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/113709.htm. Acesso em: 21 jul. 2021.

BURKS, A. An architectural theory of human consciousness. **Logic of Computers**. Ann Arbor, MI: University of Michigan, 1984.

CONSCIENCE: What is Moral Intuition?. *In*: CHURCHLAND, Patricia S. **Conscience**: The Origins of Moral Intuition. [S. l.]: W. W. Norton & Company, 2011. cap. 7, p. 150-169. ISBN 1324000899. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html>. Acesso em: 24 set. 2020.

DE MOL, Liesbeth. Turing Machines. **The Stanford Encyclopedia of Philosophy**, [s. l.], p. 1-7, 7 dez. 2019. Disponível em: <https://plato.stanford.edu/archives/win2019/entries/turing-machine/>. Acesso em: 14 jul. 2021.

EBERWINE, James; KAHN, Jeffrey. The BRAIN Initiative and Neuroethics: Enabling and Enhancing Neuroscience Advances for Society. **A JOB Neuroscience**, [s. l.], v. 11, n. 3, p. 135-139, 15 jul. 2020. DOI 10.1080/21507740.2020.1778121. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/21507740.2020.1778121?scroll=top&needAccess=true>. Acesso em: 23 jul. 2021.

FOTHERGILL, B. Tyr *et al.* Responsible Data Governance of Neuroscience Big Data. **Frontier Neuroinformatics**, [s. l.], p. 1-14, 2019. DOI <https://doi.org/10.3389/fninf.2019.00028>. Disponível em: <https://www.frontiersin.org/articles/10.3389/fninf.2019.00028/full>. Acesso em: 15 mar. 2021.

GONZALEZ, Maria Eunice Quilici; BROENS, Mariana C.; MORAES, João Antonio de. A virada informacional na Filosofia: Alguma novidade no estudo da Mente?. **Revista de Filosofia**, Aurora, Curitiba, v. 22, ed. 30, p. 137-151, 3 maio 2010. DOI 10.7213/rfa.v22i30.2230. Disponível em:

https://www.researchgate.net/publication/291828341_A_virada_informacional_na_Filosofia_Alguma_novidade_no_estudo_da_Mente. Acesso em: 21 fev. 2020.

JACOB, Pierre, "Intentionality", **The Stanford Encyclopedia of Philosophy** (Winter 2019 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/win2019/entries/intentionality/>>. Acesso em: 29 jan. 2021.

JACKSON, Frank. Epiphenomenal Qualia. **The Philosophical Quarterly**, [s. l.], v. 32, ed. 127, p. 127-136, 1982. DOI 10.2307/2960077. Disponível em: <https://www.jstor.org/stable/2960077>. Acesso em: 15 jan. 2020.

JOHNSON-LAIRD, P. A computational analysis of consciousness. **Cognition and Brain Theory**, 6, 499-508, 1983.

JUARRERO, Alicia. Dynamics in Action: Intentional Behavior as a Complex System. **The British Journal for the Philosophy of Science**, [s. l.], v. 52, n. 3, p. 635-640, 4 set. 2001. Disponível em: <https://www.jstor.org/stable/3541935>. Acesso em: 28 jul. 2021.

KANDEL, Eric R. *et al.* Neuroscience thinks big: and collaboratively. **NATURE REVIEWS**, [S. l.], v. 14, 18 set. 2013. NEUROSCIENCE, p. 659-664. Disponível em: <https://www.nature.com/articles/nrn3578>. Acesso em: 23 jul. 2021.

KOSINSKI, Michal *et al.* Private traits and attributes are predictable from digital records of human behavior. **PNAS**, [s. l.], v. 10, n. 15, p. 5802-5805, 9 abr. 2013. DOI 1218772110. Disponível em: www.pnas.org/cgi/doi/10.1073/pnas.1218772110. Acesso em: 18 jun. 2018.

KOSINSKI, Michal *et al.* Psychological targeting as an effective approach to digital mass persuasion. **PNAS**, [s. l.], v. 114, n. 48, p. 12714-12719, 28 nov. 2017. DOI 1710966114. Disponível em: <https://www.pnas.org/content/114/48/12714>. Acesso em: 18 jun. 2018.

LANDHUIS, Esther. Neuroscience: Big brain, big data. **Nature**, [s. l.], n. 541, p. 559-561, 2017. Disponível em: <https://www.nature.com/articles/541559a>. Acesso em: 16 jul. 2020.

LEE, J.A.N. Joseph Weizenbaum. **Computer Pioneers**, Los Alamitos, p. 245-246, 1995. Disponível em: <https://history.computer.org/pioneers/index.html>. Acesso em: 16 jan. 2020.

LEONELLI, Sabina. Scientific Research and Big Data. **The Stanford Encyclopedia of Philosophy**, [s. l.], Summer 2020. Edward N. Zalta (ed.), Disponível em: <https://plato.stanford.edu/archives/sum2020/entries/science-big-data/>. Acesso em: 4 nov. 2020.

MAHFOUD, Tara. Visions of unification and integration: Building brains and communities in the European Human Brain Project. **New media & society**, [s. l.], v. 23, ed. 2, p. 322-343, 9 fev. 2021. DOI 10.1177/1461444820929576. Disponível em: <https://journals.sagepub.com/doi/full/10.1177/1461444820929576>. Acesso em: 30 jun. 2021.

MAN, Kingson; DAMASIO, Antonio. Homeostasis and soft robotics in the design of feeling machines. **Nature Machine Intelligence**, [s. l.], v. 1, p. 446-452, 9 out. 2019. DOI <https://doi.org/10.1038/s42256-019-0103-7>. Disponível em: <https://www.nature.com/articles/s42256-019-0103-7>. Acesso em: 6 fev. 2020.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, Chicago, USA, v. 5, p. 115-133, 23 dez. 1943. DOI <https://doi.org/10.1007/BF02478259>. Disponível em: <https://link.springer.com/article/10.1007/BF02478259#citeas>. Acesso em: 29 jan. 2020.

NAGEL, Thomas. What Is It Like to Be a Bat?. **The Philosophical Review**, [s. l.], v. 83, ed. 4, p. 435-450, 1974. DOI 10.2307/2183914. Disponível em: <https://www.jstor.org/stable/2183914>. Acesso em: 3 jan. 2020.

PACHECO, Claudia Núñez-. Expanding our Perceptual World Through Technology: A Subjective Bodily Perspective. **The 2015 ACM International Joint Conference**, OSAKA, JAPAN, p. 951-956, 7 set. 2015. DOI 10.1145/2800835.2806206. Disponível em: [https://www.researchgate.net/publication/301419603_Expanding_our_perceptual_world_thro](https://www.researchgate.net/publication/301419603_Expanding_our_perceptual_world_through_technology)
[ugh_technology](https://www.researchgate.net/publication/301419603_Expanding_our_perceptual_world_thro). Acesso em: 15 jan. 2020.

POLDRACK, Russell A; GORGOLEWSKI, Krzysztof J. Making big data open: data sharing in neuroimaging. **Nature Neuroscience**, [s. l.], n. 17, p. 1510-1517, 2014. Disponível em: <https://www.nature.com/articles/nn.3818>. Acesso em: 15 mar. 2021.

PLACE, Ullin T. Is consciousness a brain process?. **The Mind-Brain Identity Theory**, London, p. 44-50, 1956. DOI https://doi.org/10.1007/978-1-349-15364-0_2. Disponível em: https://link.springer.com/chapter/10.1007/978-1-349-15364-0_2#citeas. Acesso em: 10 jan. 2020.

PORTMESS, Lisa; TOWER, Sara. Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data. **Ethics and Information Technology**, [s. l.], p. 1-9, 6 dez. 2015. DOI <https://doi.org/10.1007/s10676-014-9357-2>. Disponível em: <https://link.springer.com/article/10.1007/s10676-014-9357-2#citeas>. Acesso em: 10 fev. 2020.

QUILICI-GONZALEZ, José Artur et al . Complexity and information technologies: an ethical inquiry into human autonomous action. **Sci. stud.**, São Paulo , v. 12, n. spe, p. 161-179, 2014 . Available from http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1678-31662014000500009&lng=en&nrm=iso. access on 08 Mar. 2020. <https://doi.org/10.1590/S1678-31662014000400009>.

ROBINSON, William. Epiphenomenalism. **The Stanford Encyclopedia of Philosophy**, [s. l.], p. 1-7, 18 jun. 2019. Disponível em: <https://plato.stanford.edu/archives/sum2019/entries/epiphenomenalism/>. Acesso em: 29 jul. 2021.

SALERNO, Jennifer *et al.* Ethics, big data and computing in epidemiology and public health. **Ann. Epidemiol**, [s. l.], p. 297–301, 2017. DOI 10.1016/j.annepidem.2017.05.002. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28595734/>. Acesso em: 18 mar. 2021.

SANCOVSCHI, Beatriz. Sobre a Noção de Representação em S. Moscovici e F. Varela. **Psicologia & Sociedade**, Rio de Janeiro, v. 19, ed. 2, p. 7-14, 3 jul. 2007. Disponível em: <https://www.scielo.br/j/psoc/a/tdNQmsxRXS5sBfCWQ7QBH8N/?lang=pt>. Acesso em: 4 ago. 2021.

SCHRÖDER, Tobias; STEWART, Terrence C.; THAGARD, Paul. Intention, Emotion, and Action: A Neural Theory Based on Semantic Pointers. **Cognitive Science**, [s. l.], v. 38, p. 851-880, 2014. DOI 10.1111/cogs.12100. Disponível em: <http://cogsci.uwaterloo.ca/Articles/Pages/Emotion-index.html>. Acesso em: 25 set. 2020.

SEARLE, Jonh R. Minds, Brains, and Programs. **Behavioral and Brain Sciences**, [s. l.], v. 3, n. 3, p. 417-457, 1980. Disponível em: <https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/abs/minds-brains-and-programs/DC644B47A4299C637C89772FACC2706A>. Acesso em: 18 dez. 2019.

SEJNOWSKI, Terrence J *et al.* Putting big data to good use in neuroscience. **Nature Neuroscience**, [s. l.], n. 17, p. 1440–1441, 2014. Disponível em: <https://www.nature.com/articles/nn.3839>. Acesso em: 18 mar. 2021.

SHANNON, Claude E. Symbolic Analysis of Relay and Switching Circuits. **AIEE summer convention**, Washington. D.C ., p. 471-495, 24 jun. 1939. Disponível em: <https://www.google.com.br/url?sa=t&rct=j&q=&esrc=s&source=web&cd=9&cad=rja&uact=8&ved=2ahUKEwik4KC7nY7oAhXgCrkGHTUEDQ8QFjAlegQIBRAB&url=http%3A%2F%2Fwww.ccapitalia.net%2Fdescarga%2Fdocs%2F1938-shannon-analysis-relay-switching-circuits.pdf&usq=AOvVaw06BUMhdPMo-Rr3dwWB37-p>. Acesso em: 29 jan. 2020.

SHAPIRO, Lawrence; SPAULDING, Shannon. Embodied Cognition. **The Stanford Encyclopedia of Philosophy**, [s. l.], p. 1-6, 23 jul. 2021. Disponível em: <https://plato.stanford.edu/archives/fall2021/entries/embodied-cognition/>. Acesso em: 10 ago. 2021.

SIEWERT, Charles, "Consciousness and Intentionality", **The Stanford Encyclopedia of Philosophy** (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/consciousness-intentionality/>. Acesso em: 29 jan. 2021.

SMART, John Jamieson C. Sensations and Brain Processes. **The Philosophical Review**, [s. l.], v. 68, ed. 2, p. 141-156, 1959. DOI 10.2307/2182164. Disponível em: <https://www.jstor.org/stable/2182164>. Acesso em: 11 jan. 2020.

SWAN, Melaine. THE QUANTIFIED SELF: Fundamental Disruption in Big Data Science and Biological Discovery. **Big Data**, [s. l.], v. 1, ed. 2, p. 85-99, 10 jun. 2013. DOI 10.1089/big.2012.0002. Disponível em: <https://www.semanticscholar.org/paper/The-Quantified-Self%3A-Fundamental-Disruption-in-Big-Swan/23a57b1e2beb4235d2020ed57f484c947e3d0816>. Acesso em: 25 jan. 2020.

THAGARD, Paul. Parallel Computation and The Mind-Body Problem. **COGNITIVE SCIENCE**, [s. l.], v. 10, p. 301-318, 1986. Disponível em: <http://cogsci.uwaterloo.ca/parallel.1986.pdf>. Acesso em: 18 jun. 2020.

_____, Paul. How Molecules Matter to Mental Computation. **Philosophy of Science**, University of Chicago, v. 69, n. 3, p. 497-518, 2002. Disponível em: <http://www.jstor.org/stable/10.1086/342452>. Acesso em: 23 set. 2020.

_____, Paul; AUBIE, Brandon. Emotional consciousness: A neural model of how cognitive appraisal and somatic perception interact to produce qualitative experience. **Consciousness and Cognition**, [s. l.], v. 17, p. 811-834, 2008. DOI 10.1016/j.concog.2007.05.014. Disponível em: www.elsevier.com/locate/concog. Acesso em: 24 set. 2020.

_____, Paul. How Brains Make Mental Models. **Model-Based Reasoning in Science & Technology**, [s. l.], n. 314, p. 447-461, 2010. Disponível em: https://link.springer.com/chapter/10.1007/978-3-642-15223-8_25. Acesso em: 10 mar. 2021.

_____, Paul. Mind, Consciousness, and Free Will. **Frontiers of Philosophy in China**, [s. l.], p. 377-393, 2018. DOI <https://doi.org/10.3868/s030-007-018-0029-2>. Disponível em: https://brill.com/view/journals/fphc/13/3/article-p377_7.xml. Acesso em: 10 mar. 2021.

_____, Paul. Cognitive Science. **Stanford Encyclopedia of Philosophy**, Stanford University, p. 1-10, 24 set. 2019. Disponível em: <https://plato.stanford.edu/entries/cognitive-science/>. Acesso em: 13 jan. 2020.

THEIL, Stefan. Why the Human Brain Project Went Wrong: and How to Fix It. **Scientific American**, [S. l.], p. 23-35, 1 out. 2015. Disponível em: <https://www.scientificamerican.com/article/why-the-human-brain-project-went-wrong-and-how-to-fix-it/>. Acesso em: 22 jul. 2021.

WEISER, Mark. The Computer for the 21st Century. **Scientific American**, [s. l.], p. 94 - 104, September 1991. Disponível em: <https://www.lri.fr/~mbl/Stanford/CS477/papers/Weiser-SciAm.pdf>. Acesso em: 10 jan. 2020.

WILSON, Robert A. Embodied Cognition. **The Stanford Encyclopedia of Philosophy**, [s. l.], p. 1-7, 12 abr. 2017. Disponível em: <https://plato.stanford.edu/archives/spr2017/entries/embodied-cognition/>. Acesso em: 16 mar. 2021.

YOUNG, Ed. The Human Brain Project Hasn't Lived Up to Its Promise. **The Atlantic**, [S. l.], p. 45-52, 22 jul. 2019. Disponível em: <https://www.theatlantic.com/science/archive/2019/07/ten-years-human-brain-project-simulation-markram-ted-talk/594493/>. Acesso em: 22 jul. 2021.