

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE FILOSOFIA
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

MARIA LUIZA IENNACO DE VASCONCELOS

A inter-relação entre o Modelo dos Múltiplos Esboços e o
Processamento Preditivo para o estudo da consciência

São Paulo

2023

Versão Corrigida

MARIA LUIZA IENNACO DE VASCONCELOS

A inter-relação entre o Modelo dos Múltiplos Esboços e o
Processamento Preditivo para o estudo da consciência

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia do Departamento de Filosofia da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Filosofia sob a orientação do Prof. Dr. Osvaldo Frota Pessoa Junior.

São Paulo

2023

Versão Corrigida

ENTREGA DO EXEMPLAR CORRIGIDO DA DISSERTAÇÃO/TESE**Termo de Anuência do (a) orientador (a)****Nome do (a) aluno (a): Maria Luiza Iennaco de Vasconcelos****Data da defesa: 12/01/2023****Nome do Prof. (a) orientador (a): Osvaldo Frota Pessoa Junior**

Nos termos da legislação vigente, declaro **ESTAR CIENTE** do conteúdo deste **EXEMPLAR CORRIGIDO** elaborado em atenção às sugestões dos membros da comissão Julgadora na sessão de defesa do trabalho, manifestando-me **plenamente favorável** ao seu encaminhamento ao Sistema Janus e publicação no **Portal Digital de Teses da USP**.

São Paulo, 8/3/2023



(Assinatura do (a) orientador (a))

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catalogação na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

I22i	<p>Iennaco, Maria Luiza A inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo para o estudo da consciência / Maria Luiza Iennaco; orientador Osvaldo Frota Pessoa Junior - São Paulo, 2023. 111 f.</p> <p>Dissertação (Mestrado)- Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo. Departamento de Filosofia. Área de concentração: Filosofia.</p> <p>1. Consciência (Percepção). 2. Ciência Cognitiva. 3. Neurociências. 4. Filosofia da Mente. 5. Cibernética. I. Pessoa Junior, Osvaldo Frota, orient. II. Título.</p>
------	---

Agradecimentos

Tenho muito que agradecer a todos aqueles que estiveram comigo nos bons e maus momentos dessa jornada, sobretudo aqueles que, ainda que fisicamente distantes, me deram o apoio e o conforto de jamais estar sozinha. Agradeço aos meus amigos conterrâneos Thiago Ladeira (o qual foi crucial na minha transição da psicologia para a filosofia), Felipe Junqueira, Sofia Junqueira, Bruna Pengo, Igor Ferraz, Márcio Machado, Clovis Frias e Gustavo Machado. Aos amigos que fiz na graduação, agradeço a Rayssa Maluf, por ter me introduzido à vida da pós-graduação, Leonardo Guedes, Matheus Menegucci, Igor Monsores, Tomas Civinelli e Clara Ladeira da UFJF, e também a Thainan Vargês e Yasmin Torres que me acompanharam nas aventuras como professora de inglês em Juiz de Fora. Agradeço às minhas queridas psicólogas Ana Luisa Casela e Erika Pizziolo por terem me acolhido nos dias difíceis e guiado o meu desenvolvimento pessoal. Agradeço aos amigos Eduardo Quirino e João César Ramos, que me apoiaram nos meus primeiros passos dentro da filosofia. Aos amigos que fiz durante o Mestrado, agradeço a Douglas de Freitas pela elucidação política e literária e pela companhia, Thais Rodrigues, Paulo Sayeg, Fernanda Oyakawa e Pedro Gloza, bem como José Camilo pelos agradáveis encontros acadêmicos e sociais.

Agradeço ao meu companheiro de vida Thales Silva que sempre acreditou no meu potencial acadêmico e certamente é o grande responsável por todos os passos que dei no Mestrado. Agradeço a ele por ter me ouvido, por ter debatido comigo e desembaraçado meus pensamentos, além de ter sido meu mais fiel leitor e meu mais rígido corretor. Sem ele eu simplesmente não estaria na Academia. Agradeço também aos meus gatos Floki e Pat Churchland por serem meus exímios ansiolíticos e a alegria da minha rotina. Agradeço à minha mãe Daniela Vasconcelos e à minha avó Darcy Vasconcelos que me criaram e me deram as condições necessárias para conquistar meus objetivos. Agradeço aos meus sogros, Minion Silva e Rodolfo Silva pelo apoio que me deram em São Paulo.

O meu interesse pela filosofia da consciência origina-se de uma incessante admiração pelo comportamento humano e, para chegar em uma perspectiva tão inesperada e distinta daquela com a qual iniciei meu trajeto, tenho que agradecer aos brilhantes professores que tive. Agradeço principalmente ao professor Saulo Araújo (UFJF) que me fez canalizar essa admiração para o nosso comportamento *interno*, isto é, para a nossa consciência, sendo esta, desde então, meu principal objeto de estudo. Agradeço ao professor Vitor Gomes (UERJ) por ter me mostrado a neurociência da consciência e ao professor Lelio Lourenço (UFJF) por ter

me lembrado, através da psicologia clínica, da importância de considerar o ser humano que resulta desses emaranhamentos neuronais corporificados que estive tão afoita para estudar. Agradeço também à professora Briseida Resende (IPUSP) por ter me ensinado tanto e desconstruído positivamente minha postura acadêmica e política em apenas uma disciplina. Agradeço imensamente pelas contribuições dos professores Gustavo Leal-Toledo (UFSJ) e Thiago França (UNIFESP) ao meu trabalho e ao meu desenvolvimento acadêmico. Finalmente, agradeço ao meu professor de inglês André Luiz Fernandes por não somente ter me ensinado o idioma que me permitiu chegar até aqui, mas por ter me ensinado a valorizar a história e as culturas do mundo e a ter prazer pela leitura e pelo constante crescimento intelectual.

Agradeço a Daniel Friedman, Alex Vyatkin, Ivan Metelkin, Dave Douglass e Avel Carlut, colegas do *Active Inference Institute*, um ambiente virtual que me acolheu no início do Mestrado e desenvolveu meu conhecimento sobre o Processamento Preditivo e habilidades acadêmicas diversas. Agradeço ao grupo de estudos sobre o Processamento Preditivo, que não somente acrescentou ao meu conhecimento do tema, mas me preparou para lidar com apresentações em público e discussões acadêmicas. Agradeço também ao grupo de leituras enativistas, coordenado pelo querido professor Ralph Bannell, não só pelo acolhimento, mas por instigar uma ampliação dos meus interesses acadêmicos e por me fazer vislumbrar algumas interessantes aplicações do meu tema de estudo.

Agradeço especialmente ao meu orientador Osvaldo Pessoa que, independente do horário ou do dia, estava presente para tirar minhas dúvidas e/ou acalmar meus ânimos. Ele me deixou livre para explorar a interdisciplinaridade do meu trabalho e me ajudou a trazê-la à filosofia, ainda que discordasse de alguns dos meus posicionamentos filosóficos. Agradeço também à quem considero ser meu segundo orientador, o professor João Kogler (POLI-USP) que desde o início do meu Mestrado tem dedicado seu tempo a me ensinar a importância da clareza e da precisão não somente na matemática que meu trabalho envolve, mas também nas próprias palavras que uso para descrevê-la.

Parcelas do presente trabalho foram realizadas com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), processo nº 2021/03688-0. As opiniões, hipóteses e conclusões ou recomendações expressas neste material são de responsabilidade do(s) autor(es) e não necessariamente refletem a visão da FAPESP.

**“Long you live and high you fly
Smiles you'll give and tears you'll cry
All you touch and all you see
Is all your life will ever be.”**

Pink Floyd, Breathe

RESUMO

IENNACO, M.L., A inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo para o estudo da consciência. 2023. 111 f. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2023.

A consciência sempre foi uma das grandes incógnitas do conhecimento humano. Por se tratar de um fenômeno extensamente investigado, sua conceitualização tornou-se multifacetada e repleta de contradições e limitações. Um dos motivos pelos quais a consciência tem sido um tópico problemático de estudo consiste no uso desmedido de intuições acerca de sua função e existência, as quais parecem ter popularizado correntes filosóficas e debates que eternizaram a ideia de uma lacuna explicativa entre a experiência subjetiva e seus processos cognitivos subjacentes, dificultando (senão impossibilitando) avanços científicos significativos nos estudos do tema. Com isso em mente, minha proposta para solucionar esse impasse envolve uma mudança de pressupostos teóricos, em prol daqueles que tratam as intuições com a devida cautela, evidenciam as vantagens de estudos deflacionários e da utilização de ferramentas interdisciplinares na elaboração de uma proposta descritiva de sua natureza e funcionamento. Assim, o principal objetivo da presente pesquisa é estabelecer os parâmetros e contornos gerais de uma descrição possível da consciência amparada pela investigação da inter-relação entre um modelo teórico chamado Modelo dos Múltiplos Esboços e o *framework* cibernético do Processamento Preditivo. Para tanto, uma vez esclarecido aquilo que entendo por consciência, desenvolvo uma descrição sobre seu funcionamento internamente a ambos, modelo e *framework*, apontando, nesse processo, alguns de seus problemas e limitações. Posteriormente, aponto algumas das semelhanças existentes entre eles e como sua utilização conjunta poderia fornecer resoluções convincentes para seus problemas e limitações individuais. Concluo alegando que a inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo fortalece enormemente nossas capacidades de descrição satisfatória e plausível dos processos e mecanismos envolvidos no fenômeno da consciência, além de viabilizar seu trato empírico.

Palavras-chave: Processamento Preditivo, Consciência, Modelo dos Múltiplos Esboços, Ilusionismo, Neurociências Cognitivas.

ABSTRACT

IENNACO, M.L., The Interrelation between the Multiple Drafts Model and Predictive Processing for Consciousness Studies. 2023. 111 f. Dissertation (Master's Degree) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia, Universidade de São Paulo, São Paulo, 2023.

Consciousness has always been one of the great unknowns of human knowledge. Because it is an extensively investigated phenomenon, its conceptualization has become multifaceted, full of contradictions and limitations. One of the reasons why consciousness has become a problematic topic of study is the excessive use of intuitions about its function and existence, a fact that seems to have popularised philosophical currents and debates that perpetuated the idea of an explanatory gap between our subjective experiences and their underlying cognitive processes, making it difficult (if not impossible) for significant scientific advances on the topic. With that in mind, my proposal of resolution to this stalemate involves a change in our theoretical assumptions, favouring those that treat intuitions with due caution, highlight the advantages of deflationary studies and the use of interdisciplinary tools in the elaboration of a descriptive proposal of the nature and functioning of such phenomenon. Thus, the main objective of this research is to establish the parameters and general contours of a possible description of consciousness supported by the investigation of the interrelationship between a theoretical model called the Multiple Drafts Model and the Predictive Processing cybernetic framework. To this end, once the concept of conscience is clarified, I develop a description of its functioning internally to both the model and the framework, pointing out, in the process, some of its problems and limitations. Later, I point out some of the similarities between them and how their joint usage could provide convincing solutions to their individual problems and limitations. I conclude by claiming that the interrelationship between the Multiple Drafts Model and Predictive Processing enormously strengthens our capacities for a satisfactory and plausible description of the processes and mechanisms involved in the phenomenon of consciousness, in addition to enabling its empirical treatment.

Key words: Predictive Processing, Consciousness, Multiple Drafts Model, Illusionism, Cognitive Neuroscience.

SUMÁRIO

Introdução	7
1. A Consciência	15
1.1 Alguns produtos da intuição humana	16
1.2 Os Problemas da consciência	20
1.2.1 Os Problemas Fáceis e o Problema Difícil	20
1.2.2 O Metaproblema	24
1.2.3 Outros problemas da consciência	27
1.3 A fuga do Teatro Cartesiano	35
1.4 O Modelo dos Múltiplos Esboços (MME)	41
2. O Processamento Preditivo (PP)	46
2.1 Fundamentos Gerais	46
2.1.1 Contextualização histórico-filosófica	46
2.1.2 O Cérebro Preditivo	52
2.1.3 Detalhamentos da Minimização de Erro de Predição	56
2.1.4 Ação	59
2.2 A consciência preditiva	61
2.2.1 Distinguindo-se do meio	62
2.2.2 Limitações do Processamento Preditivo como uma teoria empírica da consciência	64
3. A inter-relação entre o MME e o PP para o estudo da consciência	70
3.1 Algumas semelhanças	71
3.2 As múltiplas sondagens	73
3.3 Como inferências inconscientes alcançam a consciência?	77
3.4 A unicidade da percepção consciente	82
3.5 A imagem manifesta preditiva e orientada às affordances	87
4. Considerações finais	94
5. Referências	98

Introdução

Durante toda a nossa vida somos constantemente surpreendidos pela complexidade e beleza que existe no mundo. Nossas experiências, ainda que majoritariamente rotineiras, são singulares e se somam em uma longa história biográfica, a qual somos capazes de lembrar, modificar e relatar desde um ponto de vista que possuímos acerca de nós mesmos e das nossas vivências nesse mundo. A expansão e o desenvolvimento da sociedade humana forneceram meios que facilitaram a exploração daquilo que supostamente possibilitaria essa perspectiva individualista, algo que muitas vezes foi referenciado como *psique*, alma, mente, vida mental, consciência, *self*, etc. Naturalmente, cada um dos termos eventualmente associados a essa perspectiva e àquilo que ela possibilitaria acompanha implicações distintas para o seu uso, mas a essência de que haveria alguém vivenciando um mundo, desde um ponto de vista privilegiado (e singular), é comum a todas elas. Por esses e outros motivos, a exploração desse tema sempre foi muito pluralizada, desafiadora e academicamente fragmentada. No decorrer do século XX, a disciplina da Filosofia acabou por tomar frente dessa exploração e – naquilo que tange às discussões epistemológicas e ontológicas disso que se conceio tratar como “consciência humana” – foi a responsável por impulsionar o desenvolvimento de teorias extremamente influentes e aparentemente robustas que, em sua maioria, partem de argumentos fundamentados na intuição.¹ Foi apenas em meados da década de 1990 que as ciências naturais passaram a, de fato, assumir uma parcela considerável dos estudos sobre o tema, contribuindo com *insights* significativos e até mesmo revolucionárias sobre o entendimento da consciência e seu impacto no que corriqueiramente costumamos caracterizar como nossa vida mental.² No entanto, ainda não se pode dizer que os vários problemas que, ao longo de seu desenvolvimento, vieram a caracterizar essa exploração científica foram resolvidos (ou sequer amenizados), pois, mesmo com seu evidente progresso recente, o campo

¹ Trata-se de uma narrativa comum na história dos estudos sobre o tema: com o behaviorismo, as ciências psicológicas teriam desistido, em sua maioria, de enfatizar os processos implícitos da mente humana – tais como a consciência – optando por se direcionar à análise do comportamento diretamente observável. Com isso, o tema teria sido desconsiderado ou, simplesmente, perdido seu apelo entre cientistas, acabando por ser majoritariamente relegado ao trabalho dos filósofos (CF. GARDNER. 2003).

² Em grande parte, costuma-se atribuir essa reviravolta à suposta influência científica dos questionamentos elaborados pelo filósofo David Chalmers (1995) – o qual, conforme veremos, foi responsável pela introdução da famosa distinção entre os Problemas Fáceis e o Problema Difícil da consciência. Ao que tudo indica, naquilo que diz respeito à influência de Chalmers nos estudos científico-naturalistas acerca da consciência, talvez possamos observar a existência de um fenômeno acadêmico semelhante ao que se costuma denominar “efeito guru” (Cf. LAU, MICHEL, 2020).

ainda sofre com múltiplas contradições, imprecisões e desafios substanciais (Cf. HARLEY, 2021).

Com esse cenário em mente, a motivação inicial desta pesquisa é contribuir para a viabilização de uma possível descrição do funcionamento da consciência, que seja frutífera para o trabalho e interesses tanto de filósofos quanto de cientistas cognitivos³. Isto é, uma descrição relativamente precisa do fenômeno e suas implicações que, ao mesmo tempo, seja teoricamente rigorosa e se empenhe pela testabilidade e falseabilidade empíricas. No entanto, antes de apresentar a linha de raciocínio a ser desenvolvida, e no intuito de minimizar a reintrodução de possíveis imprecisões e ambiguidades comumente presentes nos estudos sobre o tema, além de delinear um escopo mais preciso para a presente pesquisa, penso que seria fundamental o esclarecimento de alguns dos pressupostos que tomarei como centrais. Mais especificamente: (1) as diferentes imagens de mundo; (2) a tese ilusionista e (3) a ideia de que a percepção consciente seria orientada às *affordances*⁴. Conforme os concebo, esses pressupostos implicam o comprometimento dessa dissertação com uma perspectiva fisicalista (que explica o mundo integralmente em termos físicos) e funcionalista (que postula a necessidade de determinados fenômenos observáveis possuírem uma função), porém fundamentada em critérios instrumentalistas (tratando os processos, mecanismos e estruturas aqui descritos como heurísticas para o trabalho científico-filosófico, sem se comprometer, então, com seu suposto *status* ontológico). Adicionalmente, é válido esclarecer que optarei por partir de uma descrição deflacionária tanto da percepção consciente quanto do próprio fenômeno da consciência, a qual não-necessariamente elimina as propriedades mais comumente atribuídas a ambos, mas enfatiza algumas dessas propriedades, as quais considero serem mais centrais e empiricamente tratáveis, em detrimento daquelas que ainda parecem carecer de desenvolvimentos científicos e filosóficos ulteriores para serem devidamente estudadas – negando, portanto, sua centralidade ou, pelo menos, sua indispensabilidade para o estudo rigoroso do tema (DOŁĘGA, DEWHURST, 2021).

³ Alcinha genérica para um conjunto de esforços interdisciplinares visando a compreensão da mente e sua relação com o cérebro, corpo e vivência humanos. Na presente dissertação tal alcinha é utilizada como análoga às denominações “ciências comportamentais e do cérebro”, “ciências psicológicas” e “ciências da mente”.

⁴ Conforme veremos, trata-se de um termo originalmente cunhado pelo psicólogo americano James Gibson (1904-1979) para se referir aos aspectos positivos e negativos, proporcionados pelo ambiente, que são relevantes para a interação e sobrevivência de um ser vivo em/com seu meio (Cf. GIBSON, 1979). Na dissertação, tal termo será, em grande parte, destituído de sua “natureza anti-representacionista” original para designar, meramente, as acessibilidades/possibilidades de interação disponíveis aos organismos em diferentes ambientes – ou seja, aquilo que o filósofo Maxwell Ramstead denomina “*Affordance 3.0*” (RAMSTEAD, 2022, p. 195).

Dito isso, o primeiro dentre os pressupostos centrais dessa dissertação é a existência das diferentes imagens de mundo, descritas pelo filósofo Wilfrid Sellars (1912-1989) na obra “*Science, Perception and Reality*” de 1963. De acordo com Sellars, uma forma de minimizar as complicações de muitos dos problemas e paradoxos do campo filosófico é melhorar o nosso entendimento das “perspectivas dos indivíduos no mundo” (1963, p. 5) separando-as em duas imagens. Para o autor, o termo “imagem” remete a dois significados importantes: (1) o primeiro consiste em um contraste entre um objeto em si e sua projeção; (2) o segundo se baseia na ideia de que uma imagem pode consistir em algo imaginado que, talvez, não exista fora desse contexto imaginativo. Para Sellars, então, isso implicaria uma existência ambígua para os dois principais conjuntos de perspectiva de mundo com os quais os filósofos deparam: a imagem manifesta e a imagem científica que, de certa forma, coexistem como reflexos da realidade e, ao mesmo tempo, são imaginadas. Isto é, ambas as imagens representam, em termos gerais, formas relativamente distintas de idealização, as quais ele acredita serem capazes de iluminar a maior parte das dinâmicas internas ao desenvolvimento das ideias filosóficas (SELLARS, 1963, p.5). De um lado, a imagem manifesta é caracterizada pelo “encontro de alguém com si mesmo” e consiste em um refinamento empírico e categórico daquilo que Sellars chamou de “a imagem original”⁵. Aqui, um processo de refinamento empírico operaria, de maneira abrangente e em termos de inferências estatísticas e correlatos indutivos, entre as próprias experiências que os indivíduos vivem em seu mundo e as crenças que possuem sobre essas vivências. Por sua vez, um refinamento categórico operaria sobre os processos de detalhamentos daquilo que é percebido e vivido no mundo, ou seja, através da percepção daquilo que os objetos fazem e são – o que levaria à sua progressiva categorização ao longo do desenvolvimento dos indivíduos (SELLARS, 1963, p. 11). De outro lado, a imagem científica é caracterizada como uma idealização otimizada da imagem manifesta, a qual procura integrar os vários *frameworks* científicos existentes – cada um munido de suas próprias (e distintas) teorias, níveis de análise, metodologias, instrumentos, etc – em uma única imagem que seria constantemente auxiliada pelo *feedback* fornecido por nossas vivências no próprio mundo manifesto. Logo, embora surja a partir da

⁵ Para Sellars, tudo na imagem original consistiria de “pessoas” que, com o tempo, hão de se “despersonalizar” ao serem nomeadas (categorizadas) como outros objetos. Por exemplo, para uma criança, uma árvore seria uma pessoa até ela aprender que uma árvore é uma árvore, diferentemente de uma pessoa (SELLARS, 1963, p.11). Trata-se de algo que parece convergir com o desenvolvimento e amadurecimento daquilo que se costuma denominar, nas ciências psicológicas, de “psicologia popular” ou “psicologia do senso comum” – isto é, a capacidade humana de intuitivamente explicar e prever o comportamento e estados mentais dos indivíduos ao nosso redor (Cf. RAVENSCROFT, 2016).

imagem manifesta, essa imagem científica estaria, de certa forma, além dela, graças à utilização de ferramentas que possibilitariam a mensuração e a observação de fenômenos “manifestamente” imperceptíveis (SELLARS, 1963, p. 21).

Com isso em mente, Sellars argumenta que (1) os objetos manifestos podem, em várias circunstâncias, ser tratados como idênticos ao conjunto de partículas (objetos científicos) que os compõem, da mesma forma que uma floresta é, de certa forma, idêntica a um conjunto de árvores; (2) os objetos manifestos são aquilo que, para nós, de fato existe, ao passo que os conjuntos de partículas são formas (científicas) abstratas de representá-los; e (3) os objetos manifestos são aparências, para a mente humana, de uma realidade (extra-humana) que é constituída por partículas manifestamente imperceptíveis (SELLARS, 1963, p. 26). Em outras palavras, as diferentes imagens de mundo coexistem em um *continuum*, ainda que a única forma de acesso à imagem científica seja através das teorias, métodos e instrumentos científicos. Logo, uma cadeira, por exemplo, é uma cadeira para nós, em nossa imagem manifesta, mas, ao mesmo tempo, ela é um conjunto de partículas se observada através das “lentes” da imagem científica. Aqui, não haveria nenhum momento decisivo no qual esse conjunto de partículas “viraria” uma cadeira. Ambos são as duas coisas ao mesmo tempo. O que muda é quem ou o que os observa. Isso implicaria que os diversos fenômenos que compõem nossa percepção e experiência no mundo poderiam ser tratados como algo físico e cientificamente tratável na imagem científica sem necessariamente possuírem a mesma tratabilidade na imagem manifesta. Além disso, independentemente do que for a suposta realidade (extra-humana) de partículas imperceptíveis descrita por Sellars, nossa realidade (humana) sempre será aquilo que vivenciamos na imagem manifesta. O conhecimento (científico) acerca dessas partículas é apenas uma forma, dentre as possíveis, de fazer sentido do mundo.

Tais observações acerca das imagens sellarianas levam-nos ao segundo pressuposto central desta dissertação: o Ilusionismo. Conforme veremos em mais detalhes no Capítulo 1, embora seja em grande parte fundamentado nos estudos do filósofo Daniel Dennett, o Ilusionismo é, em termos gerais, uma tese recente originalmente desenvolvida por Keith Frankish (2017), a qual notoriamente postula que o fenômeno da consciência – mais especificamente a consciência fenomênica ou as *qualia* – não seria, de fato, aquilo que nos aparenta ser. Para Frankish, as teorias “ingenuamente” realistas acerca da consciência requereriam, para seu sucesso empírico, a necessidade de grandes revoluções científicas (uma demanda extremamente custosa e, talvez, até mesmo inviável), ao passo que suas contrapartes

idealistas não seriam capazes de contribuir adequadamente para a nossa compreensão científica da consciência (uma vez que propõem que a única forma de investigação legítima de tal fenômeno seria através da introspecção individual). Assim, o autor alega que, por meio de uma perspectiva ilusionista, os filósofos se veriam aptos a avançar para além dessa dicotomia entre o realismo e o idealismo que parece dominar o estudo contemporâneo do tema, abandonando programas de pesquisa que busquem encontrar, da maneira como ela nos aparenta – seja por meio de uma nova ciência ou por intermédio da introspecção –, a suposta qualitatividade da consciência no cérebro, voltando-se, por sua vez, à busca pela compreensão dos próprios motivos pelos quais achamos que possuímos uma vida mental rica, vívida e repleta de “qualidades misteriosas” (FRANKISH, 2017, p. 17). Ao longo do presente texto, o Ilusionismo será utilizado apenas como um elemento norteador, sendo desenvolvido como uma possível consequência da devida integração dos *insights* sellarianos ao estudo da consciência – isto é, a partir da imagem científica, a realidade manifesta de nossa consciência poderia ser tratada como ilusória, pois, nessa imagem, ela seria melhor concebida como “constituída por [processos, mecanismos e estruturas] imperceptíveis” (SELLARS, 1963, p. 26) a nós.⁶ Digno de nota, a adoção desse pressuposto ilusionista não implica qualquer mudança em nossa vivência manifesta, mas, conforme será argumentado, pode fazer toda a diferença para as investigações teórico-empíricas acerca da consciência (Cf. DENNETT, 2016). O que ela implica, no entanto, é uma versão radical do Ilusionismo, uma vez que o pressuposto sellariano e os desenvolvimentos do Processamento Preditivo parecem concordar com a ideia de que toda a percepção (de propriedades espaço-temporais basilares às qualia) seria ilusória e existente apenas na imagem manifesta.

Finalmente, o terceiro pressuposto central consiste na compreensão de percepção consciente como um processo orientado às *affordances*. Esse conceito foi primeiramente criado e, então, popularizado por James Gibson em 1979, consistindo, em termos gerais, naquilo que o meio pode interativamente oferecer ao animal (e vice-versa) – indicando a existência de uma complementaridade cognitiva necessária entre os organismos e seus ambientes (1979, p. 197). Para Gibson, então, a percepção seria um processo ativo e voltado à detecção de *affordances*, em vez de algo passivo e voltado à detecção e categorização das qualidades ou propriedades reais de nosso mundo. Por exemplo, algo como uma cadeira não

⁶ Aqui, não é o nosso objetivo assumir todas as nuances do Ilusionismo, conforme uma proposta teórica específica para o estudo e compreensão da natureza/função das *qualia*, mas apenas nos amparar em algumas de suas premissas e, com isso, indicar que as *qualia* e a totalidade da nossa vida mental, como parte de nossa realidade manifesta, poderiam ser, conseqüentemente, tratadas como ilusórias (desde a imagem científica).

nos seria primeiramente percebido como, por exemplo, um objeto de madeira, marrom e velho, mas como algo sentável. Essas *affordances* seriam, então, relações (ou, segundo alguns teóricos, disposições) únicas para cada organismo, dadas as suas capacidades e seu histórico de interações com o meio. Por exemplo, a minha percepção de *affordance* da “escalabilidade” de um prédio é diferente da percepção de um praticante de parkour. O motivo pelo qual esse pressuposto é importante aqui é por seu potencial de complementar aquilo já insinuado pelos outros dois: a realidade manifesta de nossa consciência não apenas poderia ser caracterizada como ilusória, mas ela (ao menos naquilo que tange à nossa percepção) também o seria por ser orientada às *affordances* – e não às qualidades ou propriedades reais de nosso mundo. Dessa forma, esse *insight* gibsoniano indica-nos um possível motivo para a existência e configuração da percepção consciente tal como a vivenciamos: possibilitar uma melhor interação com nosso meio. Afinal, se considerarmos a complexidade dos contextos sociais nos quais vivemos, precisamos de um aparato mental com capacidades igualmente complexas, sofisticadas e otimizadas para fazermos sentido daquilo que nos é importante nesses ambientes e, assim, (sobre)viver neles (Cf. GODFREY-SMITH, 2002).

Digno de nota, para além da exposição desses três pressupostos centrais, faz-se necessário também, uma breve nota acerca do uso de algumas terminologias e conceitos. Por exemplo, ao longo do texto os termos “crença” e “representação” serão utilizados segundo um entendimento relativamente “liberal”, a despeito de sua notória pluralidade semântica e das querelas existentes em torno de sua definição, especialmente na filosofia das ciências cognitivas e psicologia. A título de ilustração, na Filosofia, o uso não-pormenorizado do termo “representação” implica a adoção de um posicionamento representacionista para o entendimento da percepção e consciência. Entretanto, conforme veremos no Capítulo 2 (e em diante), sua adoção por essa dissertação não significa que ela está comprometida com o cognitivismo e/ou intelectualismo acerca da cognição, mas com um posicionamento que, apesar de considerar a atuação do cérebro como central, busca sua integração com o corpo e o meio. De fato, uma vez que (também conforme veremos) muito daquilo que será exposto na presente dissertação tem como suporte um material científico bastante técnico e, originalmente, estruturado em termos matemáticos, penso que – mediante às devidas explicações, esclarecimentos e referências – torna-se necessário, além de aceitável, adaptar seu vocabulário técnico para uma linguagem mais acessível e comum à literatura filosófica. Portanto, ao longo deste texto, a principal motivação para o uso “liberal” desses e outros termos e conceitos, os quais são filosoficamente “carregados de teoria”, é seu potencial

heurístico, facilitando a introdução e compreensão de um argumento que, de outra maneira, poderia se tornar desmedidamente técnico – uma vez que, em seu “terreno habitual”, muito daquilo que será exposto é normalmente entendido em termos de relações estatísticas (no caso, por exemplo, de “crenças”) e dinâmicas neuronais (ao tratar “representações” no contexto de padrões de ativação cerebral).

Dito isso, o principal objetivo da presente pesquisa é investigar a possível inter-relação (isto é, a possibilidade de relação e esclarecimento mútuos) entre o *framework* do Processamento Preditivo (PP) e o Modelo dos Múltiplos Esboços (MME) da consciência numa possível descrição da consciência a partir da imagem científica. Com origem na cibernética, o PP é um *framework* neurocomputacional que busca abarcar a totalidade dos mecanismos estruturais e processos hierarquicamente organizados do cérebro, tendo sido utilizado para a descrição dos processos perceptivos, atencionais e comportamentais dos organismos de maneiras (formalmente) satisfatórias e biologicamente implementáveis (Cf. HOHWY, 2013, CLARK, 2016). Por sua vez, o MME foi criado pelo filósofo Daniel Dennett em 1991 para descrever, de uma maneira teoricamente robusta e empiricamente testável, a consciência como o resultado da influência de múltiplos esboços paralelos de informação, ocorrendo tanto localmente quanto globalmente em nosso sistema nervoso.⁷ A minha hipótese é que a relação entre esse modelo e esse *framework* pode resultar em uma resolução de alguns dos problemas que eles enfrentam e, por conseguinte, viabilizar uma descrição satisfatória e precisa da consciência.

No intento de cumprir tal objetivo, o argumento apresentado por essa dissertação será segmentado em três capítulos. O primeiro capítulo é sobre a consciência. Nele, começo por argumentar como o uso desmedido das intuições filosóficas parece estar no cerne dos chamados “problemas da consciência” e, em seguida, proponho uma maneira de nos desviar desses problemas, rumo a uma descrição tanto filosoficamente plausível quanto cientificamente tratável da consciência. Feito isso, o capítulo termina fornecendo alguns detalhes adicionais acerca de como o modelo dennettiano poderia elucidar, internamente a essa descrição, o funcionamento da consciência. O segundo capítulo se volta ao *framework* do Processamento Preditivo. Nele, começo expondo as origens histórico-filosóficas do PP,

⁷ No estudo da consciência, a vantagem da adoção de modelos está na potencial capacidade de eles nos auxiliar na identificação de correlatos ou, simplesmente, respostas neuronais (para além das suas dinâmicas heterogêneas), facilitando assim nossa compreensão, não somente de seu funcionamento habitual, mas também de distúrbios e transtornos em tal funcionamento – o que poderia possibilitar, inclusive, a eventual viabilização de intervenções (Cf. SMITH, BADCOCK FRISTON, 2020, p. 11).

seguido dos detalhamentos dos principais processos e mecanismos propostos por esse *framework*. Depois, exploro algumas das evidências a seu favor e de suas aplicações, de modo a investigar como ele poderia descrever a consciência e, nesse meio termo, exponho suas principais limitações no que tange ao fornecimento de tal descrição. Finalmente, o terceiro e último capítulo trata a inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo. Nele, começo por demonstrar algumas semelhanças existentes entre ambas as propostas, seguido de uma adição importante ao modelo dennettiano que facilita a compreensão de alguns pontos que ficaram ambíguos na descrição elaborada pelo autor. Posteriormente, valendo-me de tal desenvolvimento, proponho uma resolução ilusionista para alguns dos problemas que o Processamento Preditivo possui em sua descrição da consciência e, por fim, elaboro uma hipótese acerca de como aquilo desenvolvido ao longo do capítulo poderia esclarecer a maneira como a fenomenalidade da percepção seria estruturada – tendo os pressupostos iniciais da dissertação sempre em consideração. A dissertação se encerra com algumas breves reflexões e comentários finais acerca de tal empreitada.

1. A Consciência

Embora seja conveniente definir a consciência como “*our experience of being us, here and now. It’s what is lost when we are in a dreamless sleep and what returns next morning when we wake up*”⁸ (Cf. TONONI, 2004; TONONI, ELDEMAN, 1998), essa definição, para António Damásio e Kaspar Meyer, é vazia e imprecisa. Da mesma forma, definir a consciência apenas como o estado de vigília (*awareness*) não nos diz muito, uma vez que a vigília é, por sua vez, frequentemente utilizada como sinônimo de consciência (DAMÁSIO, MEYER, 2009, p. 4). Ainda, o filósofo Thomas Nagel (1974) defende uma definição de consciência muito aclamada na filosofia da mente tradicional: “*an organism is conscious if there is something it is like to be this organism*”. Tal definição fomentou, inclusive, o desenvolvimento e popularização de perspectivas que enfatizam a suposta primazia dos aspectos qualitativos da consciência – mais tecnicamente denominados *qualia* (ou *quale*, no singular). Segundo Frankish (2012), uma *quale* pode ser caracterizada como a sensação subjetiva de experienciar algo. Sempre que há esse como-é-ser, há fenomenalidade. Uma segunda forma de caracterizar uma *quale* é dizer que ela é uma propriedade dos dados sensoriais – isto é, se o sujeito vê seu café, ele terá uma representação visualmente fidedigna do café (em relação ao objeto no mundo externo), além de suas propriedades intrínsecas (ou seja, íntimas e não-representacionais), as quais caracterizariam o como-é-ser que permitiria a experiência visual do café. Ainda, na maioria dos casos, as *qualia* são tratadas como inefáveis (ou seja, não podem ser descritas ou nomeadas); cruas (por não precisarem do raciocínio cognitivo para sua detecção) e diretamente apreendidas na percepção consciente (Cf. TYE, 2021; FRANKISH, 2012).⁹

Dito isso, é inegável que, ao longo do tempo, o termo “consciência” tenha ganhado diversos significados, de modo que, *a priori*, seja apenas possível defini-lo como um termo guarda-chuva. Isso parece ser um grande problema para a confecção de estudos precisos e rigorosos, pois como não há uma definição ou mesmo critérios delineadores claros acerca de sua ontologia, cada autor pode considerar um conjunto qualquer de propriedades, processos, mecanismos, etc, (determinados da maneira que lhes convier) como consciência, contribuído,

⁸ “nossa experiência de sermos nós, aqui e agora. É o que se perde quando estamos em um sono sem sonhos e o que retorna na manhã seguinte quando acordamos” [tradução nossa].

⁹ Mais especificamente, essas características compõem aquilo que Frankish chamou de “*qualia clássicas*”. Atualmente, as *qualia clássicas* não são tão populares quanto as “*diet qualia*”, as quais seriam equivalentes a uma definição mais ampla de *quale* (como-é-ser) e estariam presentes na maioria dos estudos filosóficos contemporâneos acerca do tema (FRANKISH, 2012, p. 669).

por conseguinte, para a crescente ambiguidade e confusão conceitual que impera sobre os estudos do tema (Cf. BLACKMORE, TROSCIANKO, 2018). Não obstante, não é o objetivo desta dissertação promover o consenso acerca de uma descrição robusta do que é a consciência e seus mecanismos, mas explorar um caminho que talvez viabilize essa descrição, evitando alguns dos habituais problemas que nos distanciam de uma melhor compreensão do fenômeno. Para tanto, o presente capítulo argumenta que alguns dentre os produtos de nossa intuição acabariam por ser os responsáveis pela maior parte dos grandes problemas filosóficos acerca da consciência e que sua atuação parece mais atrapalhar do que auxiliar no entendimento do fenômeno. Como solução, proponho não somente que devemos aceitar e partir dos pressupostos introduzidos na seção anterior, mas também considerar atentamente algumas premissas filosóficas menos populares, as quais podem nos auxiliar a viabilizar uma descrição tanto cientificamente bem-embasada quanto filosoficamente coerente e contundente.

1.1 Alguns produtos da intuição humana

Na filosofia, a intuição é comumente descrita como uma crença ou disposição para crer em algo (Cf. De CRUZ, 2014). Para o filósofo David Chalmers, as intuições seriam confeccionadas através de uma combinação entre fatores quase-universais – isto é, estados mentais e mecanismos introspectivos pan-humanos – e variáveis contextuais – tais como cultura, linguagem, religião, carga teórica, etc. (2018, p. 15). Em termos práticos, ela seria a capacidade que possuímos de sentir algo a respeito de um evento ou estado sem necessariamente possuímos alguma explicação para tal evento ou estado. E, mesmo que nos pareçam realistas, óbvias e difíceis de serem negadas, para o filósofo David Lewis (1983), as intuições seriam simplesmente opiniões, algumas delas baseadas no senso comum, outras sofisticadas por nossas interações sociais, technicalidades filosóficas, etc – sem nunca deixarem de ser, meramente, opiniões.

A intuição foi e ainda é uma ferramenta importante para os avanços intelectuais, sendo, desde muito cedo, na história do pensamento ocidental. Assim, graças ao seu uso conjunto com as introspecções (na busca por “olhar para si mesmo”), as intuições permitiram que, por séculos, a ontologia e a epistemologia da atividade mental fossem amplamente estudadas. De fato, muito daquilo acerca da mente que parecia ser realisticamente óbvio para os intelectuais de outrora foi, com o avanço da tecnologia e metodologia científicas, ganhando certo referencial empírico, mas o uso “empiricamente desacompanhado” de intuições ainda

continua a exercer influência crucial no campo de estudos da consciência. Na presente seção, dois dentre os mais influentes produtos da intuição humana serão explorados em maiores detalhes: o dualismo mente-corpo e o essencialismo.

As raízes do postulado da separação entre mente e corpo podem ser localizadas desde pelo menos a Grécia Antiga, especialmente na obra de Platão, a qual postulava que as verdadeiras substâncias ou “formas” seriam eternas e não-físicas e das quais os efêmeros corpos físicos seriam apenas cópias imperfeitas – ao passo em que, para Aristóteles, as formas eram a natureza das coisas e, portanto, a alma seria a própria forma do corpo humano (Cf. ROBINSON, 2020). No entanto, foi a obra do filósofo René Descartes a maior responsável pela consolidação, no pensamento filosófico ocidental, da perspectiva moderna de um dualismo mente-corpo. Mais especificamente, o dualismo cartesiano ou “dualismo de substâncias” defende que a mente é não-física e distinta do corpo físico, além de existir independentemente do corpo – algo que o torna compatível com diversas teologias (CARRUTHERS, 2020a, p. 199). Graças ao impacto desse paradigma dualista, muitas das descobertas teóricas e empíricas do século XVII, enfatizando o método científico nos estudos sobre o corpo, acabaram por, simplesmente, evitar falar sobre a mente em suas análises (Cf. ROBINSON, 2020). Tal cenário perdurou por pelo menos três séculos e, mesmo com diversos problemas e inconsistências filosóficas (Cf. *ibid.*), o dualismo cartesiano permanece uma presença impactante no pensamento científico e filosófico ocidentais, permanecendo inclusive no senso comum (Cf. COHEN et al, 2011) – o que parece ter influenciado a conclusão de que tal posicionamento seria universal, por surgir desde muito cedo em nosso desenvolvimento (Cf. BLOOM, 2004).

O dualismo e suas ramificações parecem fomentar uma espécie de conflito entre supostos estados mentais e corporais, ao postular a independência ou irredutibilidade da mente em relação ao corpo e com isso, possibilitar, por exemplo, a sedutora existência de zumbis filosóficos (LEAL-TOLEDO, 2006, p. 129). Mas, se os nossos corpos são físicos, por que a agência e as demais atividades mentais instanciadas por ele não o seriam também? De acordo com o filósofo Peter Carruthers (2020a, p. 201-203), há uma desconexão entre a física e a psicologia popular¹⁰, pois a última não exige que seus estados ocupem qualquer espaço – uma característica crucial de nosso universo físico. Dado tal suporte, as intuições de que estados

¹⁰ De certa forma, a *folk psychology* ou “psicologia popular” consiste na capacidade humana de explicar e prever o comportamento e estados mentais de outras pessoas, valendo-se de termos linguísticos comuns, em oposição a jargões técnicos ou científicos (Cf. RAVENSCROFT, 2021).

mentais podem não ser redutíveis ao físico rapidamente se espalhariam entre as culturas que as articulam, de modo a atuar como um “efeito atrator” que fomenta esse impasse em nosso conhecimento básico do mundo e ambiguidades em nossa ciência e filosofia. A facilidade da aquisição e existência desse dualismo tácito (Cf. LEAL-TOLEDO, De VASCONCELOS, 2022), até mesmo quando nos esforçamos pelo desenvolvimento de perspectivas estritamente fisicalistas, é algo que não só parece enviesar o pensamento e metodologia científicos, graças ao surgimento e (aparente) inevitabilidade de dicotomias excessivas, mas limitar as possibilidades de pesquisas melhor alinhadas com nossa realidade empírica (CARRUTHERS, 2020a, p. 203).¹¹ Ainda, as próprias intuições dualistas tendem a nos levar a linhas de raciocínio desnecessariamente intelectualistas, visto que, (I) se há uma característica essencial e comum a todos os seres humanos (mente), que nos diferencia dos outros animais; (II) e as crenças cartesianas sobre essa diferença se baseiam na existência de nossa vida mental (ou alma) como distinta de nossos corpos; (··) logo, parece ser possível concluir que a mente humana seria essencial (ou parte da essência de nossa existência), em detrimento de nossa estrutura física. E isso nos leva ao segundo produto da intuição que será abordado aqui: o essencialismo.

O termo “essência” costuma se referir a um tipo de núcleo interno a cada coisa, o qual causa uma manifestação das propriedades físicas específicas a elas, além de suas possíveis propriedades mentais e/ou intencionais (tais como desejos, preferências, etc.) (CARRUTHERS, 2020b, p. 231). Na psicologia, o essencialismo consiste em uma intuição de que todos os seres vivos possuem algo como uma essência imutável e inata, compartilhada e mantida de maneira semelhante por cada indivíduo de uma mesma espécie. Ademais, semelhanças entre as espécies poderiam ser determinadas por uma “natureza subjacente”, a qual não poderia ser observada diretamente (grosso modo, parece haver, em última instância, alguma propriedade do gato, por exemplo, que o faz ser diferente de um leão e de uma onça, mas, ao mesmo tempo, mais semelhante aos dois do que a uma cobra – ou seja, sua essência) (WHITE, 2021, p. 117). O essencialismo é uma intuição que também se desenvolve precocemente durante a infância humana (Cf. GELMAN, WELLMAN, 1991; GELMAN,

¹¹ Caso o leitor suspeite que as imagens de mundo propostas por Sellars seriam mais uma dessas dicotomias ou dualismos aqui criticados, ressalto que não são. Dicotomias e o dualismo de substâncias sugerem a separação de duas partes de forma exaustiva (tudo ou nada nas partes) e exclusiva (as partes não podem ter uma mesma propriedade). As imagens sellarianas, por outro lado, coexistem, ou seja, tratam de uma mesma coisa que é observada a partir de dois ângulos diferentes: a observação a “olho nu” e a observação por meio de ferramentas científicas.

2003), sendo, portanto amplamente difundido e bem-consolidado, tanto em nossas crenças quanto em nosso senso comum.

Isso exposto, seguindo a linha de raciocínio de que corpo e mente são, de alguma forma, distintos e que a mente compõe ou é, ela mesma, a essência do ser e existência humanos, caminhamos automaticamente para uma distinção final que aparenta completar esse cenário: o “eu” em contraposição ao restante daqueles ao nosso redor. É de supramportância para um organismo se diferenciar do seu contexto de modo que ele saiba, por exemplo, o que comer e o que não comer, onde ele pode entrar, qual limite de temperatura ele suporta, etc. Quando nos referimos à complexidade mental humana, talvez a primeira coisa a ser destacada é o universo mental e único de cada indivíduo, o qual não pode ser acessado por ninguém além dele mesmo. Afinal, é esse universo que faz dele, ele. Mesmo que ainda sob dúvidas, é natural pensar que, se possuímos um universo privado e único e nos comportamos e nos expressamos de certas maneiras, nossos vizinhos que se comportam e se expressam de modo semelhante talvez tenham, também, seus próprios universos únicos e privados. Ao explorarmos detalhadamente esse “reino” único e privado, imediatamente constatamos que somos intérpretes fidedignos de nós mesmos e que podemos confiar plenamente em nossa percepção (supostamente detalhada) de nosso mundo interno e externo, nossos pensamentos e comportamentos – como se não houvesse outra forma de ser a não ser essa que vivenciamos. Logo, quando realizamos um relato ou refletimos sobre nossas próprias crenças, juízos, intuições e intenções, quase sempre estaremos convictos de que estamos corretos acerca das mesmas (afinal, “é óbvio que eu sei o que se passa na minha cabeça!”). Com isso, não apenas nos é natural e intuitivo dizer coisas como que “sei que sou/estou consciente” e que “sei o que está acontecendo em meu fluxo de consciência nesse exato momento”, como nos é simplesmente impossível conceber que as coisas talvez não sejam assim (CARRUTHERS, 2020a, p. 207).

Curiosamente, embora em um primeiro momento não seja possível saber como é o universo de vivências do outro, nós possuímos uma tendência de tentar inferir ou prever o que se passa por lá. Pois, uma vez que somos convictos de que sabemos praticamente tudo o que há a de ser conhecido sobre a nossa vida mental, por que não tentar destrinchar (também) a vida mental dos outros? Tanto na psicologia quanto na filosofia isso é conhecido como teoria da mente e consiste na capacidade que possuímos de atribuir (às outras pessoas e/ou a nós mesmo) estados mentais – uma habilidade fundamental para o sucesso das interações humanas (CARRUTHERS, 2020a, p. 195-196). Essas intuições acerca das vidas mentais de nossos

semelhantes e de nós mesmos influenciam fortemente nossas vivências individuais e, conseqüentemente, a forma como interagimos com o mundo. Inclusive, há boas razões para se esperar que intuições tais quais as oriundas de nossa teoria da mente possam inadvertidamente influenciar até mesmo o pensamento científico e reflexões filosóficas acerca da própria mente humana (CARRUTHERS, 2020a, p. 214; Cf. LAU, MICHEL, 2020).

É inegável que o uso indiscriminado dessas e tantas outras intuições tenham, em vários momentos da história das ideias, fomentado o avanço do conhecimento e dos rumos por ele tomados. Mas, ao que nos parece, uma forma mais prudente e sensata de lidar com as intuições seria sempre termos em pauta a investigação sobre os motivos pelos quais elas existem, o que nos leva a tê-las, quais as suas vantagens, desvantagens, etc, para que, com isso, quando algo nos parecer intuitivo ou óbvio, sempre saibamos como questionar tal impressão, testá-la ou falseá-la. Assim o uso, tanto científico quanto filosófico, de intuições ganharia uma linha robusta de raciocínio subjacente, a qual nos permitiria (justamente) saber o porquê de algo eventualmente nos parecer óbvio. Penso que elas, então, devam ser mais propriamente consideradas como ferramentas para auxiliar o raciocínio teórico e empírico, em vez de um fundamento suficiente para programas inteiros de pesquisa. A valer, esse talvez não seja um posicionamento fácil de ser empregado – especialmente naquilo que diz respeito à reflexão ou pressuposição de *insights* altamente contraintuitivos sobre nós mesmos (LAU, MICHEL, 2020, p. 2-3; Cf. DENNETT, 2016) – mas, não obstante, trata-se de uma prática necessária, pois, conforme veremos a seguir, ao não realizarmos tais questionamentos, nos tornamos “reféns” dessas intuições, sendo levados a questões potencialmente insolúveis e ambigüidades e limitações possivelmente intransponíveis.

1.2 Os Problemas da consciência

1.2.1 Os Problemas Fáceis e o Problema Difícil

A influência dos vieses intuitivos humanos na filosofia parece ter consolidado um padrão dualista de pensamento que obteve sua mais recente e relevante contribuição intelectual na distinção entre os Problemas Fáceis e o Problema Difícil da consciência, primeiro concebida pelo filósofo David Chalmers em 1995. Segundo o autor, os Problemas Fáceis são aqueles que, atualmente ou futuramente, poderiam ser explicados pelas ciências cognitivas através de mecanismos e processos computacionais e neuronais. Por outro lado, o

Problema Difícil implicaria que a ontologia da experiência fenomênica “transcende” as capacidades explicativas de tais aparatos científicos (Cf. CHALMERS, 1995).

Especificamente, os Problemas Fáceis abrangem uma série de campos de pesquisa nas ciências cognitivas que visam explicar, por meio da análise do funcionamento cerebral humano: a habilidade de distinguir, categorizar e reagir a um estímulo externo; como ocorre a integração da informação por um sistema cognitivo; os relatos de estados mentais;¹² a capacidade de um sistema acessar seus próprios estados internos; o foco da atenção; o controle deliberado do comportamento; e a diferença entre os estados de sono e vigília. Para Chalmers, todos esses fenômenos estariam relacionados com a consciência, e, mesmo que não tenhamos explicações completas e satisfatórias acerca de seu funcionamento, possuímos uma noção de como eles funcionam (CHALMERS, 1995, p. 2).

Entretanto, aquilo que Chalmers alega ser realmente difícil de explicar, é o problema da experiência. Tal problema foi anteriormente descrito pelo filósofo Thomas Nagel (1974) em seu ilustre texto “*What is it like to be a bat?*”, no qual ele alega que todo organismo consciente possui uma experiência única de ser como ele é. Por exemplo, a experiência de uma sensação visual, a experiência do escuro e do claro, o som de uma flauta, o cheiro do café, a emoção, etc, todas são unidas pela sensação única do que é ser um organismo, consistindo, portanto, em estados de experiência (CHALMERS, 1995, p. 3). Disso, o Problema Difícil questiona como e por que nossos processos cognitivos físicos gerariam tal experiência subjetiva, cujas formas através das quais tomamos conhecimento de sua existência são apenas o relato de primeira pessoa e a introspecção. Como ilustração, ele expõe que podemos entender e analisar empiricamente os mecanismos que nos permitem enxergar o vermelho do tomate (isto é, um Problema Fácil). Entretanto, a experiência consciente, ou seja, a sensação fenomênica de experimentar o vermelho do tomate, não estaria no escopo das ciências atuais e, mesmo que os processos computacionais pelos quais enxergamos tal vermelho sejam meticulosamente explicados, as qualidades subjetivas envolvidas ainda careceriam de elucidação.

Digno de nota, Chalmers equipara aquilo que ele chama de “experiência consciente” com as qualia e a “consciência fenomênica” e separa essa experiência da “consciência” em um sentido estrito, a qual, para ele, estaria mais propriamente relacionada aos nossos processos cognitivos comuns e ao estado de vigília. E, embora essa separação possa passar a impressão

¹² “[...] explicar como um sistema realiza a função de produzir relatos [por exemplo, verbais] acerca de estados internos” (CHALMERS, 1995, p. 203).

de uma melhor definição para os aspectos centrais à questão, ela acaba por gerar uma lacuna explicativa (Cf. LEVINE, 1983) entre os processos cognitivos examinados pelos Problemas Fáceis e a experiência subjetiva referida pelo Problema Difícil. Na verdade, a intuição acerca da dificuldade ou até mesmo impossibilidade de explicarmos como os processos cognitivos seriam acompanhados de experiências subjetivas é precisamente o que dá vida ao Problema Difícil (CHALMERS, 2020 p. 12), visto que, aqui, o pressuposto básico é o de que “quaisquer que sejam os mecanismos físicos imaginados, permanecem-nos ininteligível que eles poderiam explicar a consciência” (LAU, MICHEL, 2020, p. 2).

Chalmers reconhece a existência de tentativas de se explicar ou pelo menos estruturar uma explicação de um ou mais dos aspectos da consciência e sua fenomenalidade. Por exemplo, a nível neuronal, Francis Crick e Christof Koch (1990) elaboraram uma teoria neurobiológica da consciência que concebe os padrões de frequência do cérebro como a base da consciência. A nível psicológico, Bernard Baars (1998) teria contribuído com o *Global Workspace Theory*, a qual entende que a consciência seria transmitida a nível global no cérebro por um processador central, que elegeria informações inconscientes para essa transmissão. Além disso, Chalmers também cita outras contribuições relevantes, tais como o *Neural Darwinism* de Gerald Edelman (1989), o *Multiple Drafts Model* de Daniel Dennett (1991) e a *Intermediate Level Theory* de Ray Jackendoff (1987). No entanto, ele nega que quaisquer uma dessas teorias ou modelos tenha sequer tocado no Problema Difícil – isto é, elas foram relevantes para a elucidação de Problemas Fáceis, mas ainda não responderam o como e o porquê dos processos por elas descritos serem acompanhados de experiências subjetivas. Na verdade, Chalmers indica que aquilo que essas e outras teorias costumam fazer é tomar cinco caminhos diferentes: (1) explicar algo diferente de fenomenalidade, (2) negar a fenomenalidade, (3) dizer explicar a totalidade da fenomenalidade sem de fato fazê-lo, (4) explicar a estrutura da fenomenalidade e (5) isolar o substrato da fenomenalidade.

A solução que ele começa a propor para o estudo da fenomenalidade da consciência é aceitar a existência de algo extra, que não seria necessariamente físico, para complementar as explicações fisicalistas já existentes. De acordo com sua perspectiva, a qual ele denomina “dualismo naturalista”, a forma correta de se estudar a fenomenalidade é utilizar uma teoria não-reducionista capaz de agregar as “propriedades extras” da consciência, de modo a conseguir satisfatoriamente explicar o como e o porquê de processos cognitivos – relativamente bem entendidos – virem acompanhados dessas experiências (Cf. CHALMERS, 1995). Aqui, a inspiração para a consideração de um “ingrediente extra” vem da própria física,

onde alguns fenômenos ou entidades são pressupostos e fixados como fundamentais: por exemplo, processos tais como os eletromagnéticos não poderiam ser explicados integralmente por processos mecânicos, como é o padrão, então, segundo a narrativa fornecida por Chalmers, o físico James Maxwell teria introduzido, em sua explicação, uma carga eletromagnética e forças eletromagnéticas como componentes fundamentais da teoria física – isto é, para explicar o eletromagnetismo, fez-se necessária a expansão da ontologia da física. Aquilo que Chalmers pensa ser necessário, então, é realizar o mesmo com as teorias da consciência (CHALMERS, 1995, p. 14).

Curiosamente, como ele mesmo nota em sua obra, tentativas de se expandir teorias físicas com elementos não-físicos ocorreram até mesmo antes de Maxwell, como no caso do vitalismo. Naquela época, as pessoas acreditavam que não haveria explicação física possível para tratar o mistério da vida, ou seja, havia uma forte dúvida sobre se os mecanismos físicos poderiam mesmo desempenhar as muitas funções notáveis associadas aos seres vivos. Mas, uma vez que foi descoberto que processos físicos podiam, sim, desempenhar tais funções, as dúvidas e mistérios vitalistas se dissiparam e o fenômeno pôde ser explicado integralmente em termos físicos e deflacionários. Entretanto, para Chalmers, o mesmo não pode ocorrer com a fenomenalidade, pois ela não pertence à mesma categoria ocupada pelo *elán vital* dos vitalistas, não podendo, então, ser excluída das explicações da mesma forma que ele o foi. Em suas palavras:

É tentador notar que todos os tipos de fenômenos intrigantes acabaram sendo explicáveis em termos físicos. Mas cada um deles eram problemas sobre o comportamento observável de objetos físicos, chegando a problemas na explicação das estruturas e das funções. Por causa disso, esses fenômenos sempre foram o tipo de coisa que um relato físico poderia explicar, mesmo que em alguns pontos tenha havido boas razões para suspeitar que tal explicação não viria. A indução tentadora desses casos falha no caso da consciência, que não é um problema sobre estruturas e funções físicas. O problema da consciência é intrigante de uma maneira totalmente diferente. Uma análise do problema nos mostra que a experiência consciente simplesmente não é o tipo de coisa que uma explicação totalmente reducionista conseguiria explicar. (1995, p. 13 [tradução nossa]).

Com isso, ele conclui que não há como explicar a fenomenalidade somente em termos físicos, introduzindo, por sua vez, uma possível explicação não-reduitiva baseada no dualismo naturalista – concebida a partir de intuições e voltada ao fomento de ainda mais intuições acerca de nossa própria vida mental. Conforme anteriormente mencionado, o poder das intuições, em alguns casos, pode nos levar à argumentações que se distanciam progressivamente da realidade científica. Penso que o trabalho de Chalmers seja uma boa

ilustração desse processo. Aqui, a busca por uma explicação ou teoria que trate a consciência como algo essencial, não-fragmentável e imutável e que, ainda, seja capaz de explicar seus aspectos físicos (e não-físicos) nos mínimos detalhes, parece ser uma empreitada simplesmente fadada ao fracasso. O motivo para tanto se daria pela tentativa de adicionar propriedades cientificamente intratáveis ao fenômeno, exigindo, ao mesmo tempo, que essas sejam cientificamente explicadas em conjunto com suas demais propriedades. Tal busca jamais resultará em uma explicação passível de mensuração, tornando-se, então, algo infalsificável e, portanto, fugindo da seriedade que o próprio Chalmers exigiu para aquilo proposto no estudo da consciência. De fato, não haveria problema algum em tomar a rota diametralmente oposta e, deflacionando o significado de consciência, dedicarmo-nos a sub-fenômenos cientificamente explicáveis e concluir, por conseguinte, que aquilo objetivado poderia simplesmente decorrer desse conjunto de sub-fenômenos (Cf. SETH, 2021).

Outrossim, é importante levar em consideração que a própria existência do Problema Difícil é, em si, algo controverso e contestado por vários autores (Cf. MATSON, 1966; HODGSON, 1996; KING, 2007; LAU, MICHEL, 2019; MIRACCHI, 2019; SYTSMA, OZDEMIR, 2019; WIERZBICKA, 2019; ADAMS, BROWING, 2020; KIRKEBY-HINRUP, 2020; DÍAZ, 2021), assim como sua suposta universalidade, sua suposta dificuldade intrínseca e a validade da influência das intuições nesse tipo de julgamento. Algumas dessas objeções foram, inclusive, respondidas pelo próprio Chalmers (2020) que, após investigá-las, reconheceu que, se encaradas de uma certa maneira, elas poderiam levar o leitor a duvidar do *status* epistêmico do Problema Difícil – embora ele, não obstante, continue advogando pela existência de uma dificuldade singular em explicá-lo. Assim, e conforme o enunciado pelo próprio Chalmers, talvez essa questão deva ser primeiro reformulada para algo como: “por que nos parece haver um Problema Difícil da consciência” (CHALMERS, 2018 p. 1)?

1.2.2 O Metaproblema

Da mesma forma que a metacognição é a cognição sobre a cognição, o Metaproblema é o problema sobre o problema, nesse caso, o problema de como explicar o motivo pelo qual existe um problema da consciência (CHALMERS, 2018, p. 1). Logo, o Metaproblema estaria estreitamente relacionado aos relatos fenomênicos, isto é, àquilo que dizemos sobre a consciência, como, por exemplo, que ela representa um Problema Difícil. Esses relatos fenomênicos são intuitivos e relativamente fáceis de serem obtidos independentemente do

background epistêmico dos indivíduos. Para Chalmers, o Metaproblema seria o problema de explicar esses relatos fenomênicos, tratando-se, portanto, de um dos Problemas Fáceis. Isso significa que o Metaproblema deve ser explicado em termos físicos e funcionais, ao descrever, por exemplo, os mecanismos neuronais ou computacionais que geram tais relatos. Concomitantemente, por ser um Problema Fácil, o Metaproblema não seria capaz de resolver inteiramente o Problema Difícil, mas, dependendo da perspectiva do autor, ele pode ser capaz de dissolver o Problema Difícil ou sugerir um caminho para sua possível resolução (CHALMERS, 2018, p. 2).

Para Chalmers, a grande motivação para seguir o Metaproblema se dá por ele ser mais geral e neutro às teorias da consciência – isto é, ele implica uma explicação de tópico neutro, possibilitando a abertura e ampliação de diversas pesquisas sobre o tema, dado que a pergunta sobre quais mecanismos seriam responsáveis pelos relatos fenomênicos, mesmo que filosoficamente tratável, seria essencialmente um questionamento empírico. Em termos mais específicos, os alvos da explicação do Metaproblema seriam os relatos verbais (dizer que a consciência é difícil de ser explicada), os julgamentos (julgar que a consciência é difícil de ser explicada) ou simplesmente as disposições para fazer tais relatos e julgamentos, as quais são consideradas como as intuições do problema. Intuições acerca de problema geralmente resultam de julgamentos feitos em argumentos filosóficos. Entretanto, na busca por uma solução para o Metaproblema, Chalmers alega que as intuições mais relevantes seriam aquelas oriundas de julgamentos anteriores aos argumentos filosóficos, ou seja, julgamentos sem a carga teórica provida pela filosofia. Portanto, nesse contexto, o Metaproblema seria o problema de se explicar as intuições de problema que refletem a *sensação* de que existe algum tipo de problema misterioso envolvendo a experiência consciente – ou que existe uma lacuna explicativa entre os processos físicos e a consciência. Por exemplo, dizer que não há como explicarmos a consciência em termos físicos é um relato do problema, enquanto a disposição de julgar e relatar isso seria uma intuição do problema (CHALMERS, 2018, p. 5-6).

Especificamente, as intuições centrais ao Metaproblema seriam: (1) *as intuições explicativas*, que defendem a dificuldade de se explicar a consciência, a lacuna explicativa e a ideia de que a explicação das funções comportamentais não seria suficiente para explicar a consciência; (2) *as intuições metafísicas*, tais como as dualistas e as essencialistas ou fundamentalistas; (3) *as intuições de conhecimento*, tais como a intuição do conhecimento de primeira pessoa e a intuição da ignorância da terceira pessoa em relação à consciência dos outros organismos; (4) *as intuições modais*, sobre o que é possível ou concebível e sobre a

possibilidade de inversão da fenomenalidade. Além dessas intuições centrais, existiriam instituições periféricas fortemente relacionadas com o Metaproblema, tais como (5) *as intuições de valor*, que entendem a consciência como possuindo um valor especial; (6) *as intuições de distribuição*, acerca de quais sistemas possuiriam ou não consciência; (7) *as autointuições*, relativas ao *self* ou ao sujeito da experiência; (8) *as intuições de qualidade*, referentes às qualidades fenomênicas supostamente apresentadas na experiência; e (9) *as intuições de apresentação*, referentes à maneira direta como a fenomenalidade seria apresentada à nós (CHALMERS, 2018, p. 6).

Diante de tal exposição, Chalmers reconhece a dificuldade de sabermos o quão amplamente compartilhadas as intuições do Metaproblema seriam, dada a não-universalidade e a possível rejeição de alguns desses vieses cognitivos intuitivos por certos indivíduos ou grupos. Independentemente disso, o Metaproblema deveria ser eventualmente capaz de explicar o motivo pelo qual compartilhamos algumas intuições e o motivo pelo qual não compartilhamos outras – se, de fato, esse for o caso (CHALMERS, 2018, p. 8). Para Chalmers, devemos partir do seguinte pressuposto rumo à solução do Metaproblema:

Temos modelos introspectivos implantando conceitos introspectivos de nossos estados internos que são amplamente independentes de nossos conceitos físicos. Esses conceitos são introspectivamente opacos, não revelando nenhum dos mecanismos físicos ou computacionais subjacentes. Nossos modelos perceptivos atribuem perceptivamente qualidades perceptivas primitivas ao mundo, e nossos modelos introspectivos atribuem relações mentais primitivas a essas qualidades. Parece que temos conhecimento imediato de que estamos nessas relações mentais primitivas com qualidades primitivas, e temos a sensação de estarmos familiarizados com elas (2018, p. 23 [tradução nossa]).

Esse senso de familiaridade (*acquaintance*) seria a principal característica que as teorias realistas comprometidas com o Metaproblema deveriam explicar, especificamente porque tal senso seria atribuído às qualidades primitivas das nossas experiências – ou seja, o porquê de nossas experiências com as supostas qualidades do ambiente nos parecerem primitivamente familiares e serem, elas mesmas, objetos de nossa familiaridade. Isso sugere a necessidade de uma combinação entre uma perspectiva representacional de primeira ordem para o trato da consciência (envolvendo a percepção imediata das propriedades mundanas) e uma perspectiva auto-representacional (envolvendo a percepção imediata de si mesmo) (CHALMERS, 2018, p. 28).

Com tais considerações em mente, o Metaproblema, em um primeiro momento, parece fornecer um excelente caminho a ser tomado e, de fato, trata-se de um grande avanço de

raciocínio se comparado àquilo desenvolvido por Chalmers em sua exposição dos Problemas Fáceis e Difícil da consciência. Ainda, ele busca reconhecer a raiz da maioria das problemáticas filosóficas a respeito da temática: as intuições. E, a forma como ele elabora o Metaproblema, especificamente ao tratá-lo como o problema de resolvermos as intuições do problema ou nossas disposições de julgar ser extremamente difícil explicar a consciência, parece atacar diretamente tal raiz. Porém, o conjunto de possíveis soluções e pressupostos elencados por Chalmers falha em delinear o Metaproblema em tópicos satisfatoriamente neutros. De fato, tratar a fenomenalidade da consciência como uma forma de entidade, a qual nos permite familiarizar-nos com ela mesma e com as coisas do mundo, a qual não seria necessariamente física e que talvez seja independente de nossos esforços cognitivos internos e externos, mesmo quando canalizada por nossa perspectiva de primeira pessoa, não pode ser considerado um tratamento neutro. Na verdade, durante todo o discurso a favor do Metaproblema, Chalmers indiretamente defende sua posição dualista naturalista original (Cf. 1995), de modo que, resumidamente, ele acaba por insinuar que as intuições do problema seriam explicáveis em termos físicos, mas que a fenomenalidade continuaria a ser explicável apenas de outra maneira (Cf. CHALMERS, 2018). Tal posicionamento mantém a lacuna explicativa intacta e não parece realmente permitir o progresso do campo rumo a uma resolução ou dissolução do Problema Difícil.

1.2.3 Outros problemas da consciência

A influência da formulação que Chalmers forneceu para os problemas da consciência é inquestionável, pois neles, ele conseguiu resumir muito bem aquilo que cientistas e filósofos há muito tempo estavam intuitivamente se indagando. Porém, como o próprio Chalmers (2018) veio a reconhecer, precisamos ter cuidado com os caminhos pelos quais as intuições podem nos levar, especialmente naquilo que tange ao conhecimento sobre nós mesmos. Nisso, se comparado ao Problema Difícil, o Meta-problema não parece ter gozado de muita fama e influência, provavelmente por se focar na descrição das próprias intuições do problema e em como elas também seriam difíceis de serem explicadas, justamente por serem extremamente óbvias e intuitivas para alguns. Aqui, o maior problema é que, diferentemente daquilo que Chalmers faz, não parece ser sensato descrever a raiz intuitiva das obviedades relacionadas ao Problema Difícil e, ao mesmo tempo, continuar confiando precipitadamente em algumas delas – isto é, de que existe algo como uma fenomenalidade com propriedades que impossibilitam,

direta ou indiretamente, seu trato científico. Portanto, talvez seja interessante partir daquilo levantado pelo Meta-problema, mas sem que, com isso, deixemos de questionar a suposta obviedade comumente atribuída à fenomenalidade.

Uma forma de questionar tal obviedade é através do Problema da Ilusão, formulado pelo filósofo Keith Frankish (2017). Esse problema surge da tese ilusionista, a qual defende que as descrições tradicionais¹³ da fenomenalidade da consciência seriam ilusórias. Isto é, não haveria um como-é-ser nas experiências conscientes e, aquilo que de fato ocorreria, poderia ser plenamente explicado por meio de termos físicos. Assim, o papel do filósofo e do cientista da consciência seria, aqui, explicar o motivo pelo qual acreditamos que essa fenomenalidade existe em nossas experiências – em vez de tentar explicar tal fenomenalidade, em si – e o porquê dela nos parecer tão real e vívida (FRANKISH, 2017, p. 17). Digno de nota, quando Chalmers primeiro apresentou o Meta-problema (2018), ele dedicou parte de seu argumento para apontar diversos motivos pelos quais ele acreditava que as particularidades de sua formulação tonariam-na melhor do que a formulação fornecida por Frankish, especialmente por valorizar a ideia de uma suposta neutralidade. Mas, conforme o anteriormente exposto, o argumento de Chalmers falha em ser devidamente neutro, pois ele continua a intuitivamente defender a realidade da fenomenalidade tal como ela nos aparenta ser. Logo, sob certa perspectiva, faz sentido conceber o Problema da Ilusão como uma alternativa mais neutra ao Meta-problema, justamente por ele questionar a (obviedade da) realidade de tal característica do fenômeno da consciência.

Mesmo que o termo “Ilusionismo” tenha sido primeiramente utilizado por Frankish em 2015, ao longo do percurso filosófico moderno e contemporâneo, é possível encontrar inúmeros trabalhos total ou parcialmente enquadrados pela perspectiva ilusionista sem, com isso, utilizar suas terminologias, além de pontos de vista distintos, porém (implícita ou explicitamente) favoráveis e/ou simpatizantes àquilo enunciado por ela.¹⁴ De acordo com aquilo concebido pela literatura ilusionista, ilusões estariam relacionadas com algum grau de

¹³As conceitualizações comumente fornecidas por dualistas, monistas neutros e misterianistas – ou seja, realistas radicais – fisicalistas e representacionalistas – isto é, realistas conservadores (FRANKISH, 2017, p. 16).

¹⁴ Frankish mencionou alguns desses trabalhos, começando pelo de Kant (1781), o qual argumenta que uma “ilusão transcendental” é responsável por nossa intuição de que o *self* é uma substância simples. Ullin Place (1956) atribui uma série de “falácias fenomênicas” acerca da consciência às nossas intuições dualistas, David Armstrong (1968) discorre a respeito daquilo que denomina “a ilusão da mulher sem cabeça” e, mais notoriamente, Daniel Dennett (1991) introduz a famosa “ilusão do usuário” (CHALMERS, 2018, p. 8). Atualmente, podemos também encontrar um número crescente de simpatizantes explícitos do raciocínio ilusionista, tais como o psicólogo Nicholas Humphrey, a filósofa Susan Blackmore, o cientista Michael Graziano (Cf. FRANKISH, 2017) e, conforme tentarei demonstrar, o neurocientista Anil Seth (2021) – cuja tese pode ser considerada uma ramificação do Ilusionismo.

impenetrabilidade cognitiva de nosso sistema perceptivo (FRANKISH, 2017, p. 18).¹⁵ Entretanto, conforme o desenvolvido pelo filósofo François Kammerer, a ideia central do Ilusionismo seria a de que haveria algo como uma ilusão da fenomenalidade, a qual talvez seja mais forte do que as demais ilusões com as quais nos deparamos em nosso dia-a-dia (tais como as que incidem sobre nossa percepção visual, por exemplo).¹⁶ Trata-se obviamente, de uma tese cujas implicações percorrem os interesses de, praticamente, toda a filosofia da mente, pois diz respeito ao nosso universo mental e toda à nossa preciosa individualidade – questionando premissas clássicas, tais quais as de que essa vida mental não seria volátil e não poderia ser “completamente dissolvida” por explicações físicas (KAMMERER, 2017, p. 156).

Disso, o Problema da Ilusão parece evitar a lacuna explicativa fomentada pelo Problema Difícil, uma vez que não parte do pressuposto intuitivo de que a fenomenalidade é tal qual (achamos que) a conhecemos. Em outras palavras, o Problema da Ilusão evita atribuir um tratamento especial para a perspectiva de primeira pessoa e a convicção da existência da fenomenalidade, reconhecendo a importância da introspecção e de outras formas de “agência interna”, ao mesmo tempo em que também reconhece que essa agência não seria inteiramente confiável. Para os ilusionistas, a motivação para tal posicionamento é que, em casos como os da introspecção, nós nunca estaríamos, de fato, observando ou agindo diretamente sobre nossos processamentos e dinâmicas corporais e cerebrais, mas apenas confeccionando narrativas (verbais ou não) daquilo que, nesses processos e dinâmicas, já teria sido processado e disponibilizado para relato (DENNETT, 2018, p. 2; Cf. DENNETT, 1991).

Aqui, uma vez estabelecidos o Problema da Ilusão e o pressuposto da ilusão da fenomenalidade, estabelece-se, também, um padrão para a estruturação de problemas da consciência que busque explicar tanto a ilusão da fenomenalidade quanto o motivo pelo qual não conseguimos escapar dela (Cf. DENNETT, 2016). Um dos mais recentes problemas estruturados desde essa orientação consiste no Problema Real, desenvolvido pelo neurocientista Anil Seth (2021). Em sua obra, Seth nota a importância da mensuração para os estudos científicos contemporâneos e, partindo dessa consideração, esboça os contornos gerais de uma teoria e programa de pesquisa cujos objetivos seriam mensurar a fenomenalidade. Nisso, mais especificamente, o Problema Real se refere a explicar o porquê das experiências

¹⁵ Embora Frankish tenha esboçado essa hipótese em seu texto, veremos nos últimos capítulos que a fenomenalidade, na verdade, estaria mais especificamente relacionada com o funcionamento interoceptivo do que com os demais direcionamentos de nosso aparato perceptivo (tais como o introspectivo, que seria, não obstante, importante para o relato [verbal ou não] da fenomenalidade).

¹⁶ Poderíamos diferenciar, ainda, ilusões perceptivas, como essas, de ilusões teóricas, tal como o proposto por Dennett (2017, p. 2).

conscientes serem da forma que são – isto é, (parecem) possuir as propriedades fenomênicas que possuem – em termos de processos físicos, cerebrais e corporais. Ou seja, como e por que um determinado padrão de atividade cerebral – ou outras dinâmicas/processos físicos – representa, para os indivíduos, um determinado tipo de experiência consciente? (SETH, 2021, p. 31). Partindo dessa formulação, Seth almeja o desenvolvimento de um método científico que seja capaz de prever, explicar e controlar a fenomenalidade da consciência, com a solução de seu Problema Real permitindo a previsão de quando uma experiência subjetiva ocorreria e como ela poderia ser controlada através de intervenções sobre seus mecanismos subjacentes.

A diferença entre o Problema Real e o Problema Difícil é que o primeiro nega a existência de uma lacuna explicativa e nega, no estudo da consciência, a necessidade de uma expansão das explicações físicas com algum ingrediente não-físico, ou um “tempero mágico” que elucidaria como a fenomenalidade surgiria desde mecanismos físicos (SETH, 2021, p. 30). Além disso, o Problema Real não pode ser resumido a um Problema Fácil (pelo menos não da maneira como Chalmers descreve tal questão), uma vez que (à sua própria maneira) seu objetivo é justamente a explicação da fenomenalidade e não de qualquer função ou comportamento associados a ela. Assim, seu desafio é prever e controlar as supostas propriedades fenomênicas da consciência em termos de nossa atividade corporal e cerebral, tratando-as como tal. Por exemplo, a propriedade fenomênica do vermelho (comumente tratada como “vermelhidão”) seria aqui tratada como uma propriedade distintivamente visual, frequentemente ligada aos objetos e aparentemente relativa às superfícies percebidas, com seus diversos níveis de saturação, brilho, etc, de modo que tais características sejam tratadas como parte do próprio espectro fenomênico, e não como funções ou comportamentos associados à fenomenalidade (SETH, 2021, p. 31).

Dito isso, embora Seth não elabore nem defenda explicitamente o Problema da Ilusão, é possível identificar que ele indiretamente parte do pressuposto de que a fenomenalidade não é como ela nos parece ser, sua presença e implicações não são óbvias e que é plenamente possível que a ciência contemporânea eventualmente explique os detalhamentos de nossa experiência consciente de maneira a incluir tal fenomenalidade. Mais especificamente, a solução para o Problema Real seria aquilo que Seth batizou de teoria da “Alucinação Controlada” (Cf. SETH, 2021), a qual é totalmente elaborada em termos da atuação de mecanismos cognitivos conforme sua descrição pelo Processamento Preditivo e trabalha a percepção como uma alucinação controlada. Entretanto, parece que, em sua formulação, Seth confunde e, portanto, inverte o significado e as implicações de alucinação e ilusão – o que, por

consequente, o faz negar parte da proposta ilusionista. As alucinações consistem em uma percepção de algo que não existe no mundo, enquanto as ilusões consistem na percepção de algo de uma maneira diferente daquilo que ele realmente é. Ou seja, quando falamos de ilusões perceptivas, não estamos negando sua relação com o que supostamente existe no mundo externo, mas negando que a forma manifesta através da qual o percebemos seria fidedigna à sua estruturação (extra-perceptiva); ao passo que alucinações perceptivas ocorreriam independentemente do que há no mundo externo.¹⁷ Em sua formulação, Seth tenta contornar tal problemática idealista ao se valer do adjetivo “controlada”, de modo que suas alucinações seriam, na realidade, dependentes daquilo estruturado e possibilitado pelo mundo externo (dadas as entradas sensoriais recebidas), mas, com isso, ele simplesmente passa a operar internamente dentro daquilo que o Ilusionismo denomina “ilusão”.¹⁸

Logo, parece que investir no termo ilusão seria um caminho frutífero para continuar a desenvolver, internamente ao programa ilusionista, aquilo proposto pelo Problema Real e, conseqüentemente, uma solução para o Problema da Ilusão. Na verdade, os dois problemas são significativamente semelhantes, quase como se um fosse uma ramificação do outro, dado que o Problema da Ilusão busca explicar o motivo pelo qual a fenomenalidade nos parece existir, enquanto o Problema Real pode ser formulado como um questionamento acerca da maneira através da qual essa aparência – que nos é real – ocorreria em termos cerebrais e corporais. Assim, ao simplesmente reconceituarmos aquilo que Seth chama de “alucinação”, podemos nos orientar para uma explicação satisfatória não apenas para o Problema da Ilusão, mas que caminhe em direção a uma resolução para o Problema Real, negando a necessidade de um Problema Difícil ou dissolvendo-o em algo semelhante a um conjunto de Problemas Fáceis (Cf. DENNETT, 2018). De fato, Seth alega que, da mesma forma como os vitalistas eventualmente aceitaram que não havia um *élan vital*, mas sim um conjunto bastante vago de critérios ou características físicas que, juntas, comumente denominamos “vida”, dualistas/misterianistas/fisicalistas-não-reducionistas eventualmente aceitarão que não existem propriedades fenomênicas misteriosas, intratáveis e não-físicas a serem estudadas e explicadas, mas sim um conjunto de critérios ou características que, juntas, resultam naquilo

¹⁷ Essa observação se aproxima, então, daquilo enumerado pelo posicionamento denominado “disjuntivismo ecológico” o qual alega que, diferentemente de ilusões, alucinações simplesmente não seriam fenômenos perceptivos (Cf. CARVALHO, 2021).

¹⁸ A proposta, então, é tratar ilusão e alucinação como fenômenos distintos e substituir a palavra “alucinação” pela palavra “ilusão” nas descrições preditivas de Anil Seth.

que chamamos de experiência consciente subjetiva ou, simplesmente, consciência (SETH, 2021, p. 34-35).

Entretanto, ainda precisamos introduzir um elemento importante nesse cenário de questionamentos e problematizações acerca da consciência: a Pergunta Difícil. Proposta por Daniel Dennett em 2018, ela consiste em nos questionar: uma vez que algo “entra” em nossa consciência, o que isso faz, causa, permite ou modifica? Isto é: “e depois, o que acontece?” Para Dennett, é demasiadamente comum que aqueles que estudam a consciência fiquem perplexos com o mistério da fenomenalidade e, nisso, muitas vezes falhem em reconhecer que, na realidade, eles já estariam trabalhando-a de certa maneira, sem qualquer necessidade de uma revolução filosófica, científica ou conjecturar propriedades “misteriosas” não-físicas (Cf. DENNETT, 2018). Nisso, existiriam quatro motivos principais que influenciaram os acadêmicos a seguir caminhos intuitivos e, portanto, cair em armadilhas epistêmicas que lhes fazem negligenciar os impactos de suas próprias descobertas científicas acerca da consciência e sua “misteriosa fenomenalidade”: (1) a ideia da dupla transdução – isto é, a suposta necessidade de o conteúdo dos estímulos ser primeiramente transformado em disparos neuronais para, somente depois, ser novamente transformado em conteúdo perceptivo para alguma forma de reprodução interna e projeção externa; (2) a ideia de que haveria um local para a fenomenalidade no cérebro; (3) a ideia da consciência ser uma espécie de “filme contínuo em nossas cabeças”, um processo ininterrupto, imutável e irreduzível; e (4) o fato de que não seria possível acessar os representantes¹⁹ da experiência consciente.

Ainda para Dennett, parece haver, de forma mais problemática, uma orientação equivocada nos próprios estudos científicos acerca da consciência, dado que eles geralmente são projetados para analisar as operações de transdução de estímulos artificiais para as ativações neuronais de indivíduos. Isso implica em isolar, o máximo possível, algumas variáveis de entrada sensorial para medir as possíveis variáveis dependentes de saída – isto é, confinar o sujeito em um local ou restringir os estímulos que ele recebe para que, com isso, seja possível mensurar e controlar os dados resultantes com certa precisão. Porém, o problema

¹⁹ Para Dennett, os objetos externos de nosso mundo, tais como as imagens, possuem tanto representantes (digamos, uma fotografia) quanto representados (o que foi fotografado) que seriam acessíveis à nossa experiência. Entretanto, objetos internos (os quais dizem respeito à nossa vida mental) não possuiriam representantes acessíveis à nós, mas apenas aquilo que é representado. Isto é, teríamos acesso apenas aos conteúdos representados e a uma ordem temporal aparente desses conteúdos, de modo que o que quer que sejam os representantes desses objetos, como e porque eles são gerados, além de quando eles nos “aparentam ser” é algo completamente fora dos limites de nossas capacidades introspectivas. E, para Dennett isso acabaria por gerar a intuição de que há algo a mais para ser explicado, de que as várias explicações fisicalistas deixam de lado aquilo que apenas nós mesmos teríamos a autoridade para estudar: a fenomenalidade (DENNETT, 2018, p. 4).

de restringir os sujeitos nos experimentos científicos é que eles acabam restritos a um conjunto extremamente pequeno e irrealista de possibilidades do que podem fazer com sua consciência. Isso acaba por desviar ou simplesmente impossibilitar que a atenção das ciências da consciência se voltem para aquilo que deveria ser a resposta geral da Pergunta Difícil: “quase qualquer coisa pode acontecer!” (DENNETT, 2018, p. 3).

Dada a infinidade de conexões neuronais, dinâmicas corpóreas e a quantidade de conteúdo que essas podem transmitir, a todo momento, nossas mentes são capazes de praticamente qualquer reação possível:

Não apenas podemos pensar sobre qualquer coisa que pode nos ocorrer, e não apenas quase qualquer coisa (qualquer coisa 'imaginável', qualquer coisa 'concebível') nos ocorre, mas uma vez que algo nos ocorreu, podemos responder a isso de uma forma aparentemente variedade ilimitada de maneiras e, em seguida, responder a essas respostas em outra grande variedade de maneiras e assim por diante, um conjunto crescente de possibilidades que supera até mesmo a produtividade das línguas naturais (faltam-me as palavras). É claro que, em qualquer ocasião particular, os estados momentâneos dos vários sistemas neurais componentes restringem o “possível adjacente” a uma seleção limitada de conteúdos “próximos”, mas isso muda de momento a momento e não está diretamente sob o controle de ninguém. [...] Vale a pena notar que temos poucos motivos para pensar que sistemas nervosos mais simples tenham uma produtividade semelhante. A maioria provavelmente são sistemas 'cognitivamente fechados', sem recursos representacionais para imaginar um século ou um continente, ou poesia, ou democracia, ... ou Deus. Os famosos quatro Fs (alimentar/feed, lutar/fight, fugir/flee e acasalar/mate) podem, com alguns suplementos (por exemplo, explorar, dormir) e subopções menores, esgotam os graus de liberdade dos invertebrados (DENNETT, 2018, p.3 [tradução nossa]).

Isso significa que aquilo que atualmente é mensurado pelas metodologias habituais estaria relacionado a uma parcela significativamente pequena daquilo que seria a totalidade da atividade mental consciente. Caso reconheçamos a identidade dessa parcela perante a totalidade da consciência e sua real importância para a compreensão do fenômeno, isso simplesmente não se torna um problema. Mas, se tomarmos o pressuposto de que a atividade mental consciente seria una e indivisível, esse tipo de estudo empírico acaba por estar condenado ao erro mereológico ou à sua categorização como uma análise de algum dos itens contidos nos Problemas Fáceis. Ainda, se tais limitações forem devidamente reconhecidas, poderíamos ampliar as discussões acerca de como, de fato, estudar a atividade mental consciente de maneira menos restritiva, mesmo considerando que a completude do fenômeno esteja longe do escopo da grande maioria dos estudos científicos – uma vez que, segundo Dennett, a consciência ou qualquer um de nossos estados conscientes não possuem início (ou T0) e muito possivelmente envolvem não somente milhares de conexões neuronais, mas

também uma variedade de sistemas se estendendo por todo o corpo e até mesmo pelo ambiente no qual este corpo se encontra. Infelizmente, levar esses pontos em consideração foi um posicionamento adiado ou, simplesmente, ignorado pelo atual *mainstream* científico e filosófico (DENNETT, 2018, p. 4).

A título de exemplo, sempre que estamos conscientes, estamos conscientes de algo – um conteúdo relacionado a qualquer atividade cognitiva – para (fazer) algo – interagir com o mundo interno e externo – o que significa que não há consciência sem um motivo e objetivo funcional para estarmos nesse estado, isto é, sem intencionalidade²⁰ (HARLEY, 2021, p.12). Logo, qualquer tratamento deflacionário da consciência deve considerá-la como uma função do corpo (Cf. DENNETT, 2011). Mas, por que nosso corpo possuiria tal função? Ou melhor: como e “por que” nos tornamos seres conscientes? Ainda, vários questionamentos acerca do funcionamento dos mecanismos e sub-mecanismos da consciência carecem de melhor resolução. Por exemplo, é amplamente conhecido que, a todo instante, nossos órgãos sensoriais captam informações do ambiente externo e interno. Disso, como o cérebro filtra e elege as informações que irão compor a consciência perceptiva? Essa consciência é imediata²¹ e instantânea? Se não há um ponto de chegada, onde tudo se junta, o que há? Haveria um fluxo neuronal canônico da consciência? Como ocorre a transição do conteúdo inconsciente para a consciência? Finalmente, por que a consciência é tal como a experienciamos? Obviamente, não há como fornecermos uma resposta definitiva, hoje em dia, para nenhuma dessas perguntas. No entanto, se levarmos as premissas enumeradas na introdução à sério, o cenário com o qual nos deparamos parece ser o de dados e uma literatura suficientes para que filósofos e teóricos, em termos gerais, tenham possíveis explicações capazes de compor uma descrição deflacionária convincente do fenômeno.

Por fim, a Pergunta Difícil aparenta postular um grande impasse ao Problema Real, pois, se quase qualquer coisa pode acontecer depois que um conteúdo entra na consciência, como mensurar, prever e controlar isso? Um progresso inicial pode advir de assumirmos que nossas capacidades de introspecção não nos permitem confiavelmente mensurar tais atividades, dado o estado ilusório permanente de nossas vidas mentais. No entanto, uma resposta coerente para a Pergunta Difícil parece ainda ter que lidar com uma forte integração

²⁰ Embora existam duradouros debates acerca de se a consciência seria legitimamente intencional ou não, aqui estou interessada numa forma fraca e deflacionada de intencionalidade, a qual não está relacionada somente com o nível pessoal da cognição, mas majoritariamente com o sub-pessoal, conforme será demonstrado ao longo dessa seção.

²¹ No sentido de ser um correlato imediato, ou seja, sem a necessidade de um mediador.

de informações, em múltiplos níveis de complexidade e em vários locais do cérebro, que permitiria a percepção subjetiva de um conteúdo consciente. Logo, uma saída possível partiria não somente da deflação do termo consciência em sub-fenômenos (SETH, 2021, p. 36), mas também da contextualização do resultado manifesto desses sub-fenômenos como a consciência. Ao que nos parece, as teorias filosóficas que partam dos pressupostos mencionados na Introdução devem ser capazes de realizar tal contextualização e explorar as implicações apresentadas por estudos empíricos que partam de pressupostos semelhantes (ou convergentes). De certa maneira, esse seria um dos objetivos gerais da presente dissertação.

1.3 A fuga do Teatro Cartesiano

Atualmente, os estudos cientificamente embasados acerca da consciência tendem a seguir dois caminhos distintos: (1) buscar um critério mínimo ou basilar para a consciência (Cf. KOCH, et al. 2002; TONONI, 2004) ou (2) identificar um conjunto de características neuronais, sistêmicas, psicológicas e sociais que possam configurar um organismo consciente (SETH, BAARS, EDELMAN, 2004, p. 121-122; TYLER, 2020, DAMÁSIO, MEYER, 2009, p. 4; PENNARTZ, FARISCO, EVERS, 2019, p. 3). Similarmente, muitas das teorias contemporâneas da consciência seguem uma dicotomia que lembra aquela dos problemas chalméricos. Por exemplo, Damásio e Meyer descreveram a consciência através de seus supostos aspectos internos, tais como a subjetividade e a própria vida mental dos indivíduos, as quais dizem respeito à perspectiva da primeira pessoa, e aspectos externos, ou seja, comportamentais, os quais dizem respeito à perspectiva objetiva da terceira pessoa (DAMÁSIO, MEYER, 2009, p. 4). O físico-químico Gerald Edelman propôs a divisão entre consciência primária, que seria responsável pelas atividades mais basilares de nossa cognição, e a consciência secundária, que seria responsável pelas nossas capacidades de abstração e metacognição (Cf. EDELMAN, 2003). Ainda, o filósofo Ned Block propôs uma divisão que se tornou bastante influente na filosofia da mente moderna, a qual distingue entre a consciência de acesso – isto é, o conteúdo da percepção, memória e outras parcelas da cognição acessíveis à nós – e a consciência fenomênica – que seria a sensação subjetiva de experienciar conscientemente o mundo interno e externo (Cf. BLOCK, 1995).

Aquilo que Dennett argumenta, no entanto, é que essas dicotomias, ainda que apoiadas em evidências empíricas, estariam desnecessariamente fomentando uma lacuna epistêmica entre “processos cientificamente explicáveis” e a “subjetividade fenomênica”, além de

paradoxalmente compreenderem a consciência como algo irreduzível e imutável em suas empreitadas. Para ele, existe uma forte tendência – sobretudo na filosofia – de conceber a consciência como o resultado de uma transição daquilo não percebido para aquilo percebido, em uma escala macro e manifesta de espaço-tempo, quase que como uma necessidade de localizá-la ao final do processamento das entradas sensoriais, em algum local “um pouco antes das saídas que implementam nossas ações” (DENNETT, 1991, p. 108). Curiosamente, nessas narrativas, parece ser necessária a existência de algo como um salto, no trajeto tomado pela entrada sensorial: por exemplo, do estímulos ao nervo óptico, passando pelas diferentes áreas do córtex visual e, então – passando por um “ponto de virada” para, finalmente – chegar à consciência. Esse salto, constantemente estipulado por correntes tradicionais de pensamento acerca da consciência, implica a existência daquilo que Dennett intitulou de Teatro Cartesiano (Cf. DENNETT, 1991).

O Teatro Cartesiano consiste em uma metáfora de como o cérebro processaria o conteúdo perceptivo, propondo a existência de um produto final sendo passivamente exibido em um projetor central que seria acessível ao sujeito (DENNETT, 1991, p. 39). A terminologia “cartesiano” remete à ideia de Renè Descartes (1596-1650) de que haveria algo como um ponto central de comunicação entre o corpo físico e a mente não-física, onde tudo se juntaria. Os defensores da existência desse ponto central de comunicação seriam, para Dennett, materialistas cartesianos: materialistas, pois buscam a existência de um ponto físico no cérebro (contrastando, por exemplo, com os dualistas de substância²²) e cartesianos, pois alimentam a ideia de que haveria um ponto de chegada para um fluxo único e central da consciência, tal como aquilo creditado ao pensamento de Descartes.²³ Considerando que tais ideias e tendências perduraram por séculos na filosofia e nas ciências da mente, Dennett reconhece que não é fácil superar a metáfora do Teatro Cartesiano, sobretudo porque ela diz respeito a uma das mais fortes intuições que possuímos acerca de nós mesmos. Mas, também, não se trata de algo impossível. De fato, aqui, a alegação será a de que, uma vez que algumas das premissas dennettianas são somadas aos pressupostos ilusionistas, sellarianos e às *affordances* anteriormente mencionadas, a fuga do Teatro Cartesiano se torna uma empreitada plenamente exequível.

²² A perspectiva, na filosofia da mente, de que os chamados “fenômenos mentais” são não-físicos e que a mente e o corpo seriam substâncias distintas e separáveis (CRANE; PATTERSON, 2000, p. 1-2).

²³ Atualmente, uma das localizações mais discutidas como candidata para esse “ponto de chegada” da consciência seria a conexão entre o tálamo – uma estrutura no mesencéfalo conhecida por integrar e lidar com as informações recebidas pelos órgãos sensoriais – e o córtex (Cf. MIN, 2010; ARU, et. al, 2019; REDINBAUGH, et. al, 2020).

A primeira dentre essas premissas é a compreensão de que os estados conscientes são (temporalmente) indetermináveis. Dennett buscou demonstrar o *status* ilusório da noção de um ponto de chegada para a consciência com um experimento mental, elaborado de maneira a evidenciar a indeterminação da ordem de nossos processos mentais conscientes. O experimento consiste em um indivíduo conceber, erroneamente, que uma mulher que ele viu na rua estaria usando óculos (que ele viu anteriormente, em uma outra ocasião). A partir desse evento, Dennett nos oferece duas formas de explicação para o ocorrido: (1) a denominada “orwelliana”, na qual o sujeito tem um breve momento perceptivo real – no qual a mulher é percebida sem os óculos – mas a lembrança recente de alguém usando óculos acaba por contaminar tal percepção, quase que imediatamente, fazendo com que a experiência “verdadeira” não deixasse vestígio algum na memória; (2) enquanto uma outra forma de explicação seria a conhecida como “stalinense”, na qual o sujeito, na verdade, nunca sequer percebeu a mulher sem os óculos, pois a memória dos óculos contamina a informação sensorial sobre a mulher antes mesmo do ato perceptivo ocorrer. Com isso, Dennett conclui que:

[A] distinção entre revisões pré-experimentais que mudam o que foi vivenciado e revisões pós-experimentais que têm o efeito de relatar ou registrar incorretamente o que foi vivenciado é indeterminada no limite [...] Quando fazemos uma pergunta a um sujeito sobre se uma confissão pública específica captura ou não adequadamente a verdade interior última sobre o que ele estava experimentando, o sujeito não está em melhor posição para julgar do que nós, estranhos (DENNETT, 1991, p. 247 [tradução nossa]).

Para Jared Warren (2021), essa indeterminação – que é relativamente bem-aceita e reconhecida pelas ciências da mente, mas não tanto pela filosofia – é um fator positivo para os estudos funcionais da consciência e deve ser levado a sério. Pois, se não há diferenças relevantes, para a consciência, entre as duas formas de explicação, não há como saber qual explicação seria, de fato, a correspondente com o que aconteceu. Logo, a grande conclusão desse experimento consiste na flagrante indeterminação de se um estado mental global, em um momento qualquer, ser consciente ou não – e, portanto, um apontamento do quão problemático seria operar internamente a tal distinção (WARREN, 2021, p. 9-10).

A segunda premissa a ser tomada na fuga do Teatro Cartesiano, de acordo com Dennett, consiste em tratar o cérebro como possuindo uma arquitetura de redes neuronais aninhadas, implicando um funcionamento paralelo e distribuído, porém integrado (DENNETT, 1991, p. 241). Na verdade, a ideia de que, de alguma forma, o cérebro é responsável ou está

envolvido na geração da mente através da atuação e estruturação de “sistemas dentro de sistemas” é algo bem aceito nas ciências cognitivas (WARREN, 2021, p. 10). Para desenvolver essa ideia, então, Dennett utiliza o conceito de uma arquitetura cognitiva “*pandemonium*” (Cf. McCARTHY, 1959), o qual envolve “*demônios*” aninhados em uma estrutura semi-hierárquica. Esses *demônios* competem entre si, “gritando” as informações que possuem. Aquele que melhor (ou mais) gritar, terá a sua informação transmitida adiante (DENNETT, 1991, p. 253-254). Pelo fato de haver vários *demônios* em paralelo, sem a necessidade de um supervisor, esse argumento fornece um modelo contra não somente o Teatro Cartesiano, mas também o chamado “observador interno” ou homúnculo que, em algumas teses filosóficas ou interpretações científicas compatíveis com o Teatro Cartesiano, seria um revisor final (isto é, aquela entidade que toma a decisão cognitiva), o qual preencheria e complementaria o “ponto de virada” necessário para essas perspectivas (WARREN, 2021, p. 10-12).

Finalmente, a terceira premissa necessária para se fugir do Teatro Cartesiano é compreender seriamente que a aparência nem sempre equivale à realidade. Na filosofia, a introspecção (ou o método da primeira pessoa, por excelência) sempre foi a principal ferramenta de exploração da consciência. Mais especificamente, de acordo com Schwitzgebel (2019), a introspecção é o processo através do qual o sujeito se torna ciente daquilo que se passa em sua mente em um momento específico. Para o autor, as características cognitivas que melhor qualificam um processo como introspectivo envolvem a geração de conhecimentos, juízos e crenças na condição de: (1) mentalidade ou eventos mentais, distinguindo-se de processos sensoriais que fornecem informações sobre eventos externos ou aspectos não-mentais; (2) processos na primeira pessoa, a respeito da própria mente; e (3) processos temporalmente proximais, sobre a vida mental no momento presente ou passado/futuro imediatos – ou seja, internamente a uma curta janela de tempo (Cf. SCHWITZGEBEL, 2019). Para Frankish (2017, p. 17), além dessas características, a introspecção também nos conduziria a realizar juízos fenomênicos, isto é, juízos sobre as supostas qualidades de nossas experiências conscientes. Pois, o próprio ato de introspecção parece acionar sistemas internos de automonitoramento (assim como nossos órgãos sensoriais são acionados com as estimulações do ambiente) que sondam o impacto dos objetos externos no cérebro e corpo, com os efeitos desse impacto sendo erroneamente interpretados como fenomênicos (FRANKISH, 2017, p. 28-29).

Disso, um dos problemas envolvidos na introspecção é o fato dos relatos resultantes dela serem frequentemente confundidos com aquilo que, de fato, ocorreria no (e com o) sistema: “eu realmente senti uma sensação inefável de vermelhidão ao olhar para este morango... veja só, isso acontece bem aqui neste momento, na minha neuroimagem”. Ora, se nossos relatos fossem fidedignos àquilo captado por nossos sensores, então o corpo humano não poderia, por exemplo, se valer dos olhos para a captação de estimulação perceptual, dado que os constantes movimentos sacádicos fariam com que nossa percepção visual resultasse em um cenário trêmulo e sem foco. Ainda, se, de fato, ouvíssemos tudo aquilo captado por nossos tímpanos, nossa percepção auditiva seria, constantemente, caótica e cheia de ruídos, tal como numa festa repleta e com música alta. O fenômeno da visão cega ou *blindsight* (Cf. AJINA, BRIDGE, 2016) é um importante exemplo daquilo que pretendemos ilustrar com isso tudo: nele, sujeitos com danos na área V1 do córtex visual, os quais não são capazes de conscientemente ver e, portanto, relatar o que estão vendo, se demonstram, não obstante, capazes de captar e processar as informações visuais de modo que, se forem pedidos para adivinhar quais objetos estão na sua frente, quase sempre irão acertar. Ou seja, a descrição daquilo que acontece no sistema não equivale ao que é conscientemente relatado.

Sinteticamente, para fugir do Teatro Cartesiano – isto é, evitar a tradicional concepção da consciência como consistindo numa espécie de filme a ser internamente projetado – faz-se necessário aceitar premissas tais quais a indeterminação do conteúdo consciente, o funcionamento paralelo e distribuído (porém integrado) do sistema e a distinção entre a aparência e a suposta realidade, naquilo que diz respeito à relação entre relatos introspectivos e o conteúdo captado por nossos sensores. Esse reconhecimento permite-nos melhor investigar o fenômeno nos moldes heterofenomenológicos²⁴ da imagem científica, pois se afasta da problemática ênfase e/ou exclusividade dos dados oriundos da introspecção e, ao mesmo tempo, possibilita a criação de novas hipóteses objetivando a resolução da Pergunta Difícil e, eventualmente, do Problema Real. Logo, de acordo com aquilo desenvolvido até aqui, seria somente ao escapar do Teatro Cartesiano que poderíamos começar a contornar os esboços de uma descrição teoricamente satisfatória e empiricamente coerente da consciência.

Decerto, algo semelhante a esse contorno foi recentemente estruturado por Anil Seth e Tim Bayne (2022), em sua proposta de que a única distinção necessária para uma boa teoria

²⁴ A aplicação do método científico em conjunto com um posicionamento antropológico, combinando auto-relatos subjetivos com todas as outras evidências disponíveis (em terceira pessoa) para a determinação de estados mentais. O objetivo do processo é descobrir como os sujeitos enxergam, eles mesmos, o mundo, sem levar a total precisão da perspectiva subjetiva em conta.

da consciência seria aquela entre estados globais e estados locais.²⁵ Aqui, estados globais consistem no perfil subjetivo geral de um organismo, estando ligados tanto ao espectro de *awareness* quanto às respostas comportamentais cabíveis – como, por exemplo, o estado de vigília, o sono e possíveis estados psicodélicos. Já os estados locais consistem nos conteúdos conscientes ou qualitativos, tradicionalmente caracterizados pelo “como-é-ser” da filosofia, estando ligados à percepção interoceptiva (isto é, voltada aos estados internos do organismo), exteroceptiva (voltada aos estados externos) e proprioceptiva (voltada aos estados produzidos e percebidos internamente a um organismo, mas que estão ligados ao seu posicionamento e à movimentação de seu corpo em um contexto). Por exemplo, a percepção de fome e desidratação no primeiro caso, a percepção de objetos, formas e cores no segundo caso e a percepção do espaço ocupado pelo corpo, naquilo que diz respeito ao terceiro caso. Adicionalmente, para Seth e Bayne, estados locais também abarcam conteúdo cognitivo e proposicional, tais como o raciocínio lógico e moral (SETH, BAYNE, 2022, p. 440). A relação e integração entre esses estados resulta em cenas momentâneas passíveis de serem conscientemente percebidas, de modo que a frequência de atualização desses estados gera tanto a estabilidade ilusória e manifesta que vivenciamos quanto a saliência de múltiplos de seus detalhamentos ao longo de um certo período, de acordo com os direcionamentos de nossos mecanismos atencionais.

Para os autores, a distinção entre ambos esses estados da consciência pode ser, ainda, ramificada em sub-estados, tornando possível e preferível investigar a consciência – seus estados e sub-estados – partindo de múltiplos pontos de vista, não havendo, portanto, a necessidade de tratá-la como um grande fenômeno irreduzível e indivisível, ou dificultar seus estudos com enunciados tais quais os do Problema Difícil e sua lacuna explicativa. Afinal, conforme aquilo proposto, a unicidade e a irreduzibilidade de nossa consciência seriam ilusórias e existiriam apenas em nossa imagem manifesta. Em termos científicos, então, o fenômeno da consciência poderia ser concebido como consistindo em um conjunto de múltiplas atividades e mecanismos neuronais e corporificados que (atualmente já) são ou (eventualmente) serão integralmente explicadas em termos físicos (Cf. SETH, BAYNE, 2022). Logo, tendo tal definição geral em mente, resta-nos agora focar (em maiores detalhes) o funcionamento dos estados locais da consciência, valendo-nos, para tanto, da inter-relação

²⁵ Em seu livro *Being You*, Seth descreve os estados globais como níveis de consciência (Cf. 2021), mas, aparentemente, decidiu por abandonar o termo “níveis” em estudos mais recentes, de maneira a evitar uma ordenação ou escala unidimensional para a consciência, dado que esses estados seriam melhor trabalhados internamente a um espaço de orientação multidimensional (SETH, BAYNE, 2022, p. 440).

entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo, os quais buscaremos introduzir nas próximas seções.

1.4 O Modelo dos Múltiplos Esboços (MME)

De acordo com Dennett, toda a atividade em nosso mundo interno e externo gera entradas sensoriais que são, simultaneamente, processadas como variáveis por diversas camadas de processamento que, por sua vez, geram vários fluxos de informação – que Dennett chamou de narrativas. Os “melhores” padrões dentre esses fluxos podem ganhar “fama no cérebro”, isto é, podem ficar disponíveis para o acesso consciente e aquilo que determinará quando e como isso irá ocorrer – e, com isso, a experiência consciente manifesta – é o que o autor denomina de “mecanismo de sondagem” (DENNETT, 1991, p. 111). A tese central do MME se volta para a questão da temporalidade dos processos conscientes, especificamente naquilo que diz respeito à inexistência de uma ordem de chegada das entradas sensoriais que compõem o conteúdo consciente. Dito isso, embora Dennett não entre em muitos detalhes acerca do funcionamento do MME (principalmente naquilo que diz respeito a escalas temporais distintas), objetivo, nesta seção, explorá-lo filosoficamente, de maneira a esclarecer seu funcionamento, tanto em termos manifestos quanto científicos.

O MME pressupõe que não experienciamos diretamente aquilo que nossos órgãos sensoriais captam, mas sim um produto dessa captura, o qual teria passado por diversos processos de “edição”. Essas edições seriam a união de informações sensoriais previamente processadas (e, provavelmente, armazenadas na memória) com informações novas e vindouras, advindas dos órgãos sensoriais, formando esboços – ou seja, variáveis representacionais que direcionam as atividades neuronais. Entretanto, nesse modelo, não é necessário que tais variáveis representacionais sejam cerebralmente reprocessadas (tal como numa projeção de seu conteúdo final) antes de serem experienciadas pelo sujeito. Ou seja, para o MME, esboços não precisam ser enviados para uma área cortical específica e, então, redistinguidos ou reinterpretados por algum distinguidor “mestre” (DENNETT, 1991, p. 112-113). Pois, ao passo que os esboços vão se enriquecendo ao serem atualizados com novas informações vindas de nossos sensores, eles tendem a ficar ativos por mais tempo, gerando influência sobre redes neuronais centrais e adjacentes, numa escala de “tempo cortical” medida em milissegundos. Podemos entender, então, que durante um breve (porém

significativo) tempo de ativação, os esboços se encontram “fixados” em nossos cérebros.²⁶ É esse processo de fixação que, em trabalhos posteriores, Dennett denominou de uma obtenção de “fama no cérebro” (Cf. DENNETT, 2006). Digno de nota, não seria simples (e, atualmente, sequer possível) observar exatamente, em um dado momento, quais dentre tais esboços irão compor a experiência consciente, visto que, na construção do cenário completo que experienciamos a nível pessoal, faz-se necessário “fixar” inúmeros esboços, em um mesmo momento da escala de “tempo cortical”, e desde múltiplos locais do encéfalo. Para Dennett, então, não faz sentido perguntar quando um conteúdo específico chega à consciência (DENNETT, 1991, p. 113).

Para fins ilustrativos, faça esta pequena atividade caso possua um computador com o sistema operacional Windows: em seu teclado, aperte Ctrl+Alt+Del e abra o Gerenciador de Ferramentas. Observe, na aba processos, todas as atividades que estão acontecendo nesse exato momento. Note que, provavelmente, o programa que você está utilizando ativamente (seja para ouvir música ou para ler esse texto) está consumindo porcentagens maiores dos recursos de sua máquina, embora existam vários outros processos dos quais você sequer sabia da existência que aparecem e somem periodicamente. Agora note que em alguns processos haverá a opção de expandir e, se clicar ali, verá que há uma série de subprocessos ocorrendo para que um único processo esteja ativado. Tal exemplo nos traz uma boa analogia para o funcionamento do processo de fixação de esboços, visto que, segundo o MME, a todo momento, parecem haver inúmeros processos fixados por tempos distintos, consumindo quantidades distintas de recursos e, nisso, gerando aquilo que experienciamos ao navegar pelo computador ou, melhor dizendo, aquilo que experienciamos ao navegar pelo mundo. A todo momento, então, vários esboços estão competindo pela fixação de longo período cortical, passando por constantes atualizações e sendo substituídos por novas representações quando não conseguem manter a coerência daquilo que estão propondo para o contexto espaço-temporal em questão.²⁷

Desenvolvendo aquilo anteriormente mencionado, o mecanismo responsável por delegar quais esboços fixados por tempo o suficiente alcançarão o nível pessoal da cognição é

²⁶ Digno de nota, esses processos de “fixação” podem ser localizáveis com certa precisão, tanto no espaço quanto no tempo (como possivelmente acontece em estudos da percepção com fMRI), mas o início dessas ativações não marcaria, necessariamente, o início da experiência consciente tal qual ela é experienciada pelo sujeito num determinado momento (DENNETT, 1991, p. 113).

²⁷ Conforme veremos mais tarde, podemos, aqui, supor que existam mecanismos que atuem como filtros estatísticos, os quais seriam calibrados com novas entradas sensoriais e realizariam uma espécie de checagem (“qual a probabilidade desse esboço ser coerente com a entrada sensorial?”), impedindo que o sujeito perceba anomalias nos objetos/cenários com os quais se depara.

aquilo que, no MME, é denominado sondagem. Embora Dennett não entre em muitos detalhes acerca de seus possíveis aspectos ontológicos e epistemológicos, em seu modelo, a sondagem parece ser a tão ansiada ponte que transfere o conteúdo perceptivo do nível sub-pessoal para o nível pessoal da cognição. Tal mecanismo consiste em uma espécie de “atenção cerebral”, que é atraída pelos esboços mais fortemente ativados nas redes neuronais e, então, auxilia em sua transmissão para o restante do sistema, fazendo com que eles fiquem ativados por um tempo mais longo e numa intensidade o suficiente para projetar seu conteúdo (e, com isso, sua percepção) ao nível pessoal. Dennett destaca que, com base nisso, o Modelo dos Múltiplos Esboços evita o erro de supor que há uma única narrativa – ou um “fluxo real” – para a consciência, pois, na verdade, múltiplos esboços e diversos ciclos de atualização ocorreriam a todo tempo e ao mesmo tempo, sendo em meio a esse “caos” de inter-influências que os fluxos narrativos que compõem nossa realidade manifesta emergiriam – de modo que nossa percepção consciente possa ser continuamente “atualizada” para lidar com um contexto dinâmico e em eterna mudança mas mantenha, ao mesmo tempo, coerência e sensação de continuidade (DENNETT, 1991, p. 113; 136).

Na medida em que, na escala de tempo cortical, ciclos de atualização de esboços vão se transpassando a nível sub-pessoal, em uma macro-escala, a percepção consciente também vai se atualizando. Aqui, embora esboços possam possuir conteúdos distintos das próprias entradas sensoriais, esses ciclos de atualização minimizam anomalias e a volatilidade daquilo por eles representado, causando a impressão pessoal (ilusória) de confiabilidade, estabilidade e continuidade dos cenários que experienciamos (Cf. ARATHORN et al., 2013). A experiência visual de um cômodo de nossa casa, por exemplo, é estável e contínua, e podemos atentar a diferentes detalhes desse cômodo, à medida em que movimentamos nosso corpo por ele. Entretanto, é um fenômeno bastante conhecido que, se ficarmos parados e olharmos fixamente para um ponto específico desse cômodo, tal ponto começará, com o tempo, a ficar instável ou a se movimentar. De acordo com o MME, nesse contexto, informações divergentes dos vários esboços ativados simplesmente não seriam “fomentadas” pela entrada sensorial e, portanto, esses simplesmente não seriam atualizados (ou seriam de maneira redundante). Aqui, a predominância de esboços em comparação com novas informações advindas das entradas sensoriais parece perturbar a dinâmica dos ciclos de atualização, gerando, portanto, a experiência visual de instabilidade.²⁸

²⁸ É válido salientar que, nesses contextos, a instabilidade de nossa experiência visual também advém da predominância da atualização de esboços referentes às demais modalidades de estimulação. Ainda, fora do

De acordo com Dennett, para além da estabilidade de nossa experiência perceptiva, a influência da dinâmica dos ciclos de atualização de esboços gera algo como uma “prontidão semântica”. Isto é, alguns esboços ativados podem, simplesmente, desaparecer logo em seguida, não deixando vestígio algum, já outros deixarão vestígios em relatos subsequentes de experiências, na memória, em estados emocionais, em tendências comportamentais, etc. Note que tais esboços não precisam sequer chegar ao nível pessoal da cognição (ou seja, adquirir “fama” e serem sondados), podendo permanecer ativos apenas em um nível referente à escala cortical e, mesmo assim, direta ou indiretamente instigar algum comportamento do sujeito. Por exemplo, ver uma foto de um cão pode criar um “conjunto perceptivo prévio”, tornando temporariamente mais fácil ver cães (ou apenas animais) em outros lugares e, ainda, ativar algum domínio semântico específico, tornando temporariamente mais provável que você leia a palavra manga como uma “fruta que o cão chupa” e não como uma parte de uma camisa (DENNETT, 1991, p. 135).

Diante dessa exposição, vemos que o MME consegue descrever, com certo nível de satisfação, o funcionamento dos estados locais da consciência, mas o modelo ainda possui lacunas notáveis naquilo que tange à exploração de certos aspectos desses estados e sua implicação para a percepção consciente em termos gerais e seus estados globais. Por exemplo, Dennett não deixa claro se a sondagem ocorre a nível pessoal ou a nível subpessoal ou se existem várias sondagens diferentes. Aqui, ainda que saibamos que, para o Ilusionismo, o nível pessoal é uma expressão manifesta do nível sub-pessoal, seria importante distinguir ou pelo menos esclarecer como o mecanismo de sondagem funcionaria em ambos (explicando como evitamos, por exemplo, a ocorrência de cenários paradoxais e catastróficos, tais como conscientemente escolher deixar de sondar informações captadas acerca da temperatura de nossos corpos e, portanto, ser inadvertidamente acometido por uma hipotermia). Outro ponto ao qual Dennett não parece ter dado devida atenção é a “natureza” das narrativas no MME. Uma crítica comum ao modelo dennettiano é a aparente necessidade da linguagem para a ocorrência de tais narrativas, ainda que Dennett não tenha explicitado isso em sua obra (Cf. DENNETT, 1991). Logo, o que seriam, mais especificamente, essas narrativas? Qual é a sua “natureza”? E, ainda mais importante, como poderíamos coerentemente diferenciá-las de meros “filmes”, projetados no Teatro Cartesiano? Finalmente, ainda que Dennett deixe

contexto proposto pelo experimento (ou caso não-induzida), essa ilusão de movimento ou instabilidade visual se denomina “oscilopsia”, estando, na maioria dos casos, relacionada com distúrbios ou danos neurológicos (Cf. TILIKETE, VIGHETTO, 2011).

implícito que as qualia compõem, de alguma forma, os esboços, juntamente com outros conteúdos informacionais, ele não descreve, em termos do MME, como e por que a perspectiva subjetiva emergiria desde múltiplos esboços – pois, mesmo que sejam ilusórias, as qualia fazem parte da consciência manifesta e, portanto, precisam ser devidamente tratadas e direcionadas dentro de uma descrição do fenômeno da consciência. Assim sendo, uma maneira de não apenas responder a essas perguntas, fornecendo (os devidos) esclarecimentos substanciais para o modelo, mas também melhor embasá-lo em aparatos científico-teóricos robustos seria “traduzi-lo” em termos de um *framework* atual, promissor e empiricamente bem-embasado: um posto que, nessa dissertação, será ocupado pelo Processamento Preditivo.

2. O Processamento Preditivo (PP)

O Processamento Preditivo é um *framework* teórico, cujo principal enunciado é o de que o funcionamento da cognição ocorre através da implementação e manutenção, em um sistema/organismo, de diversos modelos²⁹ de seu mundo e corpo. Em tese, o uso desses modelos é aquilo que permite a sobrevivência do sistema/organismo e a otimização de sua interação com o meio. Caso compreendido em um sentido amplo, o PP parece fomentar, então, a consolidação de um novo paradigma para as ciências comportamentais e do cérebro, fundamentado em sua suposta novidade e poder explicativo. Em seu emprego atual, ele é comumente utilizado para abordar tanto processos sub-pessoais (neste caso, físico-computacionais e neurobiológicos) quanto pessoais (isto é, as estruturas de racionalização e relato agenciadas por nossa *folk psychology*), apresentando-se como um promissor campo aberto de pesquisa (Cf. WIESE, METZINGER 2017). Com isso em mente, na Seção 2.1 do presente capítulo, veremos algumas fundamentações mais gerais, comuns às várias ramificações do PP e na Seção 2.2 algumas perspectivas que buscam viabilizar o estudo da consciência internamente ao *framework* serão desenvolvidas, com base majoritariamente no trabalho de Jakob Hohwy e Andy Clark. Digno de nota, o formalismo matemático e a conceitualização técnica que comumente se vêem relacionados à literatura sobre o Processamento Preditivo não serão desenvolvidos aqui. Isto é, ainda que (em parte) eventualmente mencionados e referenciados, tais aspectos não serão explorados ou detalhados, uma vez que o objetivo do presente capítulo é meramente fornecer uma introdução e análise filosófica dos contornos gerais do Processamento Preditivo.

2.1 Fundamentos Gerais

2.1.1 Contextualização histórico-filosófica

Uma das formas mais comuns, na literatura sobre o PP, de localizá-lo internamente a uma corrente histórico-filosófica relativamente clara inicia-se com uma interpretação específica da Alegoria da Caverna de Platão (370 a.C.). Pois, embora essa espécie de experimento mental se refira, canonicamente, ao contexto da vivência humana em sociedade, dela, faz-se possível extrair, também, alguns pontos interessantes acerca de nossa percepção.

²⁹ A literatura sobre o PP não parece distinguir entre alegações de que sistemas cognitivos implementariam um único modelo ou, na realidade, diversos modelos. De qualquer forma, como é perfeitamente plausível tratar parâmetros internos a um modelo geral como sendo, eles mesmos, modelos mais específicos, na presente dissertação também não nos preocuparemos com tal distinção.

Desses, talvez o principal seja dizer que a “natureza” de nossa percepção consciente seria como as sombras projetadas na caverna – implicando que ela seria o reflexo indireto de causas (ocultas) que nunca poderemos diretamente observar e encontrar no mundo (Cf. SETH, 2021). Seguindo uma linha de raciocínio parecida, séculos depois, nos sete volumes de seu “Perspectiva” (كتاب المناظر), o polímata árabe Alhazen (965-1040) introduziu uma série de contribuições para o entendimento da percepção visual, negando que ela pudesse fornecer um acesso direto ao mundo e explorando empiricamente como, por sua vez, ela dependeria de mecanismos como o julgamento e a inferência (AL-HAYTHAM, 1030/1989).

Já no século XVIII, a literatura sobre o tema destaca a atuação de dois filósofos cujas ideias culminariam em temáticas-chave para a filosofia do Processamento Preditivo: o escocês David Hume (1711-1776) e o alemão Immanuel Kant (1724-1804). Conforme foi notado por Hohwy (2013) e Dennett (2013), no *Tratado da Natureza Humana*, Hume desenvolve uma problemática acerca da percepção que gira em torno de nossa tendência (hábito) de estabelecer relações causais no mundo (HUME, 1739; 2000). Especificamente, para o autor, se notarmos várias vezes que B sucede A, tendemos a estabelecer uma relação causal entre A e B, sem que, necessariamente, exista uma relação como essa a ser detectada. Ou seja, mesmo que uma relação causal possa ser convencionalmente estabelecida em nosso dia-a-dia, não haveria como justificá-la em termos racionais. Curiosamente, então, aquilo que fazemos ao percebermos uma causação parece ser, na verdade, a percepção de uma espécie de sentimento interior (ou uma “imaginação”) que, de acordo com nossas vivências, “projetamos” no mundo (DENNETT, 2013, p. 355). Segundo a análise de Jason Flores (2015), Hume desenvolveu sua linha de raciocínio de maneira a concluir que todo conhecimento perceptivo poderia ser reduzido a relações probabilísticas e, portanto, tratar-se-ia de algo cuja “natureza” é a mesma das evidências (interpretativas) que utilizamos para fazer sentido de nosso mundo (FLORES 2015, p. 3) – uma observação que, internamente ao PP, foi adaptada para um vocabulário da estatística bayesiana³⁰ e explorada por Hohwy (2013) em sua análise da percepção.³¹ Ainda, de acordo com Link Swanson (2016), esse mesmo problema (da origem da estrutura causal em nossa percepção), parece ter instigado Kant a desenvolver, em sua *Crítica da Razão Pura*

³⁰ Em referência ao Teorema de Bayes, uma equação fundamental para a estatística probabilística: $P(A|B) = [P(B|A)P(A)]/P(B)$, onde P = probabilidade, H = hipótese e E = evidência. Tal teorema incorpora a possibilidade de eventos condicionais influenciarem o resultado-alvo, com tais eventos correspondendo à pluralidade de fatores pertinentes para a ocorrência daquilo previsto. Em outras palavras, ele auxilia na predição de resultados em contextos nos quais há certo grau de incerteza, descrevendo a probabilidade de um evento ocorrer ou ter ocorrido, com base no conhecimento prévio das condições possivelmente relacionadas a esse mesmo evento.

³¹ Para uma análise mais profunda das várias influências da obra Hume sobre o Processamento Preditivo, Cf. FLORES, 2015.

(1781; 1998), algumas ideias que, atualmente, podem ser consideradas como fundacionais para a filosofia do Processamento Preditivo, tais como a maneira através da qual o conteúdo perceptivo seria mentalmente construído, o papel fundamental das informações *a priori* nesse processo e, principalmente, a função daquilo que hoje se convém “denominar” modelos mentais na cognição (Cf. SWANSON, 2016).

Ainda de acordo com essa narrativa, no final do século XIX, o cientista alemão Hermann von Helmholtz (1821-1894) se inspirou no trabalho de Kant para desenvolver a perspectiva do cérebro como um testador de hipóteses e da percepção como um processo de inferência inconsciente, envolvendo inferências probabilísticas (Cf. SANTOS, VERÍSSIMO, 2018). Mais especificamente, para ele, a percepção seria um processo inferencial de combinação de sinais sensoriais com “crenças”/expectativas que o cérebro possuiria acerca das causas desses sinais. Tais inferências ocorreriam fora da percepção do sujeito – sendo, portanto, inconscientes – e acompanhariam as causas no mundo ao atualizarem nossas melhores suposições perceptivas (hipóteses) à medida em que novos sinais sensoriais fossem chegando (Cf. HELMHOLTZ, 1860). Graças a tal desenvolvimento, Helmholtz se tornou a principal inspiração para inúmeros programas de pesquisa na psicologia, neurociência, ciência da computação, aprendizado de máquina, etc, dos quais o PP pode ser considerado um “herdeiro intelectual”.

Já no século XX, outras grandes influências para a filosofia do Processamento Preditivo comumente elencadas pela literatura foram a Teoria da Informação, proposta por Claude Shannon (1916-2001) em 1948, a qual possui a premissa de que a informação seria inseparável da falta de previsibilidade (incerteza) – ou seja, se, antes mesmo de ser observado, um evento é plenamente previsível, ele não é um transmissor significativo de informação; e o Princípio da Redundância Mínima, proposto por Horace Barlow (1921-2020) em 1961, para o qual a maximização da taxa de transferência de informação só pode ocorrer se um evento for minimamente previsível e, portanto, minimamente redundante (MILLIDGE, SETH, BUCKLEY, 2022, p. 4).³² Ainda, desse mesmo período, outra inspiração significativa está no connexionismo – particularmente naquilo que diz respeito à construção de redes neurais com

³² Outra forte influência na construção da filosofia do Processamento Preditivo foi a cibernética de William Ashby (1903-1974), principalmente no que diz respeito à função exercida por modelos internos e as implicações do Teorema do Bom Regulador.

algoritmos de retropropagação para a correção de erros de classificação nos dados recebidos (Cf. RUMELHART et al., 1986).³³

Recentes trabalhos filosóficos também contribuíram consideravelmente para a expansão do quadro teórico do Processamento Preditivo. Por exemplo, em 2007, o psicólogo Chris Frith introduziu e discorreu sobre um uso particular de conceitos bayesianos no entendimento do cérebro, elaborando algumas das possíveis aplicações empíricas desse esquema em seu livro *Making up the Mind* (Cf. FRITH, 2007). Já o filósofo Jakob Hohwy tem se dedicado à exploração detalhada de um dos principais mecanismos do PP, denominado Minimização de Erro de Predição, que consiste no principal proponente, internamente ao *framework*, da tese da hipótese vencedora para a percepção consciente (a qual será desenvolvida ao final deste capítulo). Em 2013, ele publicou uma influente monografia sobre o assunto: *The Predictive Mind*, na qual defende uma resolução para o chamado “Problema da Percepção” (cuja inspiração é o supramencionado Problema da Causação de Hume) valendo-se da Inferência Bayesiana e do Processamento Preditivo. Mais tarde, em seu livro *Surfing Uncertainty* de 2016, o filósofo Andy Clark desenvolveu uma ramificação (atualmente) bastante popular do PP chamada Processamento Preditivo Orientado à Ação, na qual o papel fundamental da ação em nossa percepção corporificada é colocado em destaque. Sua obra ainda introduziu uma diversidade de conexões interessantes entre os trabalhos recentes de neurocientistas computacionais, pesquisadores em Inteligência Artificial e a sua eclética filosofia das ciências cognitivas. Finalmente, em 2021, o neurocientista Anil Seth realizou uma preciosa contribuição para a área com seu livro *Being You*, onde – contrariando as afirmações de Descartes – desenvolveu a ideia, já anteriormente introduzida, de que somos “animais-máquinas” cuja percepção é uma alucinação controlada.

Já naquilo que diz respeito, mais especificamente, aos aspectos empíricos do Processamento Preditivo (que mais comumente são adotados desde os desenvolvimentos originalmente realizados em torno de um *framework* semelhante, denominado Codificação Preditiva³⁴), Helmholtz é comumente considerado o autor-chave que possibilitou muitos dos

³³ Entretanto, o PP se difere profundamente dos algoritmos de retropropagação por não ser supervisionado e, portanto, não precisar de uma alimentação ascendente de dados para o seu treinamento. No *framework*, o mecanismo de Minimização de Erro de Predição utiliza modelos *top-down*, contrastando com a retropropagação, a qual classifica os dados de forma *bottom-up* (Cf. HINTON, 2007).

³⁴ O *insight* central da codificação preditiva é que a totalidade dos dados não precisa ser processada, mas apenas a diferença entre a expectativa de uma sequência temporal de eventos ou estímulos e sua presença real. Portanto, apenas variações não-esperadas são incluídas no código compactado e são somente essas variações, entre o previsto e o real, que, em última instância, seriam transmitidas – economizando, com isso, espaço e tempo de processamento (Cf. ATAL, 2006).

avanços subsequentes da área. Sua inspiração foi fundamental, por exemplo, para psicólogos e cientistas do comportamento, tais como Ulrich Neisser (1967), Richard Gregory (1980) e Irvin Rock (1983), trabalhando na elaboração de teorias e modelos do tipo “*analysis by synthesis*”, os quais, geralmente, estão associados ao estudo da percepção da fala e processamento de linguagem.³⁵ Também de vital importância foi o trabalho de Mandyam Srinivasan e colegas que, em 1982, foi responsável por concretizar o primeiro modelo teórico de codificação preditiva na retina humana (Cf. SRINIVASAN et al., 1982). Dez anos depois, David Mumford estendeu esse mesmo modelo, desde as propriedades da retina, transformando-o numa teoria geral de nossa função cortical (Cf. MUMFORD, 1992). Não obstante, mesmo introduzindo aspectos fundamentais da codificação preditiva ao funcionamento córtex de maneira formalmente robusta, sua teoria não gozou de grande popularidade, principalmente por não ser sustentada por simulações realistas ou experimentações empíricas subsequentes. Finalmente, em 1999, Rajesh Rao e Dana Ballard publicaram aquilo que é considerado um dos mais importantes artigos da área (talvez até mesmo o mais importante), no qual os autores desenvolveram as ideias propostas por Mumford e construíram uma pequena rede biologicamente realista de codificação preditiva para investigar empiricamente seu comportamento numa simulação do córtex visual humano. Graças a tal estudo, foi possível notar como uma interação complexa e dinâmica entre previsões e erros de previsão poderia explicar diversos fenômenos neurofisiológicos – tanto aqueles já anteriormente elucidados pelas neurociências como diversos outros que, na época, pareciam estar fora do escopo explicativo de tais disciplinas (Cf. RAO, BALLARD, 1999) – algo que, atualmente, está no cerne daquilo proposto pelo PP.

Não obstante, pode-se dizer que foi somente na década de 2000, com os trabalhos cruciais do neurocientista Karl Friston (2003, 2005, 2008), que o Processamento Preditivo ganhou o impulso decisivo para sua extraordinária expansão pelas ciências cognitivas contemporâneas. De certa forma, aquilo que Friston fez foi demonstrar que os resultados de Rao e Ballard poderiam ser trabalhados conforme uma inferência bayesiana aproximada, processada com base em modelos gerativos³⁶, o que aproximaria, portanto, a ideia geral de

³⁵ Em poucos termos, trata-se de uma abordagem que consiste em analisar as propriedades de um estímulo para, depois, determinar as informações significativas advindas dessa análise e reuni-las (sintetizá-las) em uma representação ou interpretação interna daquilo que o estímulo poderia ser – isto é, similarmente ao PP, num processamento orientado aos dados e guiado por informações previamente adquiridas. Cf. <https://dictionary.apa.org/analysis-by-synthesis>

³⁶ Geralmente, modelos tais como os modelos gerativos são entendidos, estatisticamente, como construtos matemáticos que capturam as relações entre os valores de um conjunto de variáveis aleatórias (PIEKARSKI, 2021 p. 7). Conforme veremos, no contexto do PP, eles também podem ser interpretados como representações

Mumford do Princípio da Redundância Mínima de Barlow e da noção helmholtziana de percepção como inferência inconsciente – caracterizando-a matematicamente e demonstrando que esquemas de codificação preditiva são eficientes na resolução de problemas de otimização, tanto nas neurociências quanto no aprendizado de máquina, entre outros (Cf. CLARK, 2016; FRIEDMAN, SØVIK, 2021).

Por sua vez, na filosofia contemporânea, o Processamento Preditivo é predominantemente trabalhado como um novo paradigma para a cognição humana. Segundo Hohwy (2013, p. 8), um dos motivos para tanto é o fato de que o PP aparentemente se posiciona entre as duas principais orientações atuais para as ciências cognitivas, visto que, conforme veremos, ele faz e permite o uso tanto de conceitos oriundos de abordagens de orientação *top-down* quanto *bottom-up* (Cf. GRIFFITHS et al., 2010; MCCLELLAND et al., 2010). Ainda, segundo Mark Sprevak, outra motivação para tal trato é o fato de que o Processamento Preditivo parece conseguir explicar, de forma satisfatória, a cognição humana de acordo com os três níveis de descrição de Marr (SPREVAK, 2022, p. 2): o computacional, o algorítmico e o implementacional. No nível computacional, o PP enunciaria que o objetivo da cognição é minimizar o erro de predição sensorial. No nível algorítmico, ele sugeriria que o algoritmo utilizado por nossos cérebros corporificados para resolver as tarefas cognitivas envolveria a operação de uma rede neural hierárquica de unidades de predição de erro – a qual pode ser entendida como a responsável pela execução de inferências bayesianas aproximadas. Finalmente, no nível implementacional, o Processamento Preditivo indicaria as mais prováveis candidatas à estrutura de instanciamento físico desse algoritmo, como, por exemplo, áreas anatomicamente distintas no neocórtex (as quais implementariam camadas funcionalmente distintas da rede hierárquica) e populações de células anatomicamente distintas dentro de cada uma dessas áreas (as quais implementariam as unidades de predição e de erro). Finalmente, naquilo que diz respeito à temática da presente dissertação, o Processamento Preditivo parece ser capaz de explicar uma parcela significativa das dinâmicas dos estados globais e locais da consciência (Cf. SETH, 2021), os quais foram eleitos no capítulo anterior como objetivos ideias para análises empiricamente embasadas acerca do fenômeno.

mentais do ambiente – embora alguns autores problematizem essa descrição, questionando se ela indicaria que tais concepções seriam permutáveis ou meramente comparáveis (Cf. WILLIAMS, 2018).

2.1.2 O Cérebro Preditivo

A abordagem tradicional do processamento perceptivo orientado à informação descreve que, para perceber um cenário, nossos sensores escaneiam e coletam informações/*inputs* para serem processados pelo cérebro de maneira a formular uma saída/*output* (seja interna ou externa) contextualmente adequada. Dessa forma, o papel do cérebro seria receber informações acerca do ambiente e juntá-las internamente para, assim, representá-lo tal como o experienciamos (Cf. MARR, 1982; GAZZANIGA, 2019, p.70). Na psicologia, isso ficou conhecido como percepção de processamento *bottom-up*, designando uma corrente teórica relativamente comum entre os estudiosos das ciências cognitivas e da percepção. E, embora aqui o papel do processamento *top-down* não seja negado, sua função estaria relacionada apenas com alguns processos complexos e abstratos, tais como a imaginação, o raciocínio e a memória, deixando a captação do conteúdo perceptivo a ser interpretada como um processo ascendente basilar, funcionando de forma relativamente passiva e demandando um gasto energético considerável.

O Processamento Preditivo oferece um esquema um pouco mais sofisticado para o entendimento da percepção. Pois, embora os cérebros por ele descritos trabalhem tanto de maneira *top-down* quanto *bottom-up*, seu diferencial é a característica predominantemente preditiva desse processamento, visto que, enquanto o modelo tradicional parece requerer a captação exaustiva de informações para o processamento de um cenário, no PP, os sensores captam apenas pistas do meio e o cérebro faz o uso massivo de modelos³⁷ preditivos para assimilá-las. De acordo com o *framework*, esses modelos estão constantemente ativados e vão melhorando ao passo que os sensores captam novas pistas do mundo. Por exemplo, ao abrirmos os olhos quando despertarmos, temos acesso à informação de que estamos em nossos quartos e qual é sua configuração. Mas, antes disso, possuímos informações de que, no dia anterior, dormimos nesse mesmo quarto e que ele possuía essa mesma configuração – e, portanto, possuímos fortes expectativas de que iremos acordar em tal ambiente. Caso não acordemos em nossos quartos ou caso, ao despertarmos, eles estejam completamente diferente do que nos lembramos, seremos acometidos por uma grande surpresa e, certamente,

³⁷ Uma importante observação acerca de modelos é que nesta dissertação, a natureza dos modelos não foi especificada, mas o uso de um “modelo gerativo do processo realizado pelo sistema” e “modelos gerativos dos objetos e cenários percebidos” foi intercalado nas descrições do Processamento Preditivo, visando um entendimento geral de seu funcionamento. Para uma compreensão robusta acerca dos modelos gerativos e sua natureza, Cf. PARR, PEZZULO, FRISTON, 2022.

passaremos a nos atentar energeticamente ao local no qual estamos, além de ativamente criar hipóteses sobre como fomos parar ali ou como sua configuração mudou durante a noite.

Em outras palavras, a percepção preditiva faz o uso de informações que o organismo já possui acerca do meio para criar expectativas sobre o quão similares essas informações serão àquilo que os sensores irão captar subseqüentemente. E, na medida que essas expectativas “condizem” com o que for captado, o cérebro mantém seu modelo e evita gastos de processamento. Entretanto, nossos modelos de mundo não podem ser, o tempo todo, idênticos às informações captadas, pois nossos mundos são dinâmicos e sempre haverá informação nova a ser assimilada. Dessa maneira, aquilo que divergir do esperado gerará surpresa e irá atrair a atenção do organismo, para que ele apreenda essa nova informação e atualize seu modelo de mundo. Portanto, nossos cérebros não captam inteira e constantemente as informações vindas de seu mundo, mas apenas pistas que “divergem de suas expectativas” e lhes permitem inferencialmente assimilá-lo (PARR, PEZZULO, FRISTON, 2022, p. 16).

Logo, passamos a entender a percepção como uma inferência que alcança uma conclusão local ou global desde um modelo do mundo responsável por gerar as ativações nos nossos canais sensoriais. Por esse motivo e sob essa perspectiva, então, o cérebro é mais propriamente entendido como uma máquina de inferências bayesianas ou, simplesmente, um cérebro preditivo. Mais especificamente, entende-se que o cálculo de tais inferências seria, na realidade, um problema intratável para o cérebro, visto que, para tanto, ele precisaria lidar a todo instante com um número praticamente infinito de variáveis possíveis, o que geraria um gasto energético simplesmente incoerente com a realidade. Portanto, aquilo que os proponentes do PP sustentam é que nossos cérebros preditivos se valeriam, na realidade, de uma aproximação da regra de Bayes (PARR, PEZZULO, FRISTON, 2022, p. 16; 27), com tal cálculo ocorrendo para a manutenção de um modelo gerativo estruturado em suas próprias dinâmicas de ativação – o qual consistiria numa probabilidade conjunta atribuída àquilo “observado” por nós, predições acerca de estados ocultos/latentes³⁸ de nosso mundo interno e externo e expectativas sobre como nossas “observações” seriam geradas a partir das informações captadas por nossos sensores (WIESE, METZINGER, 2017, p.14). E, conforme será desenvolvido ao longo do presente capítulo, para que tal modelagem preditiva funcione adequadamente, esse esquema de entendimento do comportamento dos organismos em seu meio parte de três pilares centrais: a necessidade de agência, o papel de um conhecimento

³⁸ Os estados ocultos ou latentes são aqueles que não podem ser diretamente observados, mas apenas inferidos desde a estimulação sensorial (PARR, PEZZULO, FRISTON, 2022, p. 16).

prévio acerca do ambiente e da forma através da qual o ambiente é captado pelo organismo (HARKNESS, KESHAHA, 2017, p. 2).

Nisso, ciclos de atualização de tal modelo gerativo promoveriam a constante aprendizagem de novas informações, as quais seriam utilizadas na confecção de “prévias” que tendem a enviesar e padronizar nossa interação com o mundo (DOŁĘGA, DEWHURST, 2021, p. 3; Cf. CLARK, 2016). Conseqüentemente, esses vieses e padrões geram expectativas (no organismo) acerca da captação de informações que estejam de acordo com aquilo que o organismo já possui (ou aprendeu) além de promoverem o aprendizado daquilo que, em contextos familiares, é informacionalmente relevante – e deve ser processado – e o que pode ser descartado como irrelevante ou ruído. Em termos gerais, no PP, essas expectativas são chamadas de “predições”, visto que, partindo daquilo mantido pelo modelo gerativo, estimam um estado do organismo e seu mundo em um dado momento ou, mais precisamente, estabelecem uma estimativa das informações que o organismo ainda não observou. Logo, aqui, o termo “predição” não se refere necessariamente ao futuro, mas às informações não-observadas (SPREVAK, 2021, p. 9).

Tais predições podem ser rápidas e de “baixo nível”, descrevendo, por exemplo, as formas de um objeto, ou lentas e de alto nível para prever, por exemplo, o comportamento de um organismo complexo. Mais especificamente, nessa história, diferentes níveis de abstração caracterizam nossas predições e, quanto mais abstrata elas forem, mais elas serão parametrizadas por crenças³⁹, informações prévias e as regularidades do meio – tais como os padrões nas estruturas do ambiente, no comportamento dos outros agentes, etc (PIEKARSKI, 2021, p. 6). Logo, de acordo com *framework* do Processamento Preditivo, só seria possível lidar com tal dinâmica preditiva por meio de uma arquitetura hierárquica, composta por múltiplas camadas cerebralmente instanciadas. Aqui, cada camada preditiva possuiria um nível de complexidade diferente e trabalharia em uma escala espaço-temporal diferente, sendo que as camadas mais baixas seriam responsáveis por predições e regularidades rápidas e basilares, ao passo que as de nível mais alto seriam responsáveis por predições e regularidades lentas e complexas (HOHWY, 2013, p. 28; WIESE, METZINGER, 2017, p. 6). Ainda, pode-se dizer que, em cada camada, existiriam duas unidades funcionais: (1) unidades representacionais que codificam a melhor crença para os eventos observados (dadas as

³⁹ Aqui, o termo “crença” pode ser entendido como uma distribuição de probabilidades, relativa aos estados das coisas ou propriedades desconhecidas em nosso mundo. Ou seja, trata-se de algo como uma crença prévia e estatística do nosso sistema, com um alto grau de abstração, relacionada ao conhecimento geral que possuímos sobre o mundo (Piekarski 2021, p. 9).

informações anteriormente captadas desde um mesmo contexto ou contextos similares), as predições das camadas superiores e as informações sensoriais oriundas das camadas inferiores; e (2) unidades de erro, ativadas por informações sensoriais que divergem das predições computadas pelas camadas superiores, enviando tal informação às mesmas (Cf. FRISTON, 2005, 2010).

Nota-se, então, que, nesse tipo de arquitetura, a comunicação ocorreria via dois tipos de fluxo bidirecional: intra-camada e inter-camadas. O primeiro diz respeito à troca de informações entre as unidades representacionais e as unidades de erro dentro de uma mesma camada, enquanto o segundo se refere à troca de informações entre camadas. Nisso, fluxos de orientação descendente enviariam as predições para as camadas inferiores, de modo a inibir ou dissolver (*explain away*) as redundância propagadas pelos fluxos ascendentes, vindos das camadas inferiores, garantindo, assim, que apenas erros residuais de predição sejam propagados para as camadas superiores, via fluxo ascendente (CLARK, 2016, p. 37-39). Dessa forma, pode-se dizer que, dentro do esquema estabelecido pelo PP, o papel funcional dos erros de predição é atuar como um *feedback* para nosso modelo interno – o que garante que ele seja supervisionado pelo próprio mundo⁴⁰ (FRISTON, 2005, p. 825; HOHWY, 2016, p. 47). Aquilo que se observa, então, é que nossos cérebros preditivos precisam constantemente buscar um equilíbrio entre tentar inibir a proliferação de erros e melhorar suas predições.⁴¹

Digno de nota, esse ponto é interessante, pois é sempre válido ressaltar a importância que o PP atribui à existência de erros de predição. Afinal, se quisermos assimilar as informações vindas do mundo, elas não poderão ser completamente inibidas pelos modelos preditivos e a única forma disso não ocorrer é se houver “novidade” nelas. Essa novidade é aquilo que será computado como erro de predição e, de maneira a atualizar o modelo gerativo e aprender tal novidade, esse erro precisa ser minimizado – mas não extinguido – e devidamente processado por diversas camadas preditivas. Caso a atualização do modelo não venha a ocorrer por conta da ausência de erros em suas predições, torna-se difícil para o organismo manter sua flexibilidade, visto que, se alguma informação radicalmente distinta daquilo esperado for eventualmente captada, ele não estará preparado para devidamente recebê-la e lidar com ela. Caso essa história esteja correta, então, ocorrências do tipo poderiam

⁴⁰ Essa ideia acaba por excluir qualquer necessidade de intermediações corticais ou “misteriosos supervisores”, tais como homúnculos ou duplas transduções (HOHWY, 2013, p. 49).

⁴¹ Naquilo que tange a esse equilíbrio, podemos, novamente, resgatar as influências de Barlow e Shannon. Afinal, mesmo que a informação seja inseparável da incerteza, se quisermos maximizar a transferência dessa informação (ou, no caso, seu aprendizado), então ela deve ser minimamente previsível e, portanto, minimamente redundante (MILLIDGE, SETH, BUCKLEY, 2022, p. 4).

gerar modelos cada vez menos próximos da realidade e o sujeito acabaria, por exemplo, sofrendo com alucinações (caso acordado). Por outro lado, se os sensores do organismo captam um número grande de informações não-devidamente “dissolvidas” em seu modelo preditivo, seu processamento pode gerar um esgotamento cognitivo e o sujeito poderá ser acometido por algum tipo de crise ou colapso (Cf. HOHWY, 2013). No Processamento Preditivo, o mecanismo responsável por garantir esse, tão vital, equilíbrio é a Minimização de Erro de Predição, a qual iremos detalhar a seguir.

2.1.3 Detalhamentos da Minimização de Erro de Predição

Sempre que divergências entre as predições e a informação sensorial captada forem detectadas, a Minimização de Erro de Predição irá sinalizá-las como *surprisal*⁴². Na literatura do PP existem diferentes significados para esse termo, mas aquele mais próximo do que será tratado nesta dissertação consiste na quantificação da diferença entre duas probabilidades ou, naquilo que nos interessa, uma medida de quanto nosso modelo precisa atualizar suas crenças após uma certa observação (Cf. TRIBUS, 1961; PARR, PEZZULO, FRISTON, 2022, p. 18-20). Assim, a *surprisal* será grande se, dado aquilo contido no modelo, a probabilidade dos dados sensoriais eventualmente “observados” for baixa. Similarmente, dado o modelo de mundo do organismo, a minimização dessa *surprisal* implicaria a maximização dessa probabilidade (FRISTON, 2010, p.128; WILLIAMS, 2022, p.12).

Com isso em mente, a Minimização de Erro de Predição possuiria uma dupla-função, explicitamente minimizando os erros de predição e, também, implicitamente minimizando a *surprisal*, de maneira a otimizar o funcionamento do organismo desde seus níveis mais basilares até os mais complexos. Na realidade, os proponentes do PP tendem a, até mesmo, considerá-la como o mecanismo (único e geral) para o qual a totalidade da atividade cerebral é orientada (Cf. HOHWY, 2015; CLARK, 2016; FRISTON, 2018). Para tanto, o funcionamento da MEP ocorreria de duas formas diferentes, porém relacionadas: (1) através de inferências perceptivas, as quais consistiriam na mera na mudança/atualização do modelo interno do organismo; ou (2) através de inferências ativas que consistem na manipulação proativa do

⁴² Existe uma diferença entre os desalinhamentos básicos dos dados sensoriais iminentes, conectando o ambiente ao que é previsto pelo cérebro em nível neuronal (*surprisal*), e o fenômeno qualitativamente acessível ao nível do agente conhecido como “surpresa”. Aqui tende-se a distinguir claramente esses dois fenômenos, operando em níveis dissimilares no cérebro. A terminologia utilizada para designar a natureza implausível (em um sentido bayesiano) de um sinal sensorial inesperado é “*surprisal*” – um conceito originalmente cunhado pelo teórico Myron Tribus (1921-2016) – de maneira a distingui-la do sentido (*folk*) de surpresa, o qual é familiar a todos nós (Cf. TRIBUS, 1961).

ambiente, de forma que o organismo minimize o nível de incertezas/erros de predição ao trazer o contexto para mais próximo de seu modelo gerativo (HOHWY, 2013, p. 42-43). Quanto menor for a previsibilidade do ambiente/estado no qual o organismo se encontra, mais seu cérebro orientará seus recursos e atenção para a captação e processamento de informações vindas do ambiente. Como vimos, essas informações passam por uma espécie de filtro que separa aquilo que é interessante (para o organismo) – ou seja, o erro de predição – do ruído ou informação desnecessária. Para que essa filtragem possa acontecer, a MEP faz o uso de pelo menos dois critérios de confiabilidade: a acurácia, que estabelece a precisão/exatidão das predições e a ponderação da precisão dos erros de predição, a qual avalia sua “confiabilidade”.

A acurácia consiste na qualidade de uma predição em explicar as causas dos efeitos absorvidos por nossos sensores (FRISTON, 2010, p. 3). Isso implica que, para cada predição de nosso modelo, existe uma expectativa de *surprisal* e, quanto maior ela for, menos precisa/exata é tal predição – e vice-versa. Predições com alta acurácia tendem a gerar menos erros preditivos e são melhores para representar a estrutura causal do mundo, utilizando a menor quantidade possível de recursos cognitivos, uma vez que não precisam ser integralmente atualizadas com frequência. Por sua vez, predições com baixa acurácia induzem a MEP a pressupor uma ponderação do grau de precisão, ou seja, um ajuste da extensão na qual as entradas sensoriais são moduladas por predições descendentes – a fim de “alimentar o modelo” com informações sensoriais mais confiáveis, ou seja, mais precisas (*Ibid.*). Quando isso ocorre, o cérebro preditivo precisa garantir que as novas informações possuirão um grau de precisão satisfatório para eficientemente atualizar o modelo gerativo (PIEKARSKI, 2021, p. 8). No PP, isso se torna possível porque, com o acúmulo de vivências, o modelo interno do organismo se torna capaz de gerar uma espécie de distribuição probabilística de erros de predição, aprendida desde as regularidades do mundo. E, se esses padrões de erros forem devidamente aprendidos, eles podem ser utilizados para prever os níveis de ruído em uma amostragem advinda de um contexto igual ou similar e, assim, atuar como informações prévias que orientarão subsequentes absorções de erros de predição (HOHWY, 2013, p. 194). Logo, para atualizar seu modelo de mundo, o cérebro busca minimizar erros em suas predições e maximizar sua precisão, graças a um mecanismo de ponderação que, dada a existência de ocorrências anteriores, elege os erros de predição mais precisos, e a um mecanismo de acurácia que, dado o sucesso preditivo em ocasiões anteriores, elege as melhores dentre as predições disponíveis.

Em termos cognitivo-comportamentais, a ponderação da precisão pode ser interpretada como a atuação da atenção (CLARK, 2016, p. 59-60; Cf. FRISTON, 2009). Quando o organismo precisa direcionar sua atenção, isso significa que sua percepção está se baseando majoritariamente em informações transmitidas ascendentemente na hierarquia preditiva, pois, nesses casos, a absorção de informação com os erros de predição é maior e, portanto, qualquer coisa inesperada tende a ser percebida. Por outro lado, quando estamos desatentos ou nosso nível de atenção é baixo, nossa percepção se baseia majoritariamente em predições descendentes e, portanto, apenas determinadas informações (já esperadas) acerca dos objetos ao nosso redor tendem a ser percebidas (PIEKARSKI, 2021, p. 8). Dada essa descrição, o cérebro preditivo parece ter que lidar com pelo menos dois tipos de atenção (FELDMAN, FRISTON, 2010, p. 2). O primeiro, denominado “atenção endógena”, consiste em processos de controle volitivo que impactam de forma descendente a influência/ganho dos erros de predição associados a características relevantes de uma certa tarefa ou localização específica. Como essas características não indicam, por si mesmas, a localização de um alvo específico, a atenção seria então direcionada para o local correto, de acordo com as informações coletadas desde o ambiente.⁴³ Já a “atenção exógena” consiste em processos que aumentam, de forma ascendente, a influência/ganho dos erros de predição amostrados durante o desempenho de uma tarefa devidamente aprendida ou em resposta a um estímulo novo/saliente no ambiente, que seja implicitamente esperado pelo organismo (FELDMAN, FRISTON, 2010, p.2; CLARK, 2016, p. 77).⁴⁴ Ainda, da mesma forma que a captação de informações relevantes é constante mesmo na presença de objetos e contextos esperados, a captação de sinais salientes também tenderia a ocorrer em ambientes ruidosos, de modo que a ponderação de precisão atue de maneira semelhante, tanto na atenção exógena quanto na endógena. Disso resulta que, de acordo com o PP a atuação da atenção na percepção implica que primeiro construímos uma essência geral do ambiente no qual os encontramos – desde os componentes já esperados e deduzidos por nossas predições – para, apenas depois, nos focarmos em quaisquer

⁴³ Um exemplo clássico de atenção endógena consiste nas famosas imagens da série de livros “Onde está Wally?”. Quando procuramos por Wally em cenários complexos, completamente lotados de pessoas e coisas, estamos orientando nossa atenção por meio de um modelo preditivo com as características de Wally, no intuito de salientá-lo em tal cenário.

⁴⁴ Um exemplo de atenção exógena consiste no impacto das notificações de celular. Quando estamos realizando uma tarefa e recebemos tal notificação, imediatamente nos atentamos a ela, mesmo que, em alguns casos, não seja necessário parar o que estamos fazendo para verificá-la.

detalhamentos desses ambientes, numa espécie de cognição “floresta primeiro, árvores depois” (FRISTON, 2005, p. 825).⁴⁵

Em suma, a percepção eficiente parece ocorrer porque o cérebro preditivo consegue manter um equilíbrio constante entre a acurácia de suas predições e a precisão de erros vindos do ambiente, ponderando os detalhamentos necessários e excluindo quaisquer ruídos desinteressantes. Nessa história, a atenção é uma peça importantíssima porque é graças a ela que o organismo consegue aumentar o ganho de informação com os erros de predição e, portanto, otimizar sua conexão com o mundo. Além disso, ela é a chave para entendermos como tudo isso se liga a um outro ponto-chave para o Processamento Preditivo: o papel da ação.

2.1.4 Ação

Uma vez que a atenção é acionada, os sensores permitem a transmissão ascendente das informações sobre o contexto ou evento-alvo. Para que isso aconteça, o organismo precisa se ajustar e alterar a disposição de seus sensores no (e sobre o) mundo, algo que ocorre por meio da ação. Digno de nota, essa ação não precisa ser um comportamento complexo e intencional, podendo se tratar meramente de atuações basilares, tais como os movimentos sacádicos dos olhos, o aumento/diminuição da pupila, a leve inclinação da cabeça, o direcionamento dos ouvidos, etc. Ainda, ela pode ser expressa por meio de movimentações endógenas relacionados a atuações internas, tais como alterações na atividade sináptica ou do sistema nervoso periférico, a ação de anticorpos, etc. Ou seja, para o PP, a atenção funcionaria como um tipo de gatilho para que o corpo aja no (e sobre o) mundo de modo a se manter em um estado ótimo (Cf. FRISTON, 2005; 2010; FRISTON, et al. 2010; FRISTON, MATTOU, KILNER, 2011).⁴⁶

Toda vez que, dado um evento, o organismo precisa realizar uma reação comportamental ou selecionar alguma ação ele realiza aquilo que é chamado de “inferência ativa”, ou seja, uma predições acerca daquilo que será “observado” caso determinada ação seja, de fato, realizada (Cf. FRISTON, et al. 2010). Isso significa que tomar decisões

⁴⁵ Logicamente, se houver algo mais saliente ou extremamente surpreendente neste cenário, orientando nossa atenção antes do detalhamento das árvores, então a ordem seria: “floresta primeiro, King Kong segundo, árvores depois” (ou algo assim).

⁴⁶ É possível notar que a descrição da ação fornecida pelo Processamento Preditivo possui raízes conceituais no Princípio Ideomotor, no qual estados neuronais ideacionais, habitualmente relacionados ao processamento das possíveis causas ambientais de nossa ativação sensorial, também mapeiam aqueles estados que promovem a preparação da ação (STOCK, STOCK, 2004, p. 176).

comportamentais implica (1) prever o que ocorrerá ao organismo para cada possível ação tomada e (2) selecionar as ações devidas previstas, de modo a (3) subsequentemente produzir as “observações preferidas” pelo organismo. No PP, tais preferências partem de expectativas prévias sobre as “observações sensoriais”, implicando que os resultados preferidos pelo organismo são aqueles que minimizam a *surprisal* esperada após a tomada de decisão. Por conseguinte, esses resultados preferidos seriam finalmente alcançados quando o organismo realiza, justamente, as ações esperadas para minimizar os erros de predição disso resultantes (SMITH, BADCOCK, FRISTON, 2020, p. 5).

Contudo, se essa história estiver correta, a ação está inteiramente envolvida com as predições, no sentido que o organismo agirá no mundo para buscar ativamente informações sensoriais que ele já espera. Assim, se uma ação ou inferência ativa não for majoritariamente compatível com uma predição, o resultado é uma cascata de erros de predição que perdura até que o modelo gerativo seja alinhado (de maneira bayesiana) com a realidade do indivíduo, através de inferências ativas adicionais. Se isso não for possível, as predições sobre o ambiente precisam ser atualizadas para se ajustarem ao contexto em questão, via inferência perceptiva (FRISTON, 2010, p. 134). Entretanto, quando esse é o caso, o organismo precisa aumentar o ganho de informação com os erros de predição por meio da atenção, e, para que esse ganho de fato ocorra, ele precisa (mais uma vez) se ajustar e alterar a disposição de seus sensores no (e sobre o) mundo. Logo, dizer que o organismo realiza uma inferência perceptiva e, com isso, atualiza suas predições implica, na realidade, dizer que ele realiza uma forma tácita de inferência ativa.

No Processamento Preditivo, o envolvimento da ação implica, então, que predições comportamentais não apenas precedem a sensação e a percepção, mas na verdade as determinam (CLARK, 2013, p. 186). Isto é, os sistemas normalmente envolvidos com o processamento de erros preditivos relativos aos nossos comandos motores se auto-suprimem através da execução de movimentos corporais, capazes de reposicionar nossos sensores para buscar e (como resultado) criar, ativamente, os estímulos sensoriais que já eram esperados por nossos modelos gerativos (FRISTON, 2003, p. 1349).⁴⁷ Logo, vemos que, no PP, percepção, cognição e ação trabalham juntas para minimizar os erros de predição através da amostragem seletiva e modelagem ativa das possíveis causas da estimulação sensorial – por meio da

⁴⁷ O organismo precisa prever também as consequências sensoriais de suas próprias ações, as quais são fortemente informadas por fatos básicos acerca de seu corpo, tais como tamanho, localização dos sensores, alcance dos efetores, etc. (CLARK, 2016, p. 185).

movimentação e da intervenção, no mundo (CLARK, 2016, p. 74). Em outras palavras, nossos cérebros podem até ser concebidos como construtores de nossa percepção do ambiente, mas essa capacidade só nos é disponível através de uma contínua e ativa sintonia corporificada com as informações sensoriais vindas do meio. Esse vínculo, tanto com as propriedades e estados externos quanto com as propriedades e estados internos de nosso ambiente, descreve a inter-relação entre o cérebro, o corpo e seu contexto sob uma perspectiva que tende a se alinhar com uma visão integrada da cognição, ofuscando a linha divisória entre a percepção e a ação. Nisso, os modelos gerativos podem ser descritos como fundamentalmente orientados para a ação, de maneira que o cérebro corporificado procura estabelecer um controle preditivo dinâmico dos estímulos ao seu redor, permitindo que o organismo percorra seu ambiente com sucesso (CLARK, 2013, p. 185; WIESE, METZINGER, 2017, p. 4). Dito (ainda) de uma outra forma, para o PP, o mundo dos seres vivos é essencialmente constituído de acessibilidades/possibilidades de interação (*affordances*) e os erros de predição podem ser entendidos como codificando informações sensoriais que ainda não foram utilizadas para o controle de uma interação relevante para o agente em seu mundo (CLARK, 2016, p. 74).

2.2 A consciência preditiva

Tudo que vimos na seção anterior descreveu o cérebro como uma máquina de inferências preditivas e ativas. Ele parece operar de forma hierárquica e seguir regras bayesianas para atualizar seus modelos preditivos com base nos *feedbacks* encaminhados pelo aparelho sensorial. Aqui, entende-se que a atenção (por meio de inferência ativas) possui um papel fundamental por selecionar a amostra de entradas e estímulos que serão processados pela Minimização de Erro de Predição. A ação promove ao sistema sensorial novas possibilidades de seleção de amostra. No entanto, essa máquina de inferências só parece funcionar porque é corporificada e precisa de um corpo inteiro se movendo para realizar suas inferências sobre as causas das informações constantemente captadas dos estados interoceptivos, estados exteroceptivos e estados proprioceptivos (FRISTON, et al. 2010, p. 233). Por isso, o organismo estaria envolto em um ciclo de ação-percepção que é dependente do mundo e garante uma otimização de utilização de recursos. Entretanto, ainda não ficou claro como a consciência poderia ser definida ou descrita nesse contexto.

2.2.1 Distinguindo-se do meio

O primeiro passo para, internamente ao Processamento Preditivo, reconhecermos a consciência como um processo funcional corporificado consiste em abordar o surgimento de seus componentes mais basilares. Nisso, talvez um bom ponto de partida seja discorrer sobre o estado de vigília, que, conforme vimos em nossa introdução é comumente entendido como o estado ordinário de consciência, no qual observamos uma plena manifestação da atividade perceptiva e motora. Mas, o que significa para um organismo estar em plena manifestação dessas atividades?

De acordo com a Segunda Lei da Termodinâmica postula que a interação entre ambas partes tende a dispersar a energia total do sistema por igual, até que ele alcance um equilíbrio e se torne homogêneo. Dada essa dissipação, pode-se dizer que a Segunda Lei prescreve a inevitabilidade do aumento da entropia nesse tipo de sistema – isto é, de seu grau de desordem ou dispersão de energia. Entretanto, assim posto, não haveria possibilidade de existência de nenhum organismo, mas somente um ambiente equilibrado com ele mesmo. Porém, para que um organismo mantenha sua integridade estrutural e funcional como uma entidade termodinamicamente (localmente) delimitável, ele precisa lutar contra esse equilíbrio global, mantendo-se, então, distinto de seu meio (STEWART, GAPENNE, DI PAOLO, 2010, p. 2). Isso, por sua vez, lhe permitiria limitar a entropia de seus próprios estados constituintes e, assim, garantir temporariamente o seu equilíbrio termodinâmico local, ao “combater” a Segunda Lei da Termodinâmica (RAMSTEAD, BADCOCK, FRISTON, 2018, p. 2-3).

Nesse sentido, os seres vivos podem ser definidos como sistemas autopoieticos termodinamicamente abertos e “longe do equilíbrio”. Isto é, sistemas que, por meio de suas interações e transformações, se regeneram continuamente, realizando redes de processos relacionais que produzem seus próprios estados constituintes, de modo a compor unidades espaciais localmente concretas, em uma constante batalha para se manter em equilíbrio (Cf. VARELA, 1979). Isso significa que, na manutenção desse equilíbrio, o organismo utiliza de suas interações com o meio, se guiando nele de maneira a poder revisitar, ao longo do tempo, estados homeostaticamente favoráveis, mesmo diante da influência contínua e aleatória de flutuações ambientais. E, ao fazer isso, ele modifica seu meio e sua relação com ele, o que modifica, também, as informações que o organismo captará em sequência (Cf. FRISTON, 2005; 2010).

O meio habitado por um organismo seria, então, nada mais do que as consequências de suas ações conforme essas são captadas por suas entradas sensoriais, o que dependeria estritamente daquilo que é propiciado por um contexto habitável. Logo, não poderíamos falar de um meio sem que exista um organismo nele atuando, visto que os seres vivos trazem à tona o meio no qual habitam/existem (Cf. MATURANA, 1987). A distinção existente entre um organismo e seu meio é, conseqüentemente, transdutiva, com a existência de ambos (como entidades delimitáveis) dependendo de suas interrelações cíclicas e dinâmicas.

Assim sendo, qualquer atuação de um organismo em seu meio dependerá, em primeiro lugar, de suas restrições morfológicas e fenotípicas, de modo que essas acabam por criar regularidades inter-relacionais que atraem o organismo para um conjunto de condições específicas, nas quais ele possivelmente se encontrará ao longo de sua vida. Quanto mais estruturalmente complexo for um organismo, mais limitado será esse conjunto de condições (por exemplo, uma bactéria tem mais chances de sobreviver em meios radicalmente distintos do que um ser humano) e, mesmo assim, como a probabilidade de um organismo sobreviver em um meio homeostaticamente favorável será maior do que em um meio desfavorável, sua sobrevivência implica a minimização das chances dele, eventualmente, violar esse conjunto atrator (SIMS, 2021, p. XXXI). Essas regularidades inter-relacionais que um organismo tende a frequentar ao longo de sua existência podem ser entendidas como um modelo de seu nicho (isto é, daquela parcela do meio que lhe é acessível) e de seu comportamento nele. E porque tais modelos estabelecem algo como um mapeamento probabilístico, ao longo do tempo, da interação indireta entre os estados internos e externos de um ser vivo, eles podem ser interpretados (pelo cientista ou filósofo) como estimativas sub-pessoais dos organismos acerca das causas ambientais de alterações em seus estados constituintes.

Com isso, tal mapeamento nos permitiria entender as dinâmicas internas aos seres vivos como realizando inferências quase-ótimas acerca de seus nichos (Cf. CORCORAN, PEZZULO, HOHWY, 2020). Aqui, a observação fundamental é a de que esse processo inferencial forneceria ao comportamento dos organismos uma natureza preditiva. Isso permite ao organismo responder a possíveis perturbações vindas de seus meios ou evitar condições fisiologicamente estressantes e homeostaticamente desestabilizadoras antes mesmo que elas ocorressem – ou seja, a alostase.

Até agora descrevi um organismo que, para sobreviver, precisa minimizar a entropia com seu meio para se distinguir dele e adquirir autonomia sobre ele. Para se manter distinto do meio, o organismo precisa agir nele e, ao fazer isso várias vezes, ele mapeia as regularidades

do ambiente, dando preferência àqueles que são favoráveis à sua existência. Nesse contexto, suas ações estarão cada vez mais otimizadas em relação a esse meio favorável, capacitando o organismo a inferir as causas dos efeitos que esse meio gerou em seu aparato sensorial.

Naturalmente, sob essa escala de análise, não existem intencionalidades, mas apenas normatividades e regularidades probabilísticas, as quais parecem guiar o funcionamento de um ser vivo e lhe dotar de uma forma basilar de vigília preditiva de seu contexto. Mas, ao estender essa narrativa para outras escalas de análise, podemos teorizar acerca de como seus detalhes iluminariam as complexidades do comportamento consciente humano. Para tanto, precisamos de um referencial que justifique esse tipo de generalização e, diante daquilo demonstrado na seção anterior é precisamente isso que o Processamento Preditivo parece ser capaz de fazer: demarcar como a narrativa da emergência de uma cognição preditiva mínima se estenderia às capacidades cognitivas e comportamentais sofisticadas que são definidoras de organismos complexos como os seres humanos.

Com isso em mente, na próxima seção, entrarei em maiores detalhes acerca de algumas formas possíveis de detalhar a maneira como a percepção consciente se daria no Processamento Preditivo. Digno de nota, a partir desse detalhamento faz-se possível começar a questionar se o Processamento Preditivo seria, de fato, um quadro capaz de fornecer uma explicação unificadora do funcionamento cerebral, pois, para tanto, ele precisa responder em maiores detalhes à questão de como e porque certos processos inconscientes atingem a percepção – isto é, a estruturação da experiência consciente (Cf. MARVAN, HAVLÍK, 2020).

2.2.2 Limitações do Processamento Preditivo como uma teoria empírica da consciência

Muitos dos autores que trabalham o Processamento Preditivo sob uma escala de análise neuronal (isto é, numa escala que se compromete com os correlatos neurais da consciência) assumem uma espécie de realismo ao alegar que o cérebro é uma máquina bayesiana aproximada, equipado com um único mecanismo para a minimização dos erros de predição que orienta todas as nossas funções cognitivas (HOHWY, 2013, p. 25). No entanto, para os filósofos Tobias Schlicht e Krzysztof Dołęga (2021), essa interpretação não é universal, havendo ainda controvérsias pendentes acerca do *status* ontológico e explicativo do PP (SCHLICHT, DOLEGA, 2021, p.7; Cf. AITCHISON, LENGYEL, 2017; COLOMBO, WRIGHT, 2017; HEILBRON, CHAIT, 2018; BRUINEBERG ET AL., 2020). Com isso, é

possível distinguir entre interpretações “fortes” – as quais alegam que o cérebro, de fato, utiliza modelos gerativos e realiza inferências bayesianas aproximadas para minimizar erros de predição – e “fracas” do PP – que alegam que o *framework* é uma mera redescrição instrumental do funcionamento dos sistemas cognitivos (“como se” estivessem utilizando modelos e inferências), de maneira que seus mecanismos físicos/causais subjacentes poderiam, na realidade, atuar de forma distinta (HOHWY, SETH, 2020, p. 15).

Ainda, Schlicht e Dołęga nos trazem uma importante observação, originalmente realizada pela filósofa Rosa Cao, sobre a possibilidade dos modelos preconizados pelo Processamento Preditivo poderem ser “re-rotulados em termos tradicionais, não-preditivos, por não possuírem consequências empíricas relevantes para a análise dos dados existentes ou futuros” (CAO, 2020, p. 517). Isto é, redescrever uma atividade neuronal em termos conceitualmente mais tradicionais ou conforme o PP não parece trazer diferenças substanciais imediatas para os dados empíricos coletados (CAO, 2020, p. 522). Logo, mesmo que uma interpretação grosseira do Processamento Preditivo seja fornecida para estudos objetivando a escala neuronal, não parece que isso facilitaria a leitura e compreensão dos dados (disso) resultantes. Para Peter Kok e Floris de Lange (2014), o PP deveria, então, buscar oferecer uma descrição de alto nível que permitisse o estabelecimento de relações precisas entre evidências de correlatos neurais e a maneira como esses mudariam ao longo da experiência consciente. No entanto, dadas as ambiguidades conceituais que ainda imperam sobre o *framework*, é improvável que o Processamento Preditivo consiga guiar esse tipo de interpretação causal e, por tanto, liderar o progresso nas pesquisas acerca dos correlatos neurais da consciência – conforme o originalmente objetivado por Seth e Hohwy (Cf. SETH, HOHWY, 2020; SCHLICHT, DOŁĘGA, 2021, p.13).

Contudo, naquilo que diz respeito à percepção consciente (e portanto, à escala pessoal de análise), embora aparente um *status* mais promissor, o Processamento Preditivo também precisa preencher algumas lacunas e resolver problemas conceituais que o distanciam de ser uma teoria filosoficamente robusta e empiricamente consistente da consciência. Para Tomáš Marvan e Marek Havlík (2020), ainda não parece haver um consenso sobre até onde o Processamento Preditivo seria capaz de ilustrar a transição entre processos inconscientes e conscientes e, assim, fornecer uma explicação para a “emergência” da fenomenalidade. Segundo os autores, a literatura contemporânea a explorar o PP como teoria da consciência geralmente segue uma das seguintes argumentações de cunho constitutivo: (1) a consciência é o resultado de atualizações inferenciais de hipóteses/predições perceptivas ou (2) a

consciência é um mosaico das hipóteses/predições perceptivas que “melhor explicam” as causas dos efeitos captados pelo aparato sensorial (MARVAN, HAVLÍK, 2020, p. 3).

A primeira posição sugere que as constantes atualizações de hipóteses perceptivas via erros de predição irão tornar os conteúdos perceptivos conscientes (SMITH, FRISTON, WHYTE, 2022, p. 7), com aquilo que não for atualizado não sendo conscientemente percebido. Isto é, a consciência, nesse contexto, seria o processo de otimização de predições/hipóteses perceptivas por meio de atualizações inferenciais (ou seja, a formação de predições/hipóteses perceptivas estatisticamente otimizadas) e, por isso, se não houver erros de predição a serem dissolvidos (*explain away*) em tais predições/hipóteses, não haveria nada para se estar ciente (HOBSON, FRISTON, 2016, p.251; HOHWY, 2012, p.6).

Embora à primeira vista ela pareça plausível, Marvan e Havlik apresentaram quatro problemas implicados por essa posição. O primeiro deles diz respeito ao fato de que nem todo conteúdo perceptivo é o resultado da atualização de nossas evidências perceptuais, visto que um sujeito pode ficar encarando um objeto sem se movimentar (ou seja, sem adquirir novas informações sobre ele) e, mesmo assim, não deixar de estar ciente dele. Logicamente, sempre existirá a presença dos movimentos sacádicos dos olhos e demais micro-ações, mas o ponto é que a atualização não parece necessariamente ocorrer para todas as propriedades conscientemente percebidas. O segundo problema seria a existência de “prévias”/*priors* perceptivos recalcitrantes – ou seja, informações prévias gerando expectativas que, mesmo com novas evidências, não são atualizadas – como é o caso de ilusões de óptica tais quais a *Hollow Mask Illusion* (Cf. GREGORY, 1973) e a *Ames Window* (Cf. AMES, 1951) que, mesmo quando sabemos seu mecanismo, não somos capazes de “escapar” de sua influência. O terceiro problema aponta o fato de que atualizações inferenciais podem ocorrer sem que, de fato, isso leve seu conteúdo a alcançar a consciência, como o ilustrado em estudos de percepção auditiva (Cf. FUJIOKA, et al., 2009). Finalmente, o último problema apontado pelos autores é que a percepção consciente pode surgir sem a presença de uma atualização inferencial, tal como ocorre em experiências conscientes ilusórias orientadas por contextos altamente preditivos – um clássico exemplo das quais é a ilusão de Kanizsa, na qual contornamos ilusoriamente um triângulo ao “preencher” os espaços em branco de uma imagem⁴⁸ (MARVAN, HAVLIK, 2020, p. 3-4).

⁴⁸ Disponível em: https://en.wikipedia.org/wiki/Illusory_contours#/media/File:Kanizsa_triangle.svg

A segunda posição sugere que a percepção consciente seria determinada pelas predições/hipóteses com a maior probabilidade posterior, isto é, aquelas que melhor explicam – em termos bayesianos – as causas dos efeitos que o aparelho sensorial captou. Tais hipóteses são chamadas por Hohwy de “Hipóteses Vencedoras” (Cf. HOHWY 2012; HOBSON, 2014). Essa posição responde bem ao primeiro problema identificado por Marvan e Havlik na perspectiva anterior, pois, aqui, o cérebro simplesmente manteria a melhor predição/hipótese sobre aquilo percebido, mesmo sem haver qualquer atualização, sugerindo que não seria tal processo que levaria o conteúdo inconsciente à consciência, mas as próprias hipóteses vencedoras, em si (MARVAN, HAVLIK, 2020, p. 4).

Existem boas razões para levarmos essa segunda posição em consideração, dado que a percepção no Processamento Preditivo funciona predominantemente de acordo com um fluxo descendente, o qual carrega as informações inferenciais – ou seja predições/hipóteses – que buscam inibir a informação sensorial ascendente. Disso, um exemplo que poderia corroborar a ideia da hohwyana da hipótese vencedora está na percepção auditiva de “palavras fantasma” (Cf. DEUTSCH, 2019). Por sua vez, outro exemplo contundente seria aquele ilustrado pelas alucinações, nas quais uma predição/hipótese vencedora dominaria a percepção consciente independentemente das entradas sensoriais (Cf. ADAMS, BROWN, FRISTON, 2014). Entretanto, não parece correto afirmar que predições/hipóteses vencedoras sempre necessariamente atingirão a consciência (MARVAK, HAVLIK, 2020, p. 5). Dada a existência de fenômenos tais como o da mobilidade aparente, Vetter e colaboradores (2014) notaram que, em alguns casos, as predições que antecipam de forma acurada os eventos subsequentes parecem atuar e ser formuladas sem, com isso, adentrar à consciência. Isso sugere que, se uma hipótese vencedora pode não chegar à consciência, então denominá-la “vencedora” não seria o suficiente para que seu conteúdo seja conscientemente percebido (mesmo que essa seja uma condição necessária). Outro fenômeno a dificultar a hipótese vencedora consiste no anteriormente mencionado *blindsight*, no qual pacientes com lesões na área V1 do córtex visual são capazes de parcialmente processar os estímulos visuais, mas não são capazes de experienciá-los conscientemente. Aqui, predições/hipóteses parecem atuar e ser formuladas de maneira suficiente, embora imprecisa, para, de alguma forma, guiar tais pacientes em seu campo cego – sugerindo a existência de uma hipótese vencedora independente da percepção consciente do sujeito (MARVAK, HAVLIK, 2020, p. 5-6).

Com isso em mente, Marvan e Havlik concluíram que o Processamento Preditivo não seria capaz de (sozinho) ilustrar como a transição de conteúdo inconsciente para a consciência

ocorreria. Ainda, o PP precisa fornecer uma explicação para a fenomenalidade. Sob a óptica de Hohwy (2013), a percepção consciente é resultante de uma unificação de conteúdos, tanto a nível local (cores, texturas, formas, etc, reunidos num único objeto) quanto global (a unicidade de cenários em sua completude). E, essa natureza una do campo perceptual seria o resultado da relação existente entre inferências perceptuais e a ação – visto que a atuação de um sujeito preditivo só é bem-sucedida (de um ponto de vista bayesiano) quando uma das inferências perceptuais por ele geradas for selecionada para compor a percepção consciente, via inferência ativa. De acordo com essa concepção, o aspecto fenomênico seria resultante da relação entre características perceptuais estáveis e de alto nível na hierarquia preditiva (inferidas a partir de padrões no ambiente) com características dinâmicas e de baixo nível hierárquico (captadas por meio dos erros de predição) (HOHWY, 2012, p. 13).

Objetivando especificamente tal posicionamento, Marvan e Havlik levantam três outras críticas. A primeira consiste no fato de que a explicação que Hohwy fornece para a fenomenalidade só trata das características estruturais quantitativas da percepção – isto é, das interações sistemáticas entre os vários conteúdos informacionais que percebemos conscientemente – ao passo que a fenomenalidade seria uma característica qualitativa. A segunda é que a existência de relações entre características perceptuais de alto e baixo nível não seria algo exclusivo do Processamento Preditivo, dado que ela é comumente utilizada na explicação da visão consciente em termos gerais. Por fim, embora pareça que tais características se relacionam estritamente em um nível consciente, esse não é o caso, visto que, conforme vimos nas seções anteriores, os mecanismos e processos descritos pelo PP seriam perfeitamente capazes de processar e unificar o conteúdo perceptivo inconscientemente – e, inclusive, essa talvez seja a alternativa mais conhecida e defendida entre os proponentes do *framework* (MARVAN, HAVLIK, 2020, p. 7).

Em suma, a literatura existente não aparenta ter esclarecido uma forma através da qual o Processamento Preditivo poderia elucidar satisfatoriamente aquilo que entendemos como consciência e sua fenomenalidade (isto é, os problemas da consciência permaneceriam intactos). Com efeito, esse último ponto talvez não seja algo necessário, afinal, teorias neurocientíficas tais como a *Global Neuronal Workspace Theory* (GNWT) propõem uma explicação de como os conteúdos de processos cognitivos entrariam no “fluxo da consciência” sem sequer discorrer a respeito de suas propriedades fenomênicas. Inclusive, atualmente parece haver uma forte movimentação a favor da aproximação entre o PP e a GNWT, que,

mesmo excluindo a fenomenalidade, propõe ser capaz de fornecer melhores explicações para a transição do conteúdo inconsciente para a consciência (HOHWY 2013; DEHAENE 2014).⁴⁹

Não obstante, uma forma mais coerente e parcimoniosa de resolver o problema explicativo do Processamento Preditivo acerca da consciência – tal como o ilustrado por Marvan e Havlik (2020, p. 7) – exigiria alinhá-lo a teorias e/ou modelos filosóficos, capazes de lhe oferecer as elucidações faltantes para que ele possa, enfim, tornar-se um quadro teórico mais robusto e compreensivo. Exemplos desse tipo de parceria envolvem a exploração de sua possível inter-relação com *Higher Order Theories* (BROWN, LAU, LEDOUX, 2019); a *Information Integration Theory* (BUCCI, GRASSO, 2017); a *Attentional AIR Theory* (MARCHI, HOHWY, 2020) e, finalmente, o *Multiple Drafts Model of consciousness* (DOŁĘGA, DEWHURST, 2021). Ao que nos parece (e conforme veremos a seguir), devido ao seu potencial de transformar o Problema Difícil da consciência em um conjunto de Problemas Fáceis ou, simplesmente, no Problema da Ilusão, esse último parece ser o melhor candidato para resolver as insuficiências em questão, rumo a um esclarecimento de como a percepção consciente poderia ser devidamente concebida dentro do Processamento Preditivo.

⁴⁹ Uma possibilidade alternativa seria, ainda, assumir o caráter preditivo da GNWT, reinterpretando seus postulados em termos do Processamento Preditivo (Cf. WHYTE, 2019).

3. A inter-relação entre o MME e o PP para o estudo da consciência

O mundo em que vivemos é extremamente complexo e desafiador, exigindo tanto estratégias rápidas e eficientes de percepção-ação quanto raciocínios lentos e trabalhosos. Conforme vimos, a forma que encontramos para (sobre)viver nesse mundo foi, muito provavelmente, antecipar seus eventos e estados (Cf. GODFREY-SMITH, 2002) e ativamente moldar nosso “véu sensorial”. Nesse contexto, o cérebro poderia ser entendido como um órgão corporificado para o controle da ação e para a regulação do organismo, o qual é ambientalmente situado (CLARK, 2016, p. 250). Contudo, ele não representa o mundo de uma maneira passiva e descritiva, com modelos internos ricos e detalhados que meramente aguardam o processamento de uma “central cognitiva” para deduzir as ações. Em vez disso, nosso cérebro tenta prever o que há lá fora, usando ativamente e dinamicamente os sinais sensoriais captados para ajustar essa previsão. Trata-se de uma atividade que envolve ciclos complexos nos quais a ação determina a percepção que, por sua vez, possibilita a própria ação, de modo a otimizar a exploração das *affordances* do contexto (CLARK, 2016, p.268).

Em sua parcela final, o Capítulo 1 expôs e desenvolveu o modelo dennettiano da consciência, no qual variáveis que influenciam o sistema sofrem constantes atualizações e, eventualmente, substituições por novas variáveis, guiadas por processos de sondagem, dadas as entradas sensoriais. Por sua vez, na Seção 2.1, o Capítulo 2 expôs e desenvolveu os processos e mecanismos que, no Processamento Preditivo, atuam subjacentemente à minimização dos erros de previsão e os quais, em tese, possibilitariam a percepção consciente conforme sua descrição na Seção 2.3. Diante disso, a tarefa do presente Capítulo será partir de tais considerações e explorar a possibilidade de relação mútua que o modelo dennettiano e o PP. Conforme veremos, existem algumas semelhanças importantes entre ambos, as quais parecem permitir certo intercâmbio conceitual, de modo que talvez seja possível traduzir o Modelo dos Múltiplos Esboços em termos preditivos e vice-versa. Aqui, meu argumento central é a de que essa tradução viabilizaria a solução de alguns problemas e paradoxos enfrentados por ambos e que isso poderia favorecer ainda mais o desenvolvimento de uma descrição robusta da percepção consciente e, de certa forma, dos estados locais da consciência – pelo menos naquilo que tange ao papel exercido por nossos cérebros contextualizados.

3.1 Algumas semelhanças

De acordo com Dołęga e Dewhurst (2021), as semelhanças mais latentes, existentes entre o MME e o Processamento Preditivo, são: (1) a estruturação e função das predições no PP se assemelham aos esboços dennettianos; (2) a Minimização de Erro de Predição parece atuar como o mecanismo que orientaria os processos de sondagem; (3) e a ideia de fama no cérebro, conforme concebida no Modelo dos Múltiplos Esboços, parece equivaler à tese hohwyana da hipótese vencedora. De fato, vimos que as predições/esboços possuem um papel fundamental na percepção humana, pois são o que estruturam e, muitas vezes, definem o conteúdo perceptivo. Graças a isso, também, o cérebro não precisa processar cenários perceptivos inteiros a todo momento, mas apenas novos detalhamentos que vão surgindo ao longo do tempo, de modo a diminuir drasticamente o custo energético de seu funcionamento durante a percepção.

Na medida em que os mecanismos do PP e, depois, do MME foram sendo descritos, o leitor pode ter ficado com a impressão de que esboços e predições poderiam ser tratados como a mesma coisa – e com razão. Pois, assim como as predições, os esboços são caricaturas ou fragmentos informacionais anteriormente processados que poderão compor narrativas que formam o conteúdo consciente. Ambos também não estão comprometidos com nenhuma linha de tempo rígida, dada a indeterminação do conteúdo consciente que auxiliam a compor (Cf. Seção 1.3). Tanto esboços quanto predições são confeccionados em múltiplos níveis de complexidade e abstração e existe uma forte competição entre esboços e entre predições para ganhar a influência do sistema. Ambos são passíveis de simples atualizações, conforme as novas informações vão chegando desde nosso aparato sensorial e, eventualmente, de sua inevitável substituição por novas predições ou esboços. Finalmente, ambos, em sua atuação, geram expectativas que alteram as dinâmicas subjacentes aos estados subsequentes do sistema, influenciando o acionamento de redes neuronais periféricas (compostas por outras predições/esboços), as quais podem vir a ser influentes, de acordo com o contexto.

Uma vez que determinados esboços/predições ganham influência no sistema, eles atrairiam o mecanismo que talvez seja o responsável por transmiti-los globalmente e cuja atuação seria, portanto, fundamental para que tal conteúdo seja acessado conscientemente: a Minimização de Erro de Predição ou a sondagem. Na MEP, as melhores predições são enviadas em um fluxo descendente rumo ao encontro do fluxo ascendente de informações sensoriais. Aqui, ela fomenta, de alguma forma, a inibição da transmissão do fluxo

ascendente, de modo que apenas as informações divergentes das previsões descendentes sejam transmitidas adiante. Especificamente, a MEP teria o poder de alternar entre previsões ou induzir um aumento na absorção de novas informações orientada a ação através da atenção. Para Dołęga e Dewhurst (2021), tanto a atenção exógena quanto endógena teriam um papel fundamental na MEP. A primeira operaria de maneira ascendente, aumentando a precisão das previsões e “explicando” os sinais de erro ao aumentar a força das hipóteses descendentes que geram tais previsões. Isso, por sua vez, aumentaria a influência dessas hipóteses, deixando-as mais suscetíveis a “dominar” o sistema. A segunda operaria de maneira descendente, aumentando a precisão esperada pelas previsões correspondentes ao objeto da atenção, resultando em uma alteração da saliência das entradas sensoriais, de maneira a confirmar tais previsões. No contexto dennettiano, então, a atuação da Minimização de Erro de Predição poderia ser equiparada à sondagem, dado que ela consiste em um processo de alto nível que prevê os estados do sistema de acordo com a precisão dos sinais de erro captados (DOŁĘGA, DEWHURST, 2021, p.17).

Nisso, da mesma forma como a MEP pode ser considerada como uma espécie de “governador virtual” da hierarquia cognitiva, o processo de sondagem também possuiria, para Dennett, uma caracterização semelhante – com ambas sendo responsáveis pelo equilíbrio entre as informações que o sistema já possui e a constante captura de novas informações sensoriais. Ainda, ambas operam tanto interna quanto externamente, a nível pessoal e supessoal, valendo-se da atenção para tanto. Digno de nota, embora a alusão a um “governador” possa fazer parecer que tanto a MEP quanto a própria sondagem sejam releituras de algo como o Teatro Cartesiano, ambas possuem natureza meramente virtual, ou seja, não se tratam, de fato, de unidades de comando ou locais centrais, nos quais tudo aquilo processado finalmente toma forma e passa a fazer sentido, mas sim múltiplas redes complexamente conectadas – o *pandemonium* de Dennett – das quais uma espécie de ordem ou sincronia parece emergir.

Por fim, em certo nível de abstração, a ideia de uma fixação ou fama no cérebro pode ser facilmente equiparada com a tese da hipótese vencedora, visto que ambas descrevem a percepção consciente como o resultado da influência, no sistema nervoso, da dinâmica entre conteúdos informacionais potencialmente concorrentes. No contexto dennettiano, esses conteúdos influenciadores são denominados fragmentos narrativos, os quais, em determinadas situações, tornam-se disponíveis para o acesso consciente – assim como ocorre com as melhores hipóteses no Processamento Preditivo. Em ambos casos não há critérios definidos de

qual exatamente será a narrativa/hipótese mais influente, além de não haver um momento correto, que seja passível de observação, no qual a próxima hipótese/narrativa é definida como aquela que, subsequentemente, ganhará tal influência. Tudo isso ocorreria dinamicamente, em termos probabilísticos e a nível sub-pessoal. Além disso, considerando a função geral que a atenção possuiria nesse esquema de entendimento, salientando determinados conteúdos a compor nossa percepção consciente, podemos considerar que, na ausência de seu devido direcionamento, hipóteses vencedoras/fama no cérebro pré-estabelecidas atuariam como um “piloto automático”, guiando o organismo de maneira leviana em seu meio.

Talvez a única diferença relevante entre hipótese vencedora e fama no cérebro é que originalmente a primeira estaria no singular, isto é, seria apenas uma hipótese vencedora influenciando o sistema, ao passo que a segunda não haveria um número definido de narrativas a chegar na fama. Entretanto, Dołęga e Dewhurst notaram que, em termos gerais, não parece ser um problema considerar que, no PP, a todo momento, múltiplas hipóteses preditivas concorreriam por sua eleição à fama – ou seja, para se tornarem uma hipótese vencedora. Na verdade, tal interpretação parece ser até mesmo algo vantajoso, fornecendo uma heurística mais realista para as dinâmicas internas ao sistema nervoso (Cf. WARREN, 2021). Veremos agora como tal inter-relação ocorreria em maiores detalhes.

3.2 As múltiplas sondagens

Um dos problemas do MME anteriormente apontados diz respeito ao mecanismo de sondagem. De fato, embora a narrativa dennettiana aparenta atribuir à sondagem a capacidade de exercer influência em ambos os níveis de nossa cognição, o fato de que Dennett não realiza claramente tal distinção acaba deixando espaço para certa ambiguidade conceitual em seu modelo. Aqui, penso que a atuação do mecanismo descrito por Dennett não deveria ser considerada algo experienciável pelo sujeito, mas atuando apenas em um nível neuronal de análise. Aliás, seria catastrófico e, de certa forma, paradoxal para o sujeito possuir acesso e agência conscientes sobre seu mecanismo de sondagem – algo como um “botão de liga/desliga” para sua percepção interna e externa – visto que, se esse fosse o caso, ele provavelmente desenvolveria diversas anomalias e descompassos em suas dinâmicas caso decidisse, digamos, por deixar (mesmo que momentaneamente) de perceber/captar os estímulos ao seu redor, o que eventualmente levaria-lo à morte. Talvez seja mais interessante, então, relegar a atuação do mecanismo de sondagem dennettiana apenas ao nível sub-pessoal,

de maneira a assemelhá-lo àquilo exercido pela atenção endógena e exógena no *framework* do Processamento Preditivo, conforme o descrito no Capítulo 2 e ulteriormente explorado na próxima Seção. Portanto, faz sentido conceber as sondagens como ocorrendo a todo momento, ativamente elegendo múltiplos esboços relevantes e coerentes com as informações captadas pelos sensores. Além disso, em si (e para além de seus efeitos), os processos de sondagem não parecem ser subjetivamente percebidos e alterados, ainda que a integração corpórea existente entre o nível pessoal e sub-pessoal da cognição indique a existência de influências e alterações indiretas diversas do acesso e agência conscientes sobre sondagens vindouras.

Grosso modo, substituindo a atuação de um mecanismo de sondagem a nível pessoal, podemos conceber a atuação dos sub-mecanismos⁵⁰ de introspecção e foco, conforme o ilustrado na **Figura 3**. O primeiro dentre esses sub-mecanismos seria o responsável pelo tipo de “atenção interna” que o sujeito pode direcionar à suas atividades mentais, além de subjetivamente utilizar para alterar seu fluxo de pensamentos, atuar sobre o ambiente, tomar decisões, etc., quando precisar e/ou desejar, sem que, com isso, acabe por “habitar” estados desfavoráveis à sua existência.⁵¹ Aqui, “introspecção” se refere a um processo ainda mais amplo do que aquele descrito no Capítulo 1 – isto é, qualquer ato que envolva a agência do sujeito tendo como foco a sua própria vida mental. Por sua vez, o segundo dentre esses sub-mecanismos seria o responsável pelo tipo de “atenção externa” que o sujeito pode direcionar ao seu contexto – isto é, os cenários e os detalhamentos para os quais ele foca a sua atenção.⁵² Ainda, e ao contrário das sondagens a nível sub-pessoal, a atuação de ambos esses sub-mecanismos não necessita ocorrer sempre, uma vez que é plenamente possível termos a sensação subjetiva de “não pensar em nada” e/ou navegar por nosso mundo no “piloto automático”.

⁵⁰ O objetivo das explicações mecanicistas é dismantelar um determinado fenômeno-alvo em suas partes (isto é, mecanismos) e inter-relações constitutivas, criando modelos de sua organização. Quando um mecanismo decorrente de tal dismantelamento se torna, por sua vez, o fenômeno-alvo de uma explicação mecanicista, suas partes são comumente denominadas sub-mecanismos (Cf. BECHTEL, HAMILTON, 2007).

⁵¹ Naturalmente, situações nas quais um indivíduo é patologicamente acometido com pensamentos e ideias suicidas ou por impulsos comportamentais altamente destrutivos não são consideradas nesse esquema. Aqui, apenas o funcionamento neurotípico, em detrimento de transtornos psicológicos graves, são utilizados para estruturar a descrição que se segue.

⁵² Graças ao foco, somos capazes de, também, destacar em nossa percepção consciente um objeto ou uma atividade específica, de modo a facilitar, por exemplo, a aprendizagem e/ou a execução de tal algo – entre outras muitas possibilidades. Para maiores detalhes sobre essa forma de atenção, Cf. GAZZANIGA, IVRY, MANGUN, 2019; WOLFE, KLUENDER, LEVI, 2018.

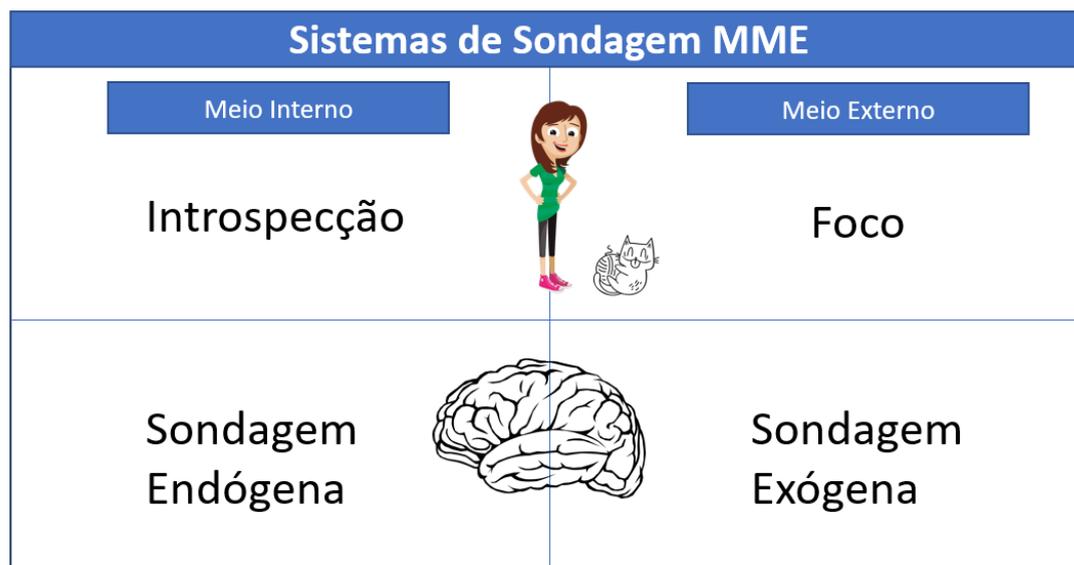


Figura 3: Ilustração das diferentes formas de atenção existentes no MME. A nível pessoal, representado pela mulher, a introspecção se volta ao meio interno e o foco ao meio externo. A nível sub-pessoal, representado pelo cérebro, processos de sondagem se voltam tanto para cenários internos quanto externos.

De qualquer forma, por partirem de descrições realizadas desde níveis distintos de análise é importante salientar que, mesmo com essas divisões, ambos esses sub-mecanismos e as múltiplas sondagens dennettianas trabalham de forma integrada, o que significa que qualquer alteração em um gera consequências e modificações imediatas nos demais, no intuito de auxiliar o alcance de um estado ótimo de funcionamento interno e o sucesso na interação com o meio. A título de ilustração, quando um sujeito olha seu gato brincando com uma bola, digamos que surja nele um desejo de se lembrar do bichano quando filhote e assim faz, alterando o fluxo “normal” de suas dinâmicas internas e atividades neuronais associadas à cenários perceptivos corriqueiros (ao introduzir uma atividade de introspecção). Como isso seria possível? Como vimos, nem todo esboço fica ativo por tempo o suficiente para atrair as sondagens, mas isso não quer dizer que ele desapareça sem deixar vestígios. Quando o sujeito olha para seu gato brincando, múltiplos esboços, referindo-se tanto a informações acerca do gato quanto adjacentes a ele, são ativados em seu sistema nervoso. Deste modo, quando o sujeito introspecta para resgatar a memória de seu gato filhote, o que ele está realizando é um direcionamento de sua atenção endógena para esboços relacionados ao gato filhote, os quais já haviam sido corticalmente ativados, mas que não foram necessariamente eleitos para a percepção externa – visto que não seriam satisfatoriamente coerentes com a informação perceptiva captada naquele momento. Isso parece implicar que, mesmo quando não “estão

aptos” a compor a percepção consciente, certos esboços neuronalmente fomentados tendem a ficar salientes à atuação do mecanismo de sondagem e, assim, à sua eleição em eventuais condições oportunas.

Particularmente – e em congruência com aquilo enunciado pela literatura do PP – essas condições oportunas possuem “natureza probabilística”. O que quero dizer é que, a captação de informações vindas de um ambiente implica o fomento, tanto direto quanto indireto, da atividade mental que se refere mais apropriadamente a esse ambiente. Dessa forma, é possível falar que o processamento de informações ambientais específicas ativa uma rede de conexões dinâmicas que é única para cada organismo, dado seu histórico de interações com ambientes similares àquele da informação processada. Dentro dessa rede de conexões, cada “membro” (ou *demon*) é estimulado por parcelas específicas da informação processada até chegar em um limiar que o permite transmiti-las para o restante da rede. Se essa transmissão acabar por ganhar força na rede em questão, ela poderá ser transmitida para as demais redes que compõem nosso sistema nervoso e possivelmente atrair a sondagem. Talvez seja por isso que, em determinados contextos, momentos, situações, etc., temos a sensação de que certos pensamentos e vontades surgem-nos “do nada”, e talvez seja por isso que, ao olhar para o gato brincando, o sujeito teve o desejo de se lembrar dos tempos em que ele era um filhote. Ainda, o resultado de introspectar acerca de esboços tanto centrais quanto periféricos à informação sensorial implica a subsequente execução de ações (o sujeito pode expressar sua recordação em palavras ou tentar abraçar o gato, etc.), concluindo nisso que também podemos considerar como um ciclo ação-percepção que parece se retro-alimentar. Em outras palavras, tudo aquilo que ocorre e acessamos acerca de nossa vida mental está inteiramente conectado às informações e nossas ações no ambiente externo. Concomitantemente, o ambiente externo exerce forte influência sobre nossa atividade mental – o que, por sua vez, nos induz a atuar no mundo e, com isso, alterar o próprio ambiente externo que nos influencia. Novamente, em concordância com o PEL, é essa dança entre um organismo e seu ambiente que fornece condições ótimas de sobrevivência para o organismo e garante a continuidade do próprio ambiente no qual ele se encontra (conforme uma parcela de seu nicho, Cf. MATURANA, 1987).

3.3 Como inferências inconscientes alcançam a consciência?

A descrição da Minimização de Erro de Predição até agora demonstrou um mecanismo extremamente complexo e cheio de variáveis que determinam seu funcionamento. Talvez a primeira pergunta que possa surgir nesse contexto é como a MEP exatamente faria com que as inferências inconscientes, por melhores que sejam, alcancem a consciência. Como esse salto entre os níveis sub-pessoal e pessoal da cognição ocorreria? No Capítulo 2, Seção 2.1.3, foi brevemente sugerido que a atenção seria a responsável, embora tal argumento não tenha sido satisfatoriamente elaborado. Portanto, vamos aos detalhes. Conforme vimos, o funcionamento da atenção implica a utilização de informações prévias acerca do contexto desde o qual os erros de predição são gerados, com o objetivo de realizar uma avaliação da confiabilidade desses erros. Nisso, o sistema emite previsões descendentes sobre os estados esperados em camadas específicas da hierarquia neuronal, prevendo o grau de precisão que os erros ascendendo desde níveis inferiores devem possuir. Efetivamente, o sistema prevê seus próprios estados (algo como “meta-predições”) para implementar a atenção e seria precisamente isso que, segundo Dołęga e Dewhurst, equivaleria à sondagem dennettiana (2021, p.17).

Para os autores, a sondagem poderia, então, se apresentar de duas formas diferentes: (1) estando presente em toda a hierarquia neuronal ou (2) em apenas algumas de suas camadas. No primeiro caso, a ideia de sondagem se distanciaria da original dennettiana e, portanto, se aproximaria daquilo previsto pela *Attention Schema Theory*, de Michael Graziano (Cf. GRAZIANO, 2015). No segundo caso, seria necessário discernir quais dentre as meta-predições operariam como sondas. Na opinião dos autores, um dos critérios possíveis para tanto consistiria na influência que determinada meta-predição teria sobre múltiplas camadas de processamento – isto é, quanto maior for sua capacidade de predição sobre os próprios estados do sistema, mais chances ela teria de ser categorizada como uma sondagem. Para tanto, parece que a estruturação dessas meta-predições precisaria gozar de um longo alcance cortical se aproximando, portanto, daquilo descrito pela *Global Workspace Theory*, de Bernard Baars (Cf. BAARS, 1997). Dessa forma, as sondas seriam distinguíveis das outras meta-predições por possuírem características funcionais e estruturais únicas (DOŁĘGA, DEWHURST, 2021, p.17).

Na seção anterior, realizei primeiramente a separação entre as sondagens que ocorrem a nível sub-pessoal e os sub-mecanismos que atuam a nível pessoal. Embora qualquer rigidez nessa separação não seja plenamente condizente com o atual estado de nosso conhecimento psicológico (Cf. WARREN, 2021), tomada como heurística, ela parece ser capaz de esclarecer e melhor delinear o funcionamento dos processos de sondagem, ao, por exemplo, explicitar que o sujeito não possui o controle total e direto daquilo que se refere, mais propriamente, aos níveis inferiores do sistema. Além disso, mencionei que as sondagens sub-pessoais possivelmente ocorrem a todo momento. Entretanto, agora torna-se claro que isso não ocorreria necessariamente por todo o sistema nervoso, pois uma implicação da existência de sondas em todas as camadas da hierarquia preditiva seria a saliência de informações perceptuais irrelevantes para o nível pessoal e, conseqüentemente, uma “poluição” perceptiva. Ainda, se esse fosse o caso, possivelmente precisaríamos de novos critérios para eleger quais das informações previstas contidas nessas várias sondas alcançaria a consciência. Com isso, parece-me mais sensato supor que essas sondagens ocorreriam apenas em locais estratégicos, embora ainda não saibamos dizer exatamente onde – além da suposição de que estariam relacionados aos circuitos neuronais responsáveis pela atenção endógena e exógena (os quais são cruciais para a MEP). Aliás, em seu modelo de consciência, o próprio Dennett não se comprometeu em pormenorizar o papel da atenção *per se* (DENNETT, 2006, p. 138) – o que parece ter influenciado sua opção pela não-separação entre sondagens sub-pessoais e sub-mecanismos pessoais. De fato, para ele, o objetivo maior é explicar a metáfora da atenção em termos funcionais (de coisas que “chamam a atenção”) sem pressupor uma fonte que “dê a atenção” (*ibid.*).

Como vimos, no Processamento Preditivo, a atenção seria entendida como a atribuição da precisão esperada para uma atenuação ou acréscimo dos erros de predição oriundos de um determinado contexto. A atenção exógena opera aumentando a precisão das predições, “explicando” os sinais de erro e aumentando a força da hipótese que gera tais predições. Isso aumenta a influência dessa hipótese e a deixa mais suscetível a dominar o sistema. De acordo com Dołęga e Dewhurst, a atenção exógena pode ser comparada com a maneira pela qual um evento súbito poderia atuar como uma sonda, gerando um relato consciente sobre o estado de um sistema que, anteriormente ao evento, estava operando no “piloto automático” (DOŁĘGA, DEWHURST, 2021, p. 18). Por exemplo, ao dirigir por um local conhecido, geralmente conduzimos de “forma automática” até que, digamos, uma moto inesperadamente cruza nosso caminho, forçando uma frenagem. Isso gera surpresa, uma cascata de erros de predição e

desperta a atenção exógena, a qual irá assimilar o ocorrido com base nas informações prévias mantidas pelo organismo, confeccionando uma hipótese retrospectiva acerca do que ocorreu antes do cruzamento (quase letal) com a moto. Ainda, quando uma aluna desatenta é questionada por seu professor sobre o que foi dito durante a aula, isso também gera surpresa, uma cascata de erros de predição (pois não é esperado que o professor interrompa a aula para isso) e desperta a atenção exógena (ou, como se diz: “volte para a Terra, Maria!”). Rapidamente, o sistema resgata as informações prévias e disponibiliza uma narrativa para ser relatada à aluna.

Aqui, a atenção exógena parece determinar o conteúdo da percepção consciente ao eleger, à consciência, hipóteses previamente inconscientes. Especificamente, a estimulação que acaba por ativar a atenção exógena gera cascatas de erros de predição a nível sub-pessoal que, por conseguinte, gera cascatas a nível pessoal (isto é, susto/surpresa manifesta). A surpresa manifesta induz o sujeito a se atentar ao contexto, o que resulta em novas amostragens confiáveis. Com isso, o sistema minimiza o erro de predição e atualiza seu modelo gerativo, de modo que, quando o susto/surpresa a nível pessoal finalmente passar, o sujeito executa um plano de ação adequado, baseado especificamente naquele contexto⁵³. Sabe-se que a estruturação desses planos de ação nem sempre percorre grandes porções do sistema nervoso, especialmente se o contexto acaba por ativar nossos mecanismos de luta ou fuga. Entretanto, de acordo com a literatura do PP, o funcionamento desses mecanismos também pode ser descrito por meio da minimização dos erros de predição (Cf. FELDMAN, FRISTON, 2010).

Por outro lado, vimos que a atenção endógena opera diretamente sobre o fluxo descendente, aumentando a precisão esperada pelas predições correspondentes ao objeto de nossa atenção, resultando em uma saliência das entradas sensoriais que confirmam tais predições. Por exemplo, quando procuramos pelas chaves de casa num quarto bagunçado, estamos conduzindo a captação de amostras sensoriais com um determinado modelo preditivo em pauta: as chaves. Isso faz com que fiquemos mais propensos a detectar objetos tais como chaves. Uma vez que o alvo é detectado, o ganho dos erros de predição será enviado para minimizar a intensidade da predição correspondente. Ou seja, a modulação atencional, nesse caso, viria do próprio sistema, o qual fornece destaque a um esboço ou predição que corresponde e possibilita a ação de procurar pelas chaves. Ainda, mesmo que o sistema

⁵³ Naturalmente, esse plano de ação tende a ocorrer instantaneamente (sob um ponto de vista comportamental), de modo que não é tão fácil (ou sequer possível) discernir quando o susto passou e quanto a ação começou.

disponibilize o acesso à hipótese correspondente às chaves, ele não se limita apenas àquilo possibilitado pelas previsões sobre a mesma, pois, durante a busca pelas chaves, esboços indiretamente relacionados também podem permitir a captura de informações úteis para o sistema por meio da atenção exógena (por exemplo, durante a busca, podemos notar que existe um pedaço de pizza debaixo da cama). Curiosamente, isso implica uma desregulação dos estados perceptivos de nível inferior, uma vez que, em casos como esses, talvez seja necessário aumentar a intensidade dos sinais de erro relevantes para que a busca das chaves continue. Em outras palavras, buscar as chaves efetivamente implica no sistema fixar algumas parcelas da hierarquia preditiva em estados que não estão em coerência com a entrada sensorial desejada, de modo que os erros de previsão esperados só possam ser “explicados” pela hipótese de ordem superior “chaves”. Isso significa que *priors* a respeito das chaves influenciam o sistema endogenamente a causar um efeito *pop-out* quando as chaves são, finalmente, encontradas (DOŁĘGA, DEWHURST, 2021, p. 18).

A atenção endógena parece determinar o conteúdo da consciência quando a hipótese anteriormente eleita, cuja influência é, aqui, maior do que em outros contextos perceptivos, é corroborada pelas entradas sensoriais restritas pelo próprio sistema. É válido ressaltar que, no caso assinalado, houve uma razão pela qual a hipótese ou o modelo referente às chaves foi o eleito para coordenar a busca: essa busca só ocorreu porque, anteriormente, as chaves não foram encontradas. Isso significa que o acionamento da atenção endógena também se dá com base nos diversos contextos nos quais habitamos e, a nível pessoal, podemos relacioná-la à introspecção. Ou seja, suponha que o sujeito procure suas chaves onde geralmente as deixa, mas não as encontra. Isso faz com que ele realize uma “busca interna” dos locais onde poderia tê-la deixado, tente resgatar narrativas acerca do trajeto que ele realizou da última vez que esteve com as chaves, etc. O fato dele não encontrar o objeto desejado no local onde ele supostamente estaria fez com que o sujeito realizasse uma introspecção – isto é, direcionasse sua “atenção interna”, a nível pessoal. Logo, da mesma forma que (conforme vimos anteriormente) os níveis inferiores da cognição influenciam os superiores, os níveis superiores também influenciam os inferiores (e, assim, sucessivamente): um ato de introspecção gera cascatas de erros de previsão que irão promover os ajustes necessários e, com isso, o plano de ação de busca do objeto perdido.

Com isso em mente, os sistemas de sondagem (atencionais) endógenos e exógenos parecem ser aquilo que determina quais previsões ou esboços irão compor os modelos ou hipóteses conscientemente percebidos. Em ambos os casos, a sonda é atraída por previsões ou

esboços altamente influentes para o sistema em um determinado momento. Em convergência com aquilo anteriormente desenvolvido nessa dissertação, Dołęga e Dewhurst esclarecem, ainda, que o critério para que uma hipótese vencedora seja sondada e, assim, alcance a percepção consciente se deve à sua condição de ser “*probe-able*”, ou à probabilidade de ser, simultaneamente, a mais coerente com os contextos interno e externo do indivíduo (2021, p. 19). Entretanto, sabemos que a consideração de apenas esse critério não seria o suficiente para que uma hipótese seja conscientemente percebida, dado que hipóteses coerentes podem nos fazer funcionar no “piloto automático” por períodos significativos até que nossa atenção seja requerida.

Desenvolvendo o anteriormente dito, para o Processamento Preditivo, a atenção exógena induz a atualização das hipóteses vencedoras via inferência perceptiva. Por sua vez, a atenção endógena induz tanto a um resgate interno de informações quanto a realização de inferências ativas – trazendo o meio para mais próximo daquilo previsto. No entanto, na Seção 2.1.4, vimos que, em um nível mais profundo de análise, ambas essas formas de inferências são, na realidade, inferências ativas. Logo, toda vez que o organismo se atenta a algum evento interno ou externo, ele realiza uma forma de inferência ativa sobre seus próprios estados. Dessa maneira, atentar-se a uma hipótese implica proativamente aumentar sua precisão e probabilidade, alterando-a e fazendo com que ela permaneça ativa por tempo e possua “força” o suficiente para assumir temporariamente o controle do sistema e, assim, disponibilizar relatos – verbais ou não. Ou seja, é a realização de uma inferência ativa ou, em outros termos, de uma sondagem ativa (aqui entendida como uma auto-sondagem) que, portanto, determina qual dentre as hipóteses influentes na hierarquia neural se tornará a vencedora e será eleita para a percepção consciente (DOŁĘGA, DEWHURST, 2021, p. 20).

Se essa história estiver correta, então, a todo instante e em todos os níveis da hierarquia neuronal, inúmeras predições ou esboços, compondo hipóteses ou modelos passíveis de sondagem, influenciam o sistema, criando estabilidade e coerência perceptiva ao sujeito. No entanto, somente quando a atenção é acionada que nos tornamos “plenamente” conscientes do que se passou ou do que está ocorrendo em um determinado momento. Aqui, é interessante notar que mesmo que, em última instância, a atenção não seja acionada (a nível pessoal), sempre que estamos em estado de vigília, as sondagens sub-pessoais endógenas e exógenas, constantemente ativas, permitem-nos “perceber inconscientemente” as florestas sem necessariamente nos atentar às árvores. Para o Processamento Preditivo, a “percepção” de tais florestas seria dada majoritariamente por *priors* acerca de sua localização e expectativas

sobre sua estrutura e composição. Assim, quando eventos surpreendentes nos ocorrem, nosso sistema não necessita se esforçar desmedidamente para sondar a totalidade do conteúdo de hipóteses que antes eram improváveis de acontecer – como, por exemplo, uma onça aparecer em meio às árvores. Isso significa que as sondagens endógenas e exógenas são capazes de aumentar a fama de parcelas de hipóteses anteriormente improváveis, unindo-as aos *priors* e, então, trazendo à consciência uma narrativa contínua acerca dos eventos dos últimos instantes – consequentemente, passando-nos a sensação manifesta de que sempre estivemos cientes de tudo (DENNETT, 1991, p. 407-408).

Logo, se considerarmos que predições equivalem ou podem ser tratadas como esboços dennettianos e que o mecanismo responsável pelo processo de sondagem equivale ou pode ser tratado como o mecanismo preditivo da atenção sub-pessoal endógena e exógena, o problema de como inferências inconscientes chegariam à consciência parece ser minimizado. Não somente isso, talvez possamos inclusive resolver o problema da percepção inconsciente, em termos gerais. Como Dołęga e Dewhurst notaram, casos de visão cega podem ser plenamente interpretados em termos dessa inter-relação entre o Processamento Preditivo e o MME. Especificamente, tais pacientes seriam capazes de formular predições e hipóteses funcionalmente aptas a responder às tarefas visuais de seu dia-a-dia, mas que não seriam passíveis de sondagem ou ajuste de precisão por parte do sistema – devido a patologias ou danos neurológicos. Ainda, seria por esse mesmo motivo que indivíduos com visão cega não são plenamente capazes de formular avaliações subjetivas de sua própria confiabilidade – isto é, eleger predições acerca de seus próprios estados (DOŁĘGA, DEWHURST, 2021, p. 20).

3.4 A unicidade da percepção consciente

O trato da atenção como responsável pelo funcionamento do mecanismo de sondagem no Processamento Preditivo também elucida uma possível solução para o problema da unicidade da percepção consciente. Esse problema (também conhecido como “*binding problem*”) consiste em questionar como múltiplas informações sensoriais poderiam resultar, na percepção consciente, em cenários completos e estáveis ao longo do tempo. Ou seja, como é que, digamos, o ato de beber um café consegue nos ser percebido como sensorialmente integrado – como o tato, o olfato, a visão e o paladar, durante o ato de beber um café são percebidos em sincronia? Para William James, a consciência é por si só uma coisa integrada e não feita de partes, correspondendo à totalidade de atividades do cérebro (1890, i, p. 177).

Ainda, para o filósofo Thomas Metzinger, essa unidade não é exclusiva da nossa percepção, pois ela também existe (de alguma forma independente de nós) no mundo e a nossa experiência da totalidade desse mundo é indubitável (2009, p. 27; 1995, p. 249). Todavia, conforme o mencionado no Capítulo 1, não há local algum no cérebro onde toda a informação sensorial se junta para ser passivamente assistido por algo ou alguém. Todas as partes de nosso sistema nervoso trabalham paralelamente e em distintos níveis de hierarquia espaço-temporal (Cf. DENNETT, 1991). Então, como essa indubitável unicidade poderia surgir?

De acordo com Susan Blackmore e Emily Troscianko (2018), existem pelo menos seis formas de abordar o problema da unicidade da consciência. A primeira delas é denominada unicidade por sincronização e alega que o motivo pelo qual tal unicidade perceptiva ocorreria consiste na atuação de disparos neuronais sincronizados (Cf. CRICK, KOCH, 1990, 2003; ENGEL et al., 1999; SINGER, 2000; ENGEL, 2003;). Por sua vez, a segunda consiste na chamada “micro-consciência”, a qual alega que a unicidade é uma característica perceptiva possibilitada pela linguagem, pois possuímos múltiplas micro-consciências ou nódulos de processamento em nosso sistema nervoso cujo *output* é integrado apenas durante o processamento linguístico (Cf. ZEKI, 2001, 2007). A terceira forma consiste na integração multissensorial, a qual alega que os neurônios geralmente respondem a disparos de mais de um sensor ou modalidade sensorial e que tal característica é otimizada ao longo do tempo, de modo a possibilitar o discernimento subjetivo das estimulações sensoriais e, ao mesmo tempo, que experienciamos os eventos ao nosso redor de forma unificada (Cf. BEIERHOLM, QUARTZ, SHAMS, 2009; FULKERSON, 2014). A quarta forma se baseia na Teoria da Informação Integrada, alegando que a consciência corresponde justamente à capacidade inata de certos sistemas integrarem informação (Cf. TONONI, 2004, 2008). A quinta forma seria a estruturação da unicidade por meio da ação, a qual afirma que a unicidade da consciência emerge desde um fluxo dinâmico entre processos causais de nível inferior e múltiplos loops de *feedback* conectando entradas e saídas sensoriais (Cf. HURLEY, 2001; O’REGAN, NOË, 2001). Finalmente, a sexta e última forma de unicidade perceptiva consistiria em uma ilusão, alegando que não há um momento no qual as várias modalidades se juntam e que, na maioria das vezes, a experiência una de primeira pessoa sequer existe – com um *self*, munido de percepção e experiências unificadas, surgindo apenas quando nos atentamos a alguma coisa (Cf. DENNETT, 1991; BLACKMORE, 2002).

De certo modo, embora essa dissertação se alinhe predominantemente com a última, no presente contexto, faz-se possível assumir algum grau de todas as 6 formas possíveis de

unicidade da consciência. Mais especificamente, para as ferramentas teóricas aqui descritas, não haveria problemas em falar de sincronização neuronal, afinal, toda vez que nos valemos de termos tais como modelos e representações, assumimos, de uma forma ou de outra, que existiria algo como uma rede de neurônios que disparam em sincronia na formação de um determinado padrão de ativação. Entretanto, talvez não seja o caso de assumirmos que essa sincronia seja necessariamente a responsável pela unicidade da experiência perceptiva. Ainda, se for possível equiparar, a grosso modo, as micro-consciências com os esboços preditivos, também não parece haver grandes problemas em assumir que a complexa unicidade manifesta de nossa percepção seja o fruto da influência de processos reflexivos – embora estes não precisariam estar apenas relacionados à linguagem, podendo ser o resultado de qualquer forma de introspecção. Por fim, a integração multissensorial e a Teoria da Informação Integrada podem ser totalmente assumidas internamente aos contextos do PP e do modelo dennettiano, conforme o desenvolvido por Clark (2016) e Seth (2021). Contudo, parece-me que as duas últimas formas de unicidade são aquelas que melhor expressam o que a inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo parece implicar.

Como acabamos de ver, a atenção ou os sistemas de sondagem endógenos e exógenos determinam o conteúdo consciente via inferência ativa. Agora veremos o papel da ação na unicidade da consciência. Hohwy argumenta que nossa percepção é unificada porque se baseia em hipóteses oriundas de um modelo hierárquico causalmente estruturado. Para ele, apenas uma única hipótese é necessária para gerar a unidade perceptiva, visto que inferências ativas impediriam que, a qualquer momento, mais de uma hipótese fosse utilizada como base para a amostragem sensorial (HOHWY, 2013, p. 217). Mais especificamente, ele se vale da lógica de funcionamento do *Global Workspace Theory* (GWT)⁵⁴ para detalhar como a hipótese vencedora ultrapassaria o limiar de uma “ignição global”, alcançando, portanto, a influência exclusiva no sistema e, com isso, orientando a ação e a tomada de decisão, além da possibilidade de ser relatada introspectivamente (HOHWY, 2013, p. 213). Nesse contexto, tal

⁵⁴ “A transmissão global permite que as informações sejam processadas com mais eficiência (porque não estão mais confinadas a um subconjunto de circuitos não-conscientes, mas podem ser compartilhadas com flexibilidade por muitos processadores corticais) e relatadas verbalmente (porque esses processadores incluem aqueles envolvidos na formulação de mensagens verbais). Estímulos não-conscientes podem ser processados de forma rápida e eficiente ao longo de rotas de processamento automatizadas ou pré-instruídas antes de decair rapidamente em alguns segundos. Em contraste, os estímulos conscientes seriam distinguidos por sua falta de “encapsulamento” em processos especializados e sua circulação flexível para vários processos de relato verbal, avaliação, memória, planejamento e ação intencional, muitos segundos após seu desaparecimento. [Nós] postulamos que essa disponibilidade global de informações é o que experimentamos subjetivamente como um estado consciente” (DEHAENE, CHANGEUX, 2011, p. 210 [tradução nossa]).

ignição seria um ponto hipotético no qual o sistema se torna “satisfeito” com a probabilidade posterior de uma hipótese, viabilizando uma mudança de inferência perceptiva para inferência ativa e, então, a produção de previsões proprioceptivas que mediam a ação. Digno de nota, a ignição seria algo integralmente dependente do contexto, o que implica que, aqui, as ações planejadas pelo sistema e sua relevância serão importantes, assim como as precisões esperadas para os erros de predição e a “confiança” que o sistema possui em suas hipóteses vencedoras e suas concorrentes (HOHWY, 2013, p. 215). Para Hohwy, portanto, a causa da unicidade da percepção seria a execução de inferências ativas no contexto de uma ignição global.

No entanto, mencionei anteriormente os problemas de assumir apenas uma hipótese como a vencedora e que isso seria um limitador significativo para a proposta de Hohwy. Além disso, na formulação hohwyana, uma vez que a hipótese vencedora ultrapassa o limiar de ignição, ela ficaria disponível para um outro mecanismo “projetar” seu conteúdo, o que implicaria a existência de algo como um Teatro Cartesiano. É possível evitar que a lógica de funcionamento do GWT tenha tal implicação caso compreendamos a ignição global como o próprio estado de consciência. Isto é, depois que a hipótese vencedora ultrapassa o limiar de ignição global, ela não fica disponível para acesso consciente, ela *é* a consciência perceptiva. No cenário aqui proposto há ainda uma significativa convergência adicional: os múltiplos esboços passíveis de sondagem ganham influência no sistema e, ao serem sondados ou atentados, são globalmente computados, o que equivaleria à percepção consciente. E, como estamos supondo uma multiplicidade de esboços passíveis de sondagem, a competição pela “fama” implica a ativação parcial de diversos esboços semelhantes e relacionados uns com os outros, ampliando as possibilidades daquilo que será elevado à consciência. Ou seja, para uma hipótese ganhar fama no sistema, ela precisa ser compatível com outras hipóteses passíveis de sondagem, além de instigar a ativação de hipóteses relacionadas a ela mesma. Conforme foi notado por Dennett, a influência que uma hipótese exerce sobre o sistema precisa ser sustentada pelas hipóteses competidoras por um determinado período para que tal influência se torne global (DENNETT, 2006, p.137). Logo, seria graças a essa competição entre hipóteses passíveis de sondagem que a unicidade da percepção pode ser considerada como ilusória.

Como foi dito no início desta dissertação, ilusão não significa inexistência de algo, mas sim algo que parece ser aquilo que não é. A descrição da percepção consciente, em termos da inter-relação entre o Processamento Preditivo e o Modelo dos Múltiplos Esboços, deixa muito claro que não faz sentido pensar que, o tempo todo, existiria um fluxo central ou

projeção unificada da consciência. As inúmeras atualizações dos modelos gerativos, ocorrendo a cada segundo subjetivo, permitem a manutenção da ativação de determinados circuitos neuronais e equivalem àquilo que, em última instância, experienciamos em nossa percepção consciente. E, se essa experiência perceptiva não se assemelha, em termos de padrões neuronais, ao que nos aparenta ser em termos manifestos, logo, ela poderia ser tratada como ilusória. Curiosamente, na literatura do Processamento Preditivo, grande parte dos autores preferem utilizar o termo “alucinação” ao invés de “ilusão”, valendo-se do famoso slogan de que “a percepção é uma alucinação controlada”⁵⁵. Porém, dado que ele coloca em dúvidas a questão de se há qualquer coisa para além das nossas limitações sensoriais, o uso do termo não somente é contraintuitivo, mas remete a uma visão deveras intelectualista da cognição – ou, em casos mais extremos, idealista. Para Hohwy, a percepção consciente:

[S]urge como resultado do apetite do cérebro para fazer o melhor sentido possível da entrada sensorial atual, mesmo que isso signifique dar muito peso às crenças anteriores. Isso se encaixa na ideia de que a experiência consciente é como uma fantasia ou realidade virtual construída para manter a entrada sensorial afastada. É diferente da experiência consciente que é verdadeiramente uma fantasia ou realidade virtual, que desfrutamos em imagens mentais ou sonhos, porque tais experiências não se destinam a manter a entrada sensorial à distância. Mas estão, no entanto, a uma distância do mundo real que está representando. (HOHWY, 2013, p. 137-138 [tradução nossa]).

Similarmente, para Seth:

O panorama multissensorial imersivo de sua cena perceptiva, aqui e agora, é uma extensão do cérebro para o mundo, uma escrita tanto quanto uma leitura. A totalidade da experiência perceptiva é uma fantasia neuronal que permanece atrelada ao mundo por meio de um contínuo fazer e refazer das melhores suposições perceptivas, de alucinações controladas. Você poderia até dizer que estamos alucinando o tempo todo. É que quando concordamos sobre nossas alucinações, isso é o que chamamos de realidade (SETH, 2021, p. 87 [tradução nossa]).

No entanto, Clark desafia essa ideia de alucinação controlada ou realidade virtual apontando dois erros. O primeiro deles é conceber mapas inferenciais para respostas adaptativas como uma espécie de “véu representacional” que separa o sujeito de seu mundo. Para Clark, esses mapas, na realidade, seriam apenas um *know-how* probabilístico estruturado através da aprendizagem orientada para previsões progressivamente mais otimizadas, o qual nos permitiria perceber, através de um “véu estatístico”, o mundo das próprias causas das

⁵⁵ No caso, uma “alucinação” por partir de inferências realizadas desde modelos preditivos e “controlada” por depender dos *feedbacks* que o meio interno e externo fornecem a tais modelos.

interações distais que afetam nossos sentidos. Por sua vez, o segundo erro seria a negligência da importância do papel da ação nas dinâmicas preditivas dos organismos, tanto nos processos de amostragem quanto nos contínuos testes preditivos ocorrendo nas múltiplas camadas da hierarquia neuronal. Assim, em vez do objetivo do PP ser um tipo de imagem neutra do mundo, para Clark, as dinâmicas de nossos cérebros preditivos corporificados oferecem a possibilidade de controle sobre as *affordances* do ambiente:

Tomados em conjunto, esses pontos sugerem que o mecanismo de inferência probabilística no cérebro não constitui uma barreira entre o agente e o mundo. Em vez disso, fornece uma ferramenta única para encontrar um mundo de significado, povoado por affordances. (CLARK, 2016, p.170 [tradução nossa]).

Com isso em mente, uma vez que: (1) as atualizações dos modelos gerativos resultam em uma percepção manifesta estável e una; (2) precisamos agir no mundo para que esses modelos sejam atualizados; e (3) nossa ação no mundo depende estritamente daquilo que nos é possível realizar; parece sensato concluir que a consciência manifesta, conforme sua descrição pela inter-relação entre o MME e a PP seja uma forma de ilusão orientada às *affordances*. E, conforme veremos a seguir, essa conclusão possui óbvias implicações para o nosso entendimento das experiências subjetivas, em sua totalidade.

3.5 A imagem manifesta preditiva e orientada às *affordances*

Sendo esta a última seção do presente capítulo, dedico-a à exploração da subjetividade de nossa percepção consciente, algo que nem Dennett (em seus escritos sobre o MME de 1991), nem Clark (2016) ou Hohwy (2013) se aventuraram a elaborar ou descrever de forma robusta – ao passo que Seth (2021), dedicou seu último livro ao *self* em termos do Processamento Preditivo. Considerando os pressupostos adotados por essa pesquisa, aqui, procurarei desenvolver, em primeiro lugar, como Frankish (2017) concebe a possibilidade de uma descrição fisicalista e ilusionista do que seria a subjetividade (ou seja, as *qualia*), seguida de uma descrição preditiva do surgimento e funcionamento da subjetividade, conforme o elaborado por Seth, e uma breve exposição da inter-relação entre alguns dos *insights* de ambos esses autores nas ideias centrais do recente texto de Clark, Wilkinson e Friston, acerca do mesmo tema (2019).

De acordo com Frankish (2017), a subjetividade seria composta por propriedades físicas “quase-fenômênicas” (ou zero *qualia*) de representações que, através do ato de

introspecção, seriam processadas como se fossem fenomênicas (p.18). Mais especificamente, para Frankish, as propriedades quase-fenomênicas são:

uma propriedade física não-fenomênica (talvez uma propriedade complexa e manipulada) que a introspecção tipicamente deturpa como fenomênica. Por exemplo, a vermelhidão quase-fenomênica é a propriedade física que normalmente desencadeia representações introspectivas de vermelhidão fenomênica. Não há nada de fenomênico em tais propriedades – nada “sensível” ou qualitativo – e elas não apresentam nenhum problema explanatório especial. (FRANKISH, 2017, p. 19 [tradução nossa]).

Em outras palavras, propriedades quase-fenomênicas parecem encadear um determinado efeito durante nossos atos de introspecção e, com isso, gerar a sensação manifesta de subjetividade. Logo, quando percebemos visualmente um objeto, nosso aparato sensorial capta uma série de informações acerca dele, as quais serão processadas e editadas, de modo que, caso a introspecção busque as representações caricaturadas que formamos acerca desse objeto, elas acabariam por transmitir “erroneamente” (como fenomênicas) suas características, em conjunto com o próprio conteúdo perceptivo.

A perspectiva ilusionista de Frankish demonstra que a ilusão seria, então, achar que as *qualia* são oriundas das informações captadas no ambiente e não da própria introspecção. Aqui, em vez de entendermos que aquilo que experienciamos como fenomênico se deve às nossas próprias propriedades físicas, como seres corporificados, a ilusão estaria em presumir que esse sentimento, essa sensação qualitativa, seria uma espécie de poder dos objetos observados de criar impactos sobre nós. Isto é, a vermelhidão seria ilusoriamente representada como um poder da superfície do tomate de nos afetar de alguma forma. Assim, o que essa ilusão parece implicar é que a maneira mais fácil de resgatarmos e transmitirmos experiências que nos foram significativas (no sentido popular do termo) é selecionar os objetos que nos causaram impactos significativos (no sentido biológico) e atribuí-los propriedades que nos pareçam significativas (Cf. LEAL-TOLEDO, De VASCONCELOS, 2022).

Isso posto, vimos que a unicidade da consciência consiste em uma ilusão criada pelos contínuos ciclos de atualização dos modelos gerativos/esboços, sendo, portanto, uma espécie de colcha dinâmica de retalhos informacionais que, de forma síncrona, ficaria ativada por tempo suficiente para ser notada pelos sistemas de sondagem. E, independentemente de sabermos (ou pelo menos acreditarmos) que essa unicidade é ilusória, não deixamos de percebê-la dessa maneira (caso consideremos, obviamente, condições neuro-normativas). Intuitivamente, quando se fala nessa percepção, se fala em termos sensoriais, ou seja, em cheirar, tocar, ouvir e ver. Essa é a exterocepção – as percepções do mundo externo.

Entretanto, ao destacar o papel da ação no PP, destaquei também o papel da propriocepção – isto é, a percepção de nossos corpos nos contextos nos quais habitamos. Juntas, exterocepção e propriocepção são aquilo que possibilita não somente a percepção dos objetos, mas a percepção de suas *affordances*: ou seja, não somente percebemos as propriedades de uma xícara, mas percebemos suas possíveis funções e as possibilidades de interação que fornecem para cada um de nós (as quais não seriam necessariamente as mesmas, dado os usos não-convencionais que às vezes damos aos objetos e as particularidades de cada organismo).

Contudo, para Seth, a grande responsável por nossa subjetividade é a interocepção. Em termos gerais, ela consiste na percepção das atividades internas ao corpo, permitindo que lidemos com nossas condições fisiológicas e transmitindo informações dos órgãos internos para o sistema nervoso, de modo que ele possa atuar na regulação de tais condições e manter o organismo vivo. A interocepção atua na percepção, por exemplo, dos batimentos cardíacos, da respiração, do nível de açúcar no sangue, etc., mas ela também é a responsável pela percepção das *affordances* do próprio organismo: pois não somente percebemos indiretamente os níveis de açúcar em nosso sangue, mas sentimos vários dos sintomas relacionados às flutuações nesses níveis e vontades de atuar no mundo para regularizá-los, caso necessário. De acordo com Seth (2021, p. 176), enquanto as predições do mundo externo servem para prever “o que está lá fora”, as predições do mundo interno servem para regular o organismo e, em última instância, gerar expectativas sobre o que faremos a seguir, o que pensaremos a seguir e o que esperamos a seguir (DENNETT, 2017, p.10).⁵⁶

As predições interoceptivas prevêm, então, as consequências internas das ações do organismo, a fim de auxiliar na regulação fisiológica. As variáveis fisiológicas, as quais determinam e regulam os parâmetros para nossa existência, são as causas ocultas dos sinais interoceptivos. Essa forma de controle preditivo pode fomentar respostas antecipatórias, através de predições sobre os estados futuros de nossos corpos e suas consequências. Geralmente, nossos modelos gerativos internos precisam possuir *priors* dos sinais interoceptivos com alta precisão, para que suas predições possam se auto-realizar por meio da inferência ativa. Esse aspecto-chave é o que garante que as melhores predições interoceptivas sejam eleitas para as flutuações fisiológicas esperadas. Por exemplo, graças à atuação da inferência ativa, prever que a temperatura de nossos corpos se mantém constante ao longo do

⁵⁶ É válido notar que Seth atribui a capacidade de percepção das *affordances* apenas às inferências interoceptivas, dado que elas atuam de maneira a controlar e regular o organismo (2021, p.177), ao passo que eu atribuo a percepção de *affordances* também à exterocepção e a propriocepção.

tempo, faz com que, de fato, ela permaneça constante – auto-realizando tal predição e, assim, garantindo nossa regulação fisiológica.

O cérebro, então, é capaz de minimizar suas expectativas interoceptivas prévias das condições do corpo para garantir que nossas variáveis fisiológicas se mantenham estáveis, ultimamente implicando que, enquanto estivermos vivos, o cérebro jamais atualizará seu *prior* de que estamos vivos (SETH, 2021, p. 181). Isso implica uma experiência corporal manifesta relativamente imutável, dado que fortes predições sobre estados corporais estáveis geram uma atenuação da detecção de flutuações nesses estados, para assegurar a otimização das variáveis fisiológicas significativas. Seria como se não pudéssemos perceber nossas condições fisiológicas mudando quando elas, de fato, mudam. Assim, o organismo, em última instância, se percebe como estável através do tempo, o que acaba por nos auxiliar a perceber variações mais inesperadas em nossos corpos e controlá-las.

Quando um evento inesperado ocorre, a percepção interna do organismo inevitavelmente captura um erro significativo de predição, dada a alta possibilidade desse evento colocar sua regulação interna em risco. Isso provoca divergências entre aquilo previsto pelos modelos gerativos internos e aquilo que os sensores externos implicam. Para minimizar tais divergências, o cérebro primeiro aumenta o ganho de erros vindos da estimulação externa para garantir que tais erros constem no modelo preditivo externo do organismo, o que, por sua vez, faz com que o organismo se modifique, tanto externa quanto internamente (ao, por exemplo, provocar o aumento de seu batimento cardíaco, tremores, suor, a dilatação da pupila, etc.). A título de exemplo, o desequilíbrio que isso acaba por gerar entre os modelos internos e externos poderá ser computado como “medo”, o que induzirá a realização de ações para tentar amenizá-lo, as quais passarão a ser incorporadas nesses modelos e influenciarão as predições subsequentes, em um circuito de *feedback*. E o que isso implica é que não fazemos coisas como chorar porque estamos tristes, mas ficamos tristes porque percebemos o nosso estado de choro (SETH, 2021, p. 183). Com efeito, mesmo que esses estados corpóreos estejam enraizados em nossas regulações fisiológicas, tendemos a experienciá-los subjetivamente, pois são incorporados como erros em nossos modelos de nós mesmos. Curiosamente, essa experiência está sempre conectada com o contexto no qual habitamos, pois, um organismo não sente, simplesmente, coisas como o medo ou a tristeza, mas medo de algo e tristeza acerca de algo. Isso, mais uma vez, demonstra a forte conexão existente entre nossos corpos e seu meio, evidenciando ainda mais a existência e centralidade de um ciclo contextualizado de ação-percepção.

Para Seth, então, as inferências interoceptivas seriam as maiores responsáveis por nossas sensações subjetivas. A ilusão de haver uma perspectiva de primeira pessoa é o resultado de inferências altamente precisas e autorealizáveis dos próprios estados do corpo e quaisquer alterações nestes acarretam mudanças diversas nessa perspectiva, como a oscilação de humor e impulsos aleatórios. Ainda, embora os exemplos aqui fornecidos se refiram apenas às emoções, após uma análise mais profunda, Seth alega não haver distinções significativas entre as emoções e as demais sensações reunidas naquilo que ele chama de “*selfhood*”, uma vez que, dentro do contexto do Processamento Preditivo, podemos também reduzi-las às inferências interoceptivas.

Finalmente, um desenvolvimento adicional para essa narrativa acerca da subjetividade foi recentemente fornecido por Clark, Friston e Wilkinson (2019). Seguindo uma linha de raciocínio semelhante à de Seth, os autores também concordam que as sensações subjetivas (isto é, *qualia*) são causas inferidas e construídas por nossas predições de maneira a melhor explicar o fluxo sensorial e o processamento hierárquico induzidos pelo ciclo ação-percepção. Porém, eles acrescentam que existe uma certeza incomum acerca dessas causas e que, para eles, isso teria aberto espaço para confusões teóricas tais quais a que gerou os Problemas da Consciência e as crenças na existência de uma lacuna epistêmica. Com isso em mente, Clark, Friston e Wilkinson afirmam que as *qualia* seriam, assim como os demais elementos de nossa percepção consciente, causas inferidas, fazendo parte daquele conjunto de inferências que melhor e mais consistentemente explicam e predizem o fluxo de informações sensoriais (CLARK, FRISTON, WILKINSON, 2019, p.21).

Os autores mencionam que os seres humanos possuem modelos gerativos hierarquicamente organizados, nos quais a atualização das hipóteses de um dado nível depende das hipóteses do nível imediatamente superior. Assim, os níveis intermediários atuam como *priors* para os níveis inferiores, de modo que tal hierarquia preditiva está repleta de “hipóteses sobre hipótese”, emparelhadas e gerando aquilo que Schwarz (2018) chamou de “fundamentos imaginários” – cujo propósito é se apresentar como uma previsão extremamente precisa, isto é, com grande certeza, ainda que não-devidamente comparável com os erros de predição. Esses fundamentos imaginários ou hipóteses intermediárias parecem permitir aos organismos realizar inferências imaginárias e/ou contrafactuais com níveis elevados de certeza, o que lhes afasta perigosamente de sua realidade ao atuar na completa ausência de *feedbacks* sensoriais diretos. Além disso, eles parecem ser capazes de gerar grande confusão quanto à relação entre padrões internos e inferências acerca de estados externos – de modo que

coisas como a vermelhidão do tomate, por exemplo, aparente-nos tão real quanto o próprio tomate (CLARK, FRISTON, WILKINSON, 2019, p. 23, 29).

Nisso, a função das hipóteses intermediárias estaria relacionada à antecipação de eventos complexos, pois, quanto mais previsões distintas um modelo gerativo englobar, mais ações adaptativas ele será capaz de gerar, otimizando, portanto, a interação do organismo com seu meio. Digno de nota, em cenários como esse, além de auto-promover os próprios padrões reativos do sistema, a realização de inferências interoceptivas tende a impactar tanto a exterocepção quanto a seleção da ação, fazendo com que tenhamos certeza de que a vivência de estados subjetivos tais como como a fofura de um bebê, seja-nos tão real quanto o próprio bebê (DENNETT, 2017, p. 210). Além disso, os fundamentos imaginários são capazes de controlar e inferir a precisão de demais previsões (por meio da atenção endógena) e até mesmo fazer com que o modelo gerativo altere os níveis de confiança que possui sobre o próprio organismo. Assim, percebemos que as qualia refletem informações altamente precisas que, no entanto, falham em ordenar disposições e estados específicos, tanto para corpo quanto para o mundo e, conseqüentemente, faz sentido concluirmos que esses estados de “alta certeza” mais condizem com “as aparências do que com a realidade” (CLARK, FRISTON, WILKINSON, 2019 p.27). E graças a complexidade de nossos modelos gerativos,

nossas capacidades distintas de perplexidade surgem porque, graças à profundidade e complexidade de nosso modelo gerativo, somos capazes de ver que esses agrupamentos (o vermelho dos objetos, a fofura de alguns animais) refletem informações altamente certas que, no entanto, falham em determinar maneiras específicas para o mundo externo (ou corpo) ser. Tornamo-nos assim conscientes de que estes estados, conhecidos com grande certeza, parecem pertencer ao lado ‘aparência’ de uma divisão aparência/realidade (Cf. ALLEN, 1997 apud CLARK, FRISTON, WILKINSON, 2019 p. 29 [tradução nossa]).

Resumidamente, talvez possamos tratar as *qualia* como o reflexo de codificações sensoriais intermediárias, as quais nos surgem como aptas para a seleção de ação e/ou repletas de informações interoceptivas. Por serem altamente precisas, sempre que realizamos uma introspecção, essas codificações ou estados intermediários seriam corriqueiramente confundidos com algo que iria além do próprio percepto. Além disso, trata-se de algo passível de manipulação, uma vez que a corroboração de *feedbacks* sensoriais não seria um pré-requisito para sua existência e manutenção. Dito isso, após sua exposição, Clark, Friston e Wilkinson reconhecem que talvez tenham deixado o Problema Difícil, em seu sentido estrito, intacto, mas chegam a uma conclusão positiva: alegando que o PP e demais aparatos teóricos e

empíricos existentes aparentam já serem capazes de estruturar descrições satisfatórias acerca das qualia (CLARK, FRISTON, WILKINSON, 2019 p. 30).

4. Considerações finais

Atualmente, os inúmeros estudos sobre a consciência parecem demonstrar que estamos cada vez mais próximos de uma compreensão significativa acerca desse fenômeno que é responsável por nos fornecer uma rica e complexa experiência de vida. Se partirmos do pressuposto de que todas as nossas vivências e suas raízes neuronais corporificadas e contextualizadas podem ser integralmente explicadas em termos físicos, acabamos abandonando as caracterizações que, à primeira vista, deixam o gap epistêmico e o Problema Difícil da consciência intactos. E, como vimos, esse abandono não é algo limitante ou pejorativo para os estudos do tema, pois é graças a ele que se torna possível alternar entre as “lentes de observação” manifesta e científica, rumo a um melhor esclarecimento daquilo que a consciência é, além daquilo que ela nos parece ser. Como resultado, evitamos nos perder em raciocínios demasiadamente complexos e (atualmente) cientificamente intratáveis. Adotar esses pressupostos desde o início da presente dissertação permitiu tratar a consciência como uma função do corpo, função essa que nos guia no mundo e nos mantém vivos.

A valer, a forma e modelo aqui escolhidos para descrever esse fenômeno estão longe de serem (plenamente) cientificamente satisfatórios e reconheço que ambos precisam ainda responder a várias questões filosóficas, mas penso ter conseguido demonstrar que seus detalhes e delineamentos gerais conseguem suprir uma narrativa teórica coerente acerca da consciência. Semelhantemente, o próprio fato de que Dennett foi amplamente criticado por seu Modelo dos Múltiplos Esboços da consciência, assim como por muitas de suas ideias ao longo dos anos, não o impediu de continuar a progressivamente desenvolver sua perspectiva até chegar a um ponto de convergência claro com o trabalho de cientistas e *frameworks* tais como o Processamento Preditivo – um fato crucial para a presente pesquisa (Cf. DENNETT, 2018; DOŁĘGA, DEWHURST, 2021). Naturalmente, seria possível partir de outros modelos que, inclusive, são mais favoravelmente reconhecidos pela literatura contemporânea (tais como a IIT, a AST, o GNWT ou a HOT), mas, considerando as intrigantes implicações diretas que o MME possui para a filosofia da mente e das ciências cognitivas – especialmente naquilo que tange aos pressupostos ilusionistas e seu posicionamento altamente contraintuitivo – e limitações de tempo e escopo, optei por não enfatizá-los e apenas mencioná-los superficialmente ao longo do texto.

Dito isso, ao contrário do que Schlicht e Dołęga (2020) parecem insinuar em sua recente revisão crítica do *framework* do PP, não me parece haver problemas em objetivar

mesclar teorias e modelos de consciência diferentes, trabalhar tais mesclas de maneira empiricamente consciente e investigar suas implicações. Afinal, um dos motivos dessa possível interação seria, justamente, demonstrar a existência de convergências e comunalidades em quadros e teses originalmente diversos – um indicador de que estamos genuinamente progredindo com o conhecimento acerca do tema (Cf. GRAZIANO et al., 2019). Contudo, teorizações científicas e inter-relações teórico-cientificamente embasadas precisam ser devidamente desafiadas para que suas limitações sejam claramente definidas e, assim, tornem-se cada vez mais precisas. Aqui, não foi possível desafiar profundamente a inter-relação estabelecida entre o MME e o PP, uma vez que o foco da pesquisa se voltou majoritariamente ao levantamento do suporte e detalhes necessários para a efetivação de tal inter-relação. Entretanto, não seria surpreendente que, uma vez em devida interação, as diversas limitações do Processamento Preditivo mencionadas ao longo do presente texto acabem por serem assimiladas pelo modelo dennettiano – pois, nesse “intercâmbio”, é altamente provável que, na medida em que o MME consiga resolver alguns dos problemas e inconsistências do PP isso, em si, acabe por também implicar a admissão e absorção de ainda outros problemas e inconsistências.

Também foi parte do objetivo desta dissertação desafiar o atual cenário filosófico de reverência aos chamados Problemas da Consciência, identificando e optando por apenas lidar com aqueles dentre esses problemas (e seus semelhantes) que fomentem estudos empíricos, em detrimento de questionamentos pura ou majoritariamente baseados em intuições. Com isso em mente, caso seja pertinente, estudos futuros poderão abordar as limitações e os desafios decorrentes dessa empreitada e, adotando um trajeto argumentativo semelhante, até mesmo efetivar o desenvolvimento de uma integração mais robusta e filosoficamente compreensiva entre o MME e o PP. Adicionalmente, estudos futuros também poderão se voltar à exploração de outras inter-relações ou possibilidades de integração no estudo interdisciplinar da consciência – demonstrando, por exemplo, como o *Attention Schema Theory* se relacionaria com o *framework* do Processamento Preditivo em sua descrição da atenção e, até mesmo, da consciência (Cf. DOŁĘGA, DEWHURST, 2019); ou o que o *Global Neuronal Workspace Theory* ganharia com os *insights* fornecidos pelo PP (Cf. WHYTE, 2019).

Outra limitação notável desta pesquisa consiste em seu foco restritivo no funcionamento do sistema nervoso humano adulto e neurotípico. A importância de tratarmos as diversidades, minorias e atipicidades na Academia e nos estudos da consciência é cada vez mais evidente e, dado o amplo espectro de suas implicações, não poderia ser diferente. De

fato, uma teoria robusta acerca da consciência deve eventualmente conseguir descrever tal fenômeno em condições diversas e em diferentes criaturas – isto é, também em não-humanos, sociedades e sistemas complexos e artificiais (Cf. SETH, BAYNE, 2022) – e, embora alguns estudos em neurociência clínica e tratamento de distúrbios mentais e transtornos psiquiátricos tenham sido mencionados ao longo do presente texto, ainda não está claro como a inter-relação entre o MME e o PP poderia funcionar dentro dessas esferas.

Em síntese, a inter-relação entre o Modelo dos Múltiplos Esboços e o Processamento Preditivo nos permitiu uma descrição de como inferências inconscientes chegam à consciência e uma possível explicação do motivo pelo qual experienciamos nossos cenários perceptivos de maneira unificada. Eventualmente, tais resoluções poderão formar a base de medidas cada vez mais confiáveis/concretas para a análise da consciência, uma vez que já parece ser possível gerar previsões heterofenomenológicas e experimentais a partir de parcelas específicas de tal descrição – ainda que o *status* experimental do Processamento Preditivo não aparente estar à par do avanços proporcionados por sua teoria (Cf. WALSH, McGOVERN, CLARK, O'CONNELL, 2020). Aqui, a grande vantagem em adotar pressupostos ilusionistas é a falta de uma ênfase nos relatos subjetivos, privilegiando outras formas mais tradicionais e tratáveis de expressão e comportamento. Por sua vez, isso também implica na indispensabilidade do reconhecimento e a necessidade de uma latente distinção científico-filosófica entre aquilo que a consciência parece ser e aquilo que ela possivelmente é.

Finalmente, a falta de evidências concretas e contundentes a favor do PP pode indicar que o argumento aqui desenvolvido não possuiria um *status* epistemológico robusto, bases empíricas sólidas e seria meramente fundamentado em um conjunto de intuições incomuns acerca do funcionamento de nossa mente. Entretanto, uma vez que o *framework* do Processamento Preditivo ainda é relativamente novo nas ciências cognitivas e que já existe uma crescente produção literária favorável às suas premissas e em prol de seu estabelecimento como novo paradigma para a compreensão da cognição humana, tais observações não parecem compor objeções reais – pelo menos não diretamente. Afinal, de maneira a conseguir compreender a cognição e o comportamento, filósofos e cientistas precisam, durante seu trabalho, realizar suposições a respeito dos propósitos mais gerais do funcionamento psicológico humano, sendo razoável, então, que qualquer candidato para esse conjunto de pressupostos – como é o caso do Processamento Preditivo – seja perfeitamente entendível e aceitável como algo abstrato, incerto e exploratório.

Sintetizando, então, o essencial de nossos esclarecimentos, conclui-se que os estados locais da consciência podem ser perfeitamente descritos como o resultado de múltiplas hipóteses preditivas, influenciando o sistema nervoso. Em termos gerais, na perspectiva manifesta, essas hipóteses constituem tudo aquilo que conseguimos vivenciar, perceber, experimentar, conhecer e sentir. Por sua vez, na perspectiva científica, elas constituem contínuos ciclos de atualização de modelos gerativos, internos e externos, numa dança dinâmica entre informações sensoriais e informações previamente armazenadas, cujo equilíbrio e harmonização se dá pela minimização das discrepância existentes entre ambas. Por fim, o fato de tanto uma concepção manifesta quanto uma científica coexistirem para a consciência não significa que essas não tratem da mesma coisa, mas apenas que se referem a um fenômeno corporificado e contextualizado que ocorre em múltiplas e diferentes escalas espaço-temporais de composição – com a “melhor” definição de consciência dependendo das lentes devidamente requeridas e utilizadas em sua investigação. Afinal, se quisermos utilizar as lentes científicas e buscar sensações subjetivas nas atividades neuronais dos sujeitos, devemos nos contentar com a “nada empolgante” realidade de que tudo que encontraremos serão redes neuronais sendo periodicamente ativadas. Observa-se, então, que uma chave para a dissolução de muitas das dificuldades existentes em se estudar a consciência parece se encontrar nessa distinção sellariana entre diferentes perspectivas de mundo e na identificação das “lentes” adequadas a serem utilizadas em cada uma de nossas investigações. Logo, uma descrição da consciência que reconheça tal distinção e saiba definir quais “lentes de observação” seus questionamentos mais propriamente requerem, parece estar especialmente apta a se desenvolver numa tese robusta, cientificamente plausível e empiricamente testável – e, portanto, trilhar um caminho mais frutífero para a exploração filosófica.

5. Referências

ADAMS, R.; BROWN, H.; FRISTON, K. Bayesian inference, predictive coding and delusions. **AVANT: Journal of Philosophical-Interdisciplinary Vanguard**, v.5, p.51–88, 2014.

ADAMS, Z., BROWNING, D. How color qualia became a problem. **Journal of Consciousness Studies**, v. 27, n. 5-6, p. 14-25, 2020.

AITCHISON, L., LENGYEL, M. With or without you: Predictive coding and Bayesian inference in the brain. **Current Opinion in Neurobiology**, v. 46, p. 219–227, 2017.

AJINA, S., BRIDGE, H. Blindsight and Unconscious Vision: What They Teach Us about the Human Visual System. **Neuroscientist**. v. 23, n. 5, p. 529-541, 2016.

AL-HAYTHAM, I. A. **The Optics of Ibn al-Haytham**. Tradução: Sabra. A. I. Londres: Warburg Institute, 1030/ 1989.

AMES, A. Visual perception and the rotating trapezoidal window. **Psychological Monographs: General and Applied**, v. 65, n. 7, 1951.

ARATHORN, D.; STEVENSON, S.; YANG, Q.; TIRUVEEDHULA, P.; ROORDA, A. How the unstable eye sees a stable and moving world. **Journal of Vision**. v. 13, n. 10, 2013.

ARMSTRONG, D. **A Materialist Theory of the Mind**. Routledge, 1968.

ARU, J., et. al. Coupling the State and Contents of Consciousness. **Frontiers in Systems Neuroscience**, v. 13, 2019.

ATAL, B. The history of linear prediction. **IEEE Signal Processing Magazine** v. 23, n.2, p. 154-158, 2006.

BARLOW, H. Possible principles underlying the transformation of sensory messages. In: ROSENBLITH, W. (Orgs.) **Sensory Communication**. Cambridge, MIT Press, p.217–34, 1961.

BAARS, B. **A Cognitive Theory of Consciousness**. Cambridge University Press, 1988.

BECHTEL, W.; HAMILTON, A. Reduction, integration, and the unity of science: Natural, behavioral, and social sciences and the humanities. In: Kuipers, T. (Org.), **Philosophy of Science: Focal Issues**. Elsevier, 2007.

BEIERHOLM, U., QUARTZ, S., SHAMS, L. Bayesian priors are encoded independently from likelihoods in human multisensory cortex. **Journal of Vision**, v. 9, n. 23, p. 1–9, 2009.

BERKES, P., ORBAN, G., LENGYEL, M., FISER, J. Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment. **Science**, v. 331, p. 83–87, 2011.

BLACKMORE, S. TROSCIANKO, E. **Consciousness: An Introduction**. Routledge, 2018.

BLACKMORE, S. There is no stream of consciousness. **Journal of Consciousness Studies**, n. 9, n. 5–6, p. 17–28, 2002.

BLOOM, P. **Descartes' baby: how the science of child development explains what makes us human**. New York: Basic Books, 2004.

BLOCK, N. On a confusion about a function of consciousness. **Brain and Behavioral Sciences**, v. 18 n. 2, p. 227–247, 1995.

BLOCK, N. Mental Paint In: HAHN, M.; RAMBERG, B. (Orgs.) **Reflections and replies: Essays on the philosophy of Tyler Burge**. MIT Press, Cambridge, p. 165-200, 2003.

BROWN, R.; LAU, H.; LEDOUX, J. Understanding the higher-order approach to consciousness. **Trends in Cognitive Sciences**, v. 23, n. 9, p. 754–768, 2019.

BRUIN, L., MICHAEL, J. Prediction error minimization: Implications for Embodied Cognition and the Extended Mind Hypothesis. **Brain and Cognition**, v.112, p.58-63, 2017.

BRUINEBERG, J., DOŁĘGA, K., DEWHURST, J., BALTIERI, M. The emperor's new Markov blankets. **Behavioral and Brain Sciences**, v.45, n. 183, 2021.

BUCCI, A.; GRASSO, M. Sleep and dreaming in the predictive processing framework. Philosophy and predictive processing. **Open MIND**. Frankfurt am Main: MIND Group, 2017.

CARRUTHERS, P. How Mindreading Might Mislead Cognitive Science. **Journal of Consciousness Studies**, v. 27, n. 7-8, p. 195-219, 2020a.

CARRUTHERS, P. Explaining the Empiricist Bias. **Journal of Consciousness Studies**, v. 27, n. 7–8, p. 230–35, 2020b.

CARVALHO, E. M. An Ecological Approach to Disjunctivism. **Synthese**. v. 198, p. 285–306, 2021.

CAO, R. New labels for old ideas: Predictive processing and the interpretation of neural signals. **Review of Philosophy and Psychology**, v. 11, n. 3, p. 517–546, 2020.

CLARK, A.; FRISTON, K.; WILKINSON, S. Bayesing Qualia: consciousness as inference, not raw datum. **Journal of Consciousness Studies**, v. 26 n. 9-10, p. 19-33, 2019.

CLARK, A. Whatever next? Predictive brains, situated agents, and the future of cognitive science. **Behavioral and brain sciences**, v. 36, n. 3, p. 181-204, 2013.

CLARK, A. **Surfing uncertainty: Prediction, action, and the embodied mind**. Oxford: Oxford University Press, 2016.

CHALMERS, D. Is the Hard Problem of Consciousness Universal? **Journal of Consciousness Studies**, v.27, n. 5-6, p. 227-257, 2020.

CHALMERS, D. The Meta Problem of Consciousness. **Journal of Consciousness Studies**, v. 25, n. 9-10, p. 6-61, 2018.

CHALMERS, D. Facing up to the problem of consciousness. **Journal of Consciousness Studies**, v. 2, n. 3, p. 200-219, 1995.

CRANE, T.; PATTERSON, S. (Orgs.) **History of the Mind-Body Problem**. Nova York: Routledge, 2000.

CRICK, F., KOCH, C. Toward a neurobiological theory of consciousness. **Seminars in the Neurosciences**, v. 2, p. 263-275, 1990.

COHEN, E.; BURDETT, E.; KNIGHT, N.; BARRETT, J. Cross-cultural similarities and differences in person-body reasoning: Experimental evidence from the United Kingdom and Brazilian Amazon, **Cognitive Science**, v. 35, p. 1282–1304, 2011.

COLOMBO, M., IRVINE, E., STAPLETON, M. (Orgs.) **And Clark and His Critics**. Oxford University Press, 2019.

COLOMBO, M., WRIGHT, C. Explanatory pluralism: An unrewarding prediction error for free energy theorists. **Brain and Cognition**, v. 112, p. 3–12, 2017.

CONSTANT, A., CLARK, A., FRISTON, K. Representation Wars: Enacting an Armistice Through Active Inference. **Frontiers of Psychology**. v. 11, n. 598733, 2021.

CORCORAN, A., PEZZULO, G., HOHWY, J. From allostatic agents to counterfactual cognisers: active inference, biological regulation, and the origins of cognition. **Biology & Philosophy**, v. 35, n. 32, p. 1-45, 2020.

DAMÁSIO, A., MEYER, K. Consciousness: An Overview of the Phenomenon and of Its Possible Neural Basis. In: LAUREYS, S., TONONI, G. **The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology**. Elsevier, 2009.

DEHAENE, S. **Consciousness and the brain**: Deciphering how the brain codes our thoughts. New York: Penguin, 2014.

DEHAENE S., CHANGEUX J. Experimental and theoretical approaches to conscious processing. **Neuron**. v. 7, n. 2, p. 200-227, 2011.

De CRUZ, H. Where Philosophical Intuitions Come From. **Australasian Journal of Philosophy**, v. 93, n. 2, p. 233-249, 2015.

De JAEGHER, H., Di PAOLO, E., GALLAGHER, S. Can social interaction constitute social cognition? **Trends in Cognitive Sciences**, v.14, n. 10, p.441-447, 2010.

DENNETT, D. Facing up to the hard question of consciousness. **Philosophical Transactions: Royal Society**, v. 373, 2018.

DENNETT, D. Illusionism as the obvious default theory of consciousness. **Journal of Consciousness Studies**, v. 23, n. 11-12, p. 65-72, 2016.

DENNETT, D. **Intuition Pumps and Other Tools for Thinking**. New York: Norton, 2013.

DENNETT, D. Consciousness cannot be separated from function. **Trends in Cognitive Sciences**, v. 15, n. 8, 2011.

DENNETT, D. **Sweet Dreams**: Philosophical Obstacles to a Science of Consciousness. The MIT Press, 2006.

DENNETT, D. **Consciousness Explained**. Londres: The Penguin Press, 1991.

DEUTSCH, D. **Musical illusions and phantom words**: How music and speech unlock mysteries of the brain. Oxford University Press, New York, 2019.

DÍAZ, R. Do People Think Consciousness Poses a Hard Problem? Empirical Evidence on the Meta-Problem of Consciousness. **Journal of Consciousness Studies**, v. 28, n. 3-4, p. 55-75, 2021.

DOŁĘGA, K.; DEWHURST, J. Bayesian Frugality and the Representation of Attention. **Journal of Consciousness Studies**, v. 26, n. 3-4, p. 38-63, 2019.

DOŁĘGA, K.; DEWHURST, J. Fame in the predictive brain: a deflationary approach to explaining consciousness in the prediction error minimization framework. **Synthese**, v. 178, p. 7781–7806, 2021.

EDELMAN, G. **The Remembered Present: A Biological Theory of Consciousness**. Basic Books, 1989.

ENGEL, A. Temporal binding and the neural correlates of consciousness. In: CLEEREMANS, A. (org.), **The unity of consciousness: Binding, integration and dissociation** (p. 132–152). New York: Oxford University Press, 2003.

ENGEL, A., FRIES, P., KÖNIG, P., BRECHT, M., SINGER, W. Temporal binding, binocular rivalry, and consciousness. **Consciousness and Cognition**, v. 8, p. 128–151, 1999.

FELDMAN, H.; FRISTON, K. Attention, uncertainty, and free-energy. **Frontiers of Human Neuroscience**, v. 4, n. 215, 2010.

FLORES, J. Hume in the Light of Bayes: Towards a Unified Cognitive Science of Human Nature, **Res Cogitans**: v. 6, n. 1, 2015.

FRANKISH, K. Quining diet qualia. **Consciousness and Cognition**, v. 21, n. 2, p. 667-676, 2012.

FRANKISH, K. (Org.). **Illusionism as a Theory of Consciousness**. Exeter: Imprint Academic, 2017.

FRIEDMAN, D.; SØVIK, E. The ant colony as a test for scientific theories of consciousness. **Synthese**, n. 198, p. 1457–1480, 2021.

FRISTON, K. Learning and inference in the brain. **Neural Networks**, v. 16, n. 9, p. 1325-1352, 2003.

FRISTON, K. A theory of cortical responses. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 360, n. 1456, p. 815-36, 2005.

FRISTON, K. Hierarchical models in the brain. **PLoS computational biology**, v. 4, n.11, 2008.

FRISTON, K. The free-energy principle: A rough guide to the brain? **Trends in Cognitive Sciences**, v. 13, n. 7, p. 293-301, 2009.

FRISTON, K. The free-energy principle: a unified brain theory? **Nature reviews. Neuroscience**, v. 11, n.2, p. 127-138, 2010.

FRISTON, K. Precision psychiatry. **Biological Psychiatry. Cognitive Neuroscience and Neuroimaging**, v. 2, p. 640–643, 2017.

FRISTON, K. Am I Self-Conscious? (Or Does Self-Organization Entail Self-Consciousness?). **Frontiers of Psychology**, v.9, n. 579, 2018.

FRISTON, K. et al. Action and behavior: a free energy formulation. **Biological Cybernetics**, v. 102, n. 3, p. 227-260, 2010.

FRISTON, K; MATTOU, J; KILNER, J. Action understanding and active inference. **Biological Cybernetics**, v. 104, n. 1-2, p. 137-160, 2011.

FRISTON, K, THORNTON, C., CLARK, A. Free-energy minimization and the dark-room problem. **Frontiers in Psychology**, v. 3, n.130, 2012.

FRITH, C. **Making Up the Mind: How the Brain Creates Our Mental World**. Oxford: Blackwell, 2007.

FUJIOKA, T.; TRAINOR, L.; LARGE, E.; ROSS, B. Beta and gamma rhythms in human auditory cortex during musical beat processing. **Annals of the New York Academy of Sciences**, v. 1169, n. 1, p. 89-92, 2009.

FULKERSON, M. Rethinking the senses and their interactions: The case for sensory pluralism. **Frontiers in Psychology**, v. 145, n. 1426, 2014.

GARDNER, H. **A Nova Ciência da Mente**. Uma História da Revolução Cognitiva. São Paulo; Edusp, 2003.

GAZZANIGA, M.; IVRY, R.; MANGUN, G. **Cognitive Neuroscience: The Biology of the Mind**, 2019.

GELMAN, S. **The Essential Child: Origins of Essentialism in Everyday Thought**, Oxford: Oxford University Press, 2003.

GELMAN, S.; WELLMAN, H. Insides and essence: Early understandings of the non-obvious. **Cognition**, v. 38, n. 3, p. 213–244, 1991.

GODFREY-SMITH, P. Environmental complexity and the evolution of cognition. In: STERNBERG, R. KAUFMAN, J. (Orgs.), **The Evolution of Intelligence**. Lawrence Erlbaum. p. 233-249, 2002.

GRAZIANO, M.; GUTERSTAM, A.; BIO, B.; WILTERSON, A. Toward a standard model of consciousness: Reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. **Cognitive Neuropsychology**, v.37, n. 3-4, p. 155-172, 2019.

GREGORY, R. L. Perceptions as hypotheses. **Philosophical Transactions of the Royal Society B: Biological Sciences** v. 290, n.1038, p. 181–97, 1980.

GREGORY, R. The confounded eye. In: GREGORY, R.; GOMBRICH, E. (Orgs.) **Illusion in nature and art**, p. 49-96. Londres, 1973.

GRIFFITHS, T.; CHATER, N.; KEMP, C.; PERFOR, A.; TENENBAUM, J. Probabilistic models of cognition: exploring representations and inductive biases. **Trends in Cognitive Sciences** v. 14, n. 8, p. 357–64, 2010.

HARLEY, T. **The Science of Consciousness: Waking, Sleeping and Dreaming**. Cambridge press, 2021.

HARKNESS, D.; KESHAVA, A. Moving from the What to the How and Where: Bayesian Models and Predictive Processing. In: METZINGER, T; WIESE, W. (Orgs.). **Philosophy and Predictive Processing: 1**. Frankfurt am Main: MIND Group, 2017.

HELMHOLTZ, H. **Handbuch der Physiologischen Optik**. Leipzig: Leopold Voss, 1867.

HEILBRON, M., CHAIT, M. Great expectations: Is there evidence for predictive coding in auditory cortex? **Neuroscience**, v. 389, p. 54–73, 2018.

HINTON, G. Learning multiple layers of representation. **Trends in Cognitive Sciences**, v. 11, n. 10, p. 428-434, 2007.

HOBSON, J.; FRISTON, K. A response to our theatre critics. **Journal of Consciousness Studies**, v. 23, n, 3-4, p. 245-254, 2016.

HOBSON, J. Consciousness, dreams, and inference: The Cartesian theatre revisited. **Journal of Consciousness Studies**, v. 21, n. 1-2, p. 6-32, 2014.

HODGSON, D. 1996. The Easy Problems Ain't so Easy. **Journal of Consciousness Studies**, v. 3, p. 69-75, 1996.

HOHWY, J. Attention and Conscious Perception in the Hypothesis Testing Brain. **Frontiers of Psychology**. v. 3, p. 1-14, 2012.

HOHWY, J. **The Predictive Mind**. Oxford: Oxford University Press, 2013.

HOHWY, J., SETH, A. Predictive processing as a systematic basis for identifying the neural correlates of consciousness. **Philosophy and the Mind Sciences**, v. 1, n. 3, 2020.

HUME, D. **A Treatise of Human Nature**. In: NORTON, D.; NORTON, M. (Orgs.). Oxford: Oxford UP, 2000.

HURLEY, S. L. Perception and action: Alternative views. **Synthese**, v. 129, n. 1, p. 3–40, 2001.

HUTTO, D., MYIN, E. **Radicalizing Enactivism: Basic Minds Without Content**. MIT Press, 2013.

JACKENDOFF, R. **Consciousness and the Computational Mind**. MIT Press, 1987.

KAHNEMAN, D.; TVERSKY, A. Prospect Theory: An Analysis of Decision under Risk. **Econometrica**, v. 47, n. 2, p. 263–291, 1979.

KAMMERER, F. The Hardest Aspect of the Illusion Problem. In FRANKISH, K. (Org.) **Illusionism as a Theory of Consciousness**, Exeter: Imprint Academic, 2017.

KANT, I. **Critique of Pure Reason**. In: GUYER, P.; WOOD, A. (Orgs.). Cambridge Press, 1998.

KING, P. Why Isn't the Mind-Body Problem Medieval? In: LAGERLUND, H. (Orgs.) **Forming the Mind**. Springer Verlag, 2007.

KIRKEBY-HINRUP, A. Two caveats to the meta-problem challenge. **Journal of Consciousness Studies**, v. 27, n. 5-6, p. 74-81, 2020.

KOK, P., DE LANGE, F. Shape perception simultaneously up- and downregulates neural activity in the primary visual cortex. **Current Biology**, v. 24, n. 13, p. 1531–1535, 2014.

LAU, H.; MICHEL, M. A socio-historical take on the meta-problem of consciousness. **Journal of Consciousness Studies**, v. 26, n. 9-10, 2020.

LEAL-TOLEDO, G. Dennett e Chalmers: Argumentos e Intuição. **Trans/Form/Ação**, São Paulo, v. 29, n.2, p. 123-132, 2006.

LEAL-TOLEDO, G.; De VASCONCELOS, M.L., O Problema da Ilusão: uma breve introdução e defesa da teoria ilusionista de Keith Frankish [Preprint], 2022.

LEWIS, C. **Mind and the world-order**: Outline of a theory of knowledge. Scribners, New York, 1929.

LEWIS, D. **Philosophical Papers**: Volume I, New York: Oxford University Press, 1983.

MARCHI, F.; HOHWY, J. The intermediate scope of consciousness in the predictive mind. **Erkenntnis**, v. 87, p. 891-912, 2022.

MARR, D. **Vision**: A Computational Investigation into the Human Representation and Processing of Visual Information, The MIT Press, 1982.

MARVAN, T.; HAVLÍK, M. Is Predictive Processing a Theory of Perceptual Consciousness? **New Ideas in Psychology**, v. 61, n. 100837, 2020.

MATSON, W. Why isn't the mind-body problem ancient? In FEYERABEND, P., MAXWELL, G. (orgs.) **Mind, Matter, and Method**: Essays in Philosophy and Science in Honor of Herbert Feigl. University of Minnesota Press, 1966.

MATURANA, H. Everything is said by an observer. In: THOMPSON, W. (Org.). **Gaia**: A way of knowing. Political implications of the new biology, p. 65-82. New York: Lindisfarne Press, 1987.

MCCARTHY, J. **Pandemonium**: A Paradigm for Learning. In: Symposium on the Mechanization of Thought Processes. H.M. Stationery Office, p.527-531, 1959.

MCCLELLAND, J.; et. al. Letting structure emerge: connectionist and dynamical systems approaches to cognition. **Trends in Cognitive Sciences**, v. 14, n. 8, p. 348-356, 2010.

MILLIDGE, B.; SETH, A.; BUCKLEY, C. Predictive Coding: a Theoretical and Experimental Review. **arXiv, Artificial Intelligence**, v.4, 2022.

MIN, B. A thalamic reticular networking model of consciousness. **Theoretical biology & medical modelling**, v. 7, n. 10, 2010.

MIRACCHI, L. 2019. None of These Problems Are That 'Hard' ... or 'Easy': Making Progress on the Problems of Consciousness. **Journal of Consciousness Studies**, v. 26 n. 9-10, p. 160-172, 2019.

MUMFORD, D. On the computational architecture of the neocortex. II. The role of cortico-cortical loops. **Biological Cybernetics**, v. 66, p. 241-251, 1992.

MURRAY, S., KERSTEN, D., OLSHAUSEN, B., SCHRATER, P., WOODS, D. Shape perception reduces activity in human primary visual cortex. **Proceedings of the National Academy of Sciences**, v. 99, n. 23, p.15164–15169, 2002.

NAGEL, T. What is it like to be a bat? **Philosophical Review**, v.83, p.435-450, 1974.

NEISSER, U. **Cognitive psychology**. New York: Appleton-Century-Crofts, 1967.

O'REGAN, J., NOË, A. A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, 24, 939–1031, 2001.

PARR, T.; PEZZULO, G.; FRISTON, K. **Active Inference**: The Free Energy Principle in Mind, Brain, and Behavior. 1^aed. The MIT Press, 2022.

PLACE, U. Is consciousness a brain process? **British Journal of Psychology**, v. 47 n. 1, p. 44-50, 1956.

PENNARTZ, C., FARISCO, M., EVERS, K. Indicators and criteria of consciousness in animals and intelligent machines: An inside-out approach. **Frontiers in Systems Neuroscience**, v. 13, n. 25, 2019.

PIEKARSKI, M. Understanding Predictive Processing: A Review. **AVANT**, v. 12, n. 1, 2021.

RAO, R.; BALLARD, D. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. **Nature Neuroscience**, v. 2, p. 79–87, 1999.

RAMSTEAD, M. Easy as 1, 2, 3: On the Short History of the Use of Affordance in Active Inference. In: DJEBBARA, Z. (Org.) **Affordances in Everyday Life**. Springer, 2022.

RAMSTEAD, M., BADCOCK P., FRISTON K. Answering Schrödinger's question: A free-energy formulation. **Physics Life Reviews**, v. 24, p. 1-16, 2018.

RATCLIFFE, M. **Rethinking Commonsense Psychology**: A Critique of Folk Psychology, Theory of Mind and Simulation. Palgrave-Macmillan, 2007.

RAUSS, K., SCHWARTZ, S., POURTOIS, G. Top-down effects on early visual processing in humans: A predictive coding framework. **Neuroscience and Biobehavioral Reviews**, v. 35, n. 5, p. 1237-1253, 2011.

RAVENS-CROFT, I. Folk Psychology as a Theory. In: ZALTA, E. (Org.). **The Stanford Encyclopedia of Philosophy**, 2021.

REDINBAUGH, M., et. al. Thalamus Modulates Consciousness via Layer-Specific Control of Cortex. *Neuron*, v. 106, n. 1, p. 66-75, 2020.

REES, G, KREIMAN G, KOCH C. Neural correlates of consciousness in humans. **Nature Reviews Neuroscience**, n.4, p. 261-270, 2002.

ROBINSON, H. Dualism. In: ZALTA, E. (Org.). **The Stanford Encyclopedia of Philosophy**, 2020.

ROCK, I. **The Logic of Perception**. Cambridge, MIT Press, 1983.

RUMELHART, D.; HINTON, G.; WILLIAMS, R. Learning representations by back-propagating errors. **Nature** v. 323, p. 533–536, 1986.

SANTOS, H.; VERISSIMO, D. A teoria de Helmholtz sobre a percepção espacial: psicofísica e filosofia transcendental. **Memorandum**, v. 34, p. 14-32, 2018.

SCHLICHT, T., DOŁĘGA, K. You can't always get what you want: Predictive processing and consciousness. **Philosophy and the Mind Sciences**, v.2, n.8, p. 1-25, 2021.

SCHWARZ, W. Imaginary Foundations. **Ergo: An Open Access Journal of Philosophy**, 2018. Disponível em: <https://philpapers.org/rec/SCHIF-6>

SCHWITZGEBEL, E. Introspection. In: ZALTA, E. **The Stanford Encyclopedia of Philosophy**, 2019.

SCHØJEDT, U; JENSEN, J. Depletion and deprivation: Social functional pathways to a shared metacognition. In: PROUST, J; FORTIER, M. (Orgs.). **Metacognitive diversity: An interdisciplinary approach**. Oxford: Oxford University Press, p. 319-342, 2018.

SETH, A. **Being You**. Faber & Fabe, 2021.

SETH, A. The Cybernetic Bayesian Brain: From Interoceptive Inference to Sensorimotor Contingencies. In: METZINGER, T., WINDT, J. (Orgs.). **Open MIND**, 35(T). Frankfurt am Main: MIND Group, 2015.

SETH, A. BAARS, B., EDELMAN, D. Criteria for consciousness in humans and other mammals. **Consciousness and Cognition**, v. 14, 2005.

SETH, A., BAYNE, T. Theories of consciousness. **Nature Reviews Neuroscience** v. 23, p. 439–452, 2022.

SETH, A.; HOHWY, J. Predictive processing as an empirical theory for consciousness science. **Cognitive Neuroscience**. v. 12, n. 2, p. 89-90, 2021.

SHANNON, C. A Mathematical Theory of Communication. **Bell System Technical Journal**, v. 27, p. 379-423, 1948.

SIMS, M. **Strong continuity of life and mind**: the free energy framework, predictive processing and ecological psychology. Edinburgh: The University of Edinburgh, 2021.

SINGER, W. Phenomenal awareness and consciousness from a neurobiological perspective. In: METZINGER, T. (Org.), **Neural correlates of consciousness**: Empirical and conceptual questions (pp. 121–137). Cambridge, MA: MIT Press, 2000.

SMITH, R.; BADCOCK, P.; FRISTON, K. Recent advances in the application of predictive coding and active inference models within clinical neuroscience. **Psychiatry and Clinical Neurosciences**, v. 75, n. 1, p. 3-13, 2020.

SMITH, R.; FRISTON, K.; WHYTE C. A step-by-step tutorial on active inference and its application to empirical data. **Journal of Mathematical Psychology**, v. 107, 2022.

SRINIVASAN, M.; LAUGHLIN, S.; DUBS, A. Predictive coding: a fresh view of inhibition in the retina. **Proceedings of the Royal Society of London. Series B. Biological Sciences**, v. 216, n. 1205, p. 427–459, 1982.

SPREVAK, M. *Predictive coding I*: Introduction. [Preprint] URL: <http://philsci-archive.pitt.edu/id/eprint/19365>, 2021.

STOCK, A.; STOCK, C. A short history of ideo-motor action. **Psychological Research**, v. 68, p. 176-188, 2004.

SUTHERLAND, S. **The International Dictionary of Psychology**. Crossroad, 1996.

STEWART, J., GAPENNE, O., DI PAOLO, E. **Enaction**: toward a new paradigm for cognitive science. Bradford: MIT Press, 2010.

SWANSON, L. The Predictive Processing Paradigm Has Roots in Kant. **Frontiers in Systems Neuroscience** v. 10. n.79, 2016.

SYTSMA, J., OZDEMIR, E. No Problem: Evidence that the Concept of Phenomenal Consciousness is Not Widespread. **Journal of Consciousness Studies**, v. 26, n. 9-10, p. 241-256, 2019.

TILIKETE, C.; VIGHETTO, A. Oscillopsia: causes and management. **Current Opinion in Neurology**, v. 24, n. 1, p. 38-43, 2011.

TONONI, G. An information integration theory of consciousness. **BMC Neuroscience** v.5, n. 42, 2004.

TONONI, G. Consciousness as integrated information: A provisional manifesto. **The Biological Bulletin**, v. 215, n. 3, p. 216–242, 2008.

TONONI G, EDELMAN G. Consciousness and complexity. **Science**, v. 282, n. 5395, p. 1846-1851, 1998.

TRIBUS, M. **Thermodynamics and thermostatics**: An introduction to energy, information and states of matter, with engineering applications. New York: D Van Nostrand, 1961.

TYE, M. Qualia. In: ZALTA, E. **The Stanford Encyclopedia of Philosophy**, 2021.

TYLER, C. Ten Testable Properties of Consciousness. **Frontiers in Psychology**, v. 11, 2020.

VARELA, F. **Principles of biological autonomy**. New York: Elsevier North Holland, 1979.

WADE, P. Confucian Free Energy: The Predictive Mind in Ancient China. **PsyArXiv Preprints**, 10.31234/osf.io/zm245, 2020.

WALSH, K., McGOVERN, D., CLARK, A., O'CONNELL, R. Evaluating the neurophysiological evidence for predictive processing as a model of perception. **Annals of the New York Academy of Sciences**, v. 1464, p. 242-268, 2020.

WARREN, J. This Quintessence of Dust – Consciousness Explained, at Thirty. **Philosophical Papers**, v. 1, n. 28, 2021.

WIESE, W; METZINGER, T. Vanilla PP for Philosophers: A Primer on Predictive Processing. In: METZINGER, T; WIESE, W. (Orgs.). **Philosophy and Predictive Processing**: 1. Frankfurt am Main: MIND Group, 2017.

WILLIAMS, D. Predictive Processing and the Representation Wars. **Minds & Machines**, v. 28, p. 141-172, 2018.

WILLIAMS, D. Is the Brain an Organ for Prediction Error Minimization? **PhiSci Archive**, 2022.

WIERZBICKA, A. From 'Consciousness' to 'I Think, I Feel, I Know'. **Journal of Consciousness Studies**, v. 26, n. 9-10, p.257-269, 2019.

WHITE, C. **An Introduction to the Cognitive Science of Religion: Connecting Evolution, Brain, Cognition, and Culture**. Routledge, 2021.

WHYTE, C. Integrating the global neuronal workspace into the framework of predictive processing: Towards a working hypothesis. **Consciousness and Cognition**, v.73, n.102763, 2019.

WOLFE, J.; KLUENDER, K.; LEVI, D. **Sensation and Perception**. Oxford University Press, 2018.

ZEKI, S. Localization and globalization in conscious vision. **Annual Review of Neuroscience**, v. 24, p. 57–86, 2001.