

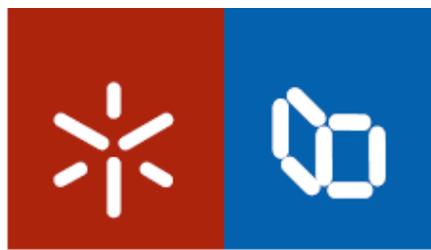


Abordagens Neurofilosóficas Redutivas e Não-Redutivas à Informação e aos Qualia: Uma Comparação Metodológica

Steven Sequeira Gouveia

Universidade do Minho
Instituto de Letras e Ciências Humanas





Universidade do Minho
Instituto de Letras e Ciências Humanas

Steven Sequeira Gouveia

**ABORDAGENS NEUROFILOSÓFICAS REDUTIVAS E NÃO-
REDUTIVAS À INFORMAÇÃO E AOS QUALIA: UMA
COMPARAÇÃO METODOLÓGICA**

Tese de Doutoramento em Filosofia,
Especialidade de Filosofia da Mente

Trabalho realizado sob a orientação do
Professor Doutor Manuel Curado
Professor Doutor Georg Northoff

julho de 2021

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-SemDerivações

CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

AGRADECIMENTOS

Gostaria de agradecer, primeiramente, ao Professor Doutor Manuel Curado por me ter fascinado, desde a primeira aula de Filosofia Antiga, com o rigoroso pensar filosófico. De seguida, ao Professor Doutor Georg Northoff por me mostrar de que forma se pode investigar o objeto de estudo mais complexo e obscuro da história intelectual da humanidade, o cérebro consciente. À Ana Mafalda, que me deu um auxílio final fundamental. Ainda, a todos os colegas e amigos que fizeram parte deste percurso incrível, longo e sofrido. E concluo esta nota agradecendo à Fundação da Ciência e Tecnologia o apoio fundamental para a realização desta investigação internacional.

Trabalho realizado com o Apoio Financeiro

Fundação para a Ciência e a Tecnologia (FCT) do Ministério da Educação e Ciência (MEC) da República Portuguesa.

Referência da bolsa

SFRH/BD/128863/2017.



III



DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração da presente tese. Confirmando que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Código de Conduta Ética da Universidade do Minho.

ABORDAGENS NEUROFILOSÓFICAS REDUTIVAS E NÃO-REDUTIVAS À INFORMAÇÃO E AOS QUALIA: UMA COMPARAÇÃO METODOLÓGICA

Resumo: O objetivo central desta tese doutoral é descobrir e identificar de que forma as concepções metafilosóficas de diversas posições metodológicas em relação ao problema da relação das metodologias da investigação filosófica com as metodologias da investigação neurocientífica influenciam essas mesmas posições. Ademais, depois de reconstruirmos os fundamentos de cada abordagem e de refletirmos criticamente sobre os mesmos, faremos notar de que formas essas mesmas concepções metafilosóficas de cada abordagem influenciam as suas posições em relação a dois conceitos centrais de cada uma das áreas científicas: do lado da Filosofia, o conceito de *Qualia*; e do lado da Neurociência, o conceito de Informação. O problema metodológico surge de se tentar encetar uma possível relação entre as duas disciplinas de modo a alcançar uma metodologia que possa investigar o fenômeno da mente consciente e da sua relação com o cérebro eficazmente. De um lado, a forma tradicional de definir o trabalho filosófico como uma metodologia *a priori* para lidar com os diversos problemas filosóficos parece opor-se claramente à forma tradicional de conceber o trabalho neurocientífico, nomeadamente como uma investigação puramente *a posteriori*. Deste modo, iremos analisar quatro abordagens que procuram responder a este problema metodológico: a Abordagem do Isolacionismo, que procurará negar a existência do mesmo; a Abordagem do Reducionismo, cujo foco será detalhar de que forma se pode reduzir o trabalho filosófico ao trabalho neurocientífico; a Abordagem da Neurofenomenologia que demonstrará a importância de uma abordagem corporal ao estudo da mente; e, finalmente, a Abordagem da Neurofilosofia Não-Redutiva, que irá expor uma metodologia muito particular que concebe a utilidade epistêmica tanto do trabalho filosófico como do neurocientífico para investigar o cérebro consciente. Depois de analisarmos as nomeadas abordagens, aplicaremos as mesmas ao conceito de Qualia e Informação para compreender de que forma os pressupostos metafilosóficos de cada abordagem influenciam as definições desses mesmos conceitos. De seguida, iremos aplicar a metodologia que nos parece ser mais plausível ao conceito da percepção, para concluir esta tese doutoral com a análise de uma teoria inovadora que procurará explicar todos os fenômenos mentais apelando ao mesmo princípio-base, concluindo que a mesma, seguindo a Abordagem da Neurofilosofia Não-Redutiva, terá de ser incorporada de forma mais holística e equilibrada entre a atividade neuronal do cérebro, por um lado, e o impacto do ambiente no corpo, por outro.

Palavras-Chave: Filosofia da Mente, Informação, Neurociência, Neurofilosofia, Qualia.

REDUCTIVE AND NON-REDUCTIVE NEUROPHILOSOPHICAL APPROACHES TO INFORMATION AND QUALIA: A METHODOLOGICAL COMPARISON

Abstract: The central goal of this doctoral thesis is to discover and identify how the metaphilosophical conceptions of different methodological approaches in relation to the problem of the relationship between the philosophical research methodologies and the neuroscientific research methodologies influence those same approaches. Furthermore, after reconstructing the basic foundations of each approach and reflecting critically on them, we will note how the metaphilosophical presuppositions influence how two central concepts in both scientific areas are defined: on the philosophy side, the concept of Qualia; and on the neuroscience side, the concept of Information. The methodological problem arises precisely from trying to conceive a possible relationship between the two disciplines in order to achieve a methodology that can effectively investigate the phenomenon of the conscious mind and its relationship with the brain. On the one hand, the traditional way of defining philosophical work as an *a priori* methodology for dealing with different philosophical problems seems to be clearly opposed to the traditional way of conceiving neuroscientific work, namely as a purely *a posteriori* investigation. Therefore, we will analyze four approaches that seek to answer to the methodological problem: the Isolationist Approach, which will try to deny the existence of the problem; the Reductionist Approach, which will focus on specifying how philosophical work can be reduced to neuroscientific work; the Neurophenomenology Approach, that will demonstrate the importance of an embodied approach to the study of the conscious mind; and, finally, the Non-Reductive Neurophilosophical Approach, which will expose a very particular methodology that conceives the epistemic utility of both philosophical and neuroscientific work to investigate the human brain. After analyzing these approaches, we will apply them to the concept of Qualia and Information to understand how the metaphilosophical assumptions of each approach influence the definitions of those concepts. Then, we will apply the methodology that seems to be more plausible to the concept of perception, to conclude this doctoral thesis with the analysis of an innovative theory that tries to explain all mental phenomena using the same basic foundation, concluding that, following the Non-Reductive Neurophilosophical Approach, this theory will need to be incorporated in a more holistic and balanced structure between the neuronal activity of the brain, on the one hand, and the impact of the environment on the body, on the other.

Keywords: Information, Neurophilosophy, Neuroscience, Philosophy of Mind, Qualia.

ÍNDICE

AGRADECIMENTOS.....	III
RESUMO.....	V
ABSTRACT.....	VI
LISTA DE FIGURAS.....	XI
CAPÍTULO I – PROBLEMA METODOLÓGICO EM FILOSOFIA E NEUROCIÊNCIA.....	1
I.1. Introdução Geral.....	1
I.2. A Metafilosofia: Introdução e Análise.....	7
I.3. A Filosofia da Neurociência: Introdução e Análise.....	17
I.3.1. Natureza e Estrutura da Neurociência.....	18
I.3.2. Conceitos, Princípios e Métodos da Neurociência.....	19
I.3.3. Questões Tradicionais da Filosofia e Neurociência.....	20
I.3.4. Filosofia da Neurociência, Relação com outras Áreas da Filosofia e novas Áreas de Investigação.....	21
I.4. Breve Introdução à Neurociência.....	25
CAPÍTULO II – ISOLACIONISMO METODOLÓGICO.....	28
II.1. Abordagem Isolacionista.....	29
II.1.1. Os Conceitos Antecedem os Factos.....	31
II.1.2. A Influência de Wittgenstein.....	40
II.1.3. Exemplo de Investigação Isolacionista: o Determinismo.....	41
II.1.4. Várias defesas da Abordagem Isolacionista.....	50
II.2. Reflexão Crítica.....	53
II.2.1. A Objeção da “Invariância Conceptual”.....	53
II.2.2. A Objeção de Dennett.....	56
II.2.3. A Objeção de Searle.....	65
II.2.4. A Objeção da “Prática e Anti-excepcionalidade Filosófica”.....	70
II.3. Conclusão.....	75

CAPÍTULO III – NEUROFILOSOFIA REDUTIVA.....76

III.1. Pressupostos da Abordagem Reducionista.....	77
III.1.1. Negação do Dualismo.....	77
III.1.2. O Conceito de Reducionismo.....	78
III.1.2.1. Redução como Tradução (RcT).....	80
III.1.2.2. Redução como Derivação (RcD).....	83
III.1.2.3. Redução como Explicação (RcE).....	86
III.1.3. Redução em Patricia Churchland.....	88
III.1.4. Teoria da Psicologia Popular.....	93
III.1.5. Redução em John Bickle.....	97
III.1.6. A Teoria da Redução Interteórica.....	100
III I.2. Reflexão Crítica.....	107
III.2.1. A Objeção da “Não-Teoria”.....	107
III.2.2. A Objeção da “Realização Múltipla”.....	109
III.2.3. A Objeção da “Redução Ontológica”.....	113
III.2.4. A Objeção do “Isolacionismo Metodológico”.....	115
III.2.5. A Objeção Metafilosófica.....	117
III.3. Conclusão.....	122

CAPÍTULO IV – NEUROFENOMENOLOGIA.....123

IV.1. Pressupostos da Abordagem da Neurofenomenologia.....	124
IV.1.1. A Fenomenologia.....	124
IV.1.2. A Importância do Corpo e a Crítica ao Cérebro.....	127
IV.1.3. Os 4E's e a Abordagem Neurofenomenológica de Francisco Varela.....	130
IV.1.4. Um exemplo de Investigação Neurofenomenológica: o caso da “Rivalidade Binocular”.....	138
IV.1.5. O Eu (<i>Self</i>) em Evan Thompson.....	141
IV.1.6. Três Lições Pragmáticas para uma Neurofenomenologia Rigorosa.....	145
IV.1.6.1. Reformulação da Ciência.....	145
IV.1.6.2. Melhor Ciência.....	146
IV.1.6.3. O Foco no Entrevistador.....	147
IV.2. Reflexão Crítica.....	150

IV.2.1. A Objeção da “Lacuna Explicativa” (<i>Explanatory Gap</i>).....	150
IV.2.2. A Objeção da “Introspecção”	153
IV.2.3. A Objeção da “Acusação Reducionista”	157
IV.2.4. A Objeção da “Metafilosofia Inconclusiva”	159
IV.3. Conclusão.....	166
CAPÍTULO V – NEUROFILOSOFIA NÃO-REDUTIVA.....	167
V.1. Pressupostos da Abordagem Não-Redutiva.....	168
V.1.1. Equilíbrio Metodológico.....	168
V.1.2. Estratégias Metodológicas da Neurofilosofia.....	169
V.1.2.1. Princípio da Assimetria.....	175
V.1.2.2. Princípio de Bidirecionalidade.....	177
V.1.2.3. Princípio de Circularidade Transdisciplinar.....	178
V.1.3. A Natureza e Metafilosofia da Hipótese Neurofilosófica.....	181
V.1.3.1. Perspetiva Baseada (e não reduzida) no Cérebro.....	183
V.1.3.2. Neurociência da Primeira Pessoa.....	185
V.1.4.1. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o caso do Eu (<i>Self</i>).....	188
V.1.4.2. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o Eu e a COVID-19.....	191
V.1.4.3. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o caso das Emoções.....	194
V.2. Reflexão Crítica.....	200
V.2.1. Objeção do “Erro de Categoria e da Modalidade”	200
V.2.2. Objeção da “Irrelevância Metodológica”	202
V.2.3. Objeção da “Metafilosofia Imprecisa”	205
V.3. Conclusão.....	214
CAPÍTULO VI – QUALIA E INFORMAÇÃO EM NEUROCIÊNCIA E FILOSOFIA.....	215
VI. 1. Conceito de <i>Qualia</i>	215
VI.1.1. A Abordagem Isolacionista e <i>Qualia</i>	219
VI.1.2. A Abordagem Reducionista e <i>Qualia</i>	226

VI.1.3. A Abordagem da Neurofenomenologia e <i>Qualia</i>	233
VI.1.4. A Abordagem da Neurofilosofia Não-Redutiva e os <i>Qualia</i>	238
VI.2. Conceito de “Informação” e as Diversas Abordagens.....	245
VI.3. Conclusão.....	253
CAPÍTULO VII – PERCEÇÃO: UMA ABORDAGEM NEUROFILOSÓFICA	
NÃO-REDUTIVA.....	255
VII.1. Introdução Geral.....	255
VII.1.1. Teorias da Percepção I: Realismo Direto.....	256
VII.1.2. Teorias da Percepção II: Realismo Indireto.....	260
VII.2. Realismo da Percepção e Modelos do Cérebro em Neurociência.....	266
VII.2.1. Realismo Ingênuo e Modelo Passivo do Cérebro.....	266
VII.2.2. Realismo Indireto e Modelo Ativo do Cérebro.....	271
VII.2.3. Realismo Direto e Modelo Espaço-temporal do Cérebro.....	273
VII.3. Conclusão.....	278
CAPÍTULO VIII – PROCESSAMENTO PREDITIVO E CONSCIÊNCIA: FALÁCIA DA PREVISÃO	
E A SUA RESOLUÇÃO ESPAÇOTEMPORAL.....	279
VIII.1. Introdução Geral.....	279
VIII.2. Processamento Preditivo e Falácia da Previsão.....	285
VIII.3. Processamento Preditivo e Dinâmica Espaço-temporal.....	288
VIII.3. Conclusão.....	292
CONCLUSÃO.....	293
ANEXO.....	297
BIBLIOGRAFIA.....	328

LISTA DE FIGURAS

Figura 2.1.: Métodos utilizados no Século XX	71
Figura 2.2.: Métodos utilizados entre 2009-2013.....	71
Figura 3.1. O espectro das reduções interteóricas.....	104
Figura 4.1.: Redução fenomenológica.....	135
Figura 4.2.: Exemplo do desenho experimental e estímulos.....	138
Figura 5.1.: Estratégia metodológica de Churchland.....	171
Figura 5.2.: Estratégia da adaptação unilateral dos conceitos aos factos.....	172
Figura 5.3.: Estratégia da Neurofilosofia Não-Redutiva.....	174
Figura 5.4.: As várias posições que procuram responder ao problema da relação (ou não) entre a investigação filosófica e a investigação neurocientífica.....	211
Figura 5.5.: As posições metafisológicas das abordagens analisadas.....	212
Figura 6.1.: Espectro biofísico-computacional do cérebro.....	241
Figura 6.2.: Exemplo de atividade neuronal.....	250
Figura 6.3.: Exemplo das diferenças de análise entre as Abordagens Reducionistas e da Neurofilosofia Não-Redutiva.....	251
Figura 7.1.: Ilusão de Müller-Lyer (original).....	257
Figura 7.2.: Ilusão de Müller-Lyer (corrigida).....	257
Figura 8.1.: Exemplo da ilusão da mão de borracha.....	281
Figura 8.2.: Exemplo da ilusão do vaso de Rubin.....	285

I. PROBLEMA METODOLÓGICO EM FILOSOFIA E NEUROCIÊNCIA

I.1. Introdução Geral

O foco desta tese doutoral será o problema metodológico que deriva da relação entre as disciplinas da Filosofia e da Neurociência. Tal afirmação poderá causar estranheza logo de princípio. O que significa “metodológico”? Porque existe um problema em primeira instância? A que se está a referir quando se usa os conceitos “Filosofia” e “Neurociência”?

Ora, todas estas questões são imensamente relevantes e abrangem um escopo teórico colossal que dificilmente poderemos conciliar numa tese deste tipo. Apesar disso, o nosso objetivo principal será exatamente delimitar cada um dos conceitos principais de modo a que possamos (i) expressar o problema claramente; (ii) aplicá-lo ao estudo da mente consciente; (iii) enquadrar as diferentes posições que tentam oferecer uma resposta a (i) para alcançar (ii); e, finalmente, (iv) demonstrar qual das posições em (iii) é a mais plausível e que merece a nossa proposta final.

Esta tese caracterizar-se-á, naturalmente, por uma forte diversidade de abordagens que poderão fazer o filósofo mais tradicional duvidar da validade desta investigação. Pelos mesmos motivos, um neurocientista tradicional estranhará muitos dos temas e conceitos que discutiremos. Por isso mesmo, é necessária uma caridade inicial para os assuntos a desenvolver que poderão acabar numa reflexão que, mesmo não se concordando com a conclusão, poderá ser proveitosa.

Comecemos, assim, a nossa investigação.

Qualquer disciplina académica possui um método próprio ou um conjunto de métodos que podem partilhar uma propriedade específica que a distinga de outros métodos particulares. Nesse sentido, a Filosofia, como área genérica que estuda certos problemas particulares – como por exemplo, o que é o conhecimento ou se podemos reproduzir a mente consciente artificialmente – e a Neurociência, como uma área igualmente geral que estuda outros (ou semelhantes, como veremos) problemas – como os mecanismos neurológicos na esquizofrenia ou como é que o nosso cérebro percebe o mundo externo – não fogem a esta regra.

A Filosofia partilha diversos métodos, como analisaremos em detalhe ao longo desta reflexão. Todavia, grande parte destes métodos parecem poder reduzir-se ao que podemos chamar de “metodologia *a priori*”, isto é, a Filosofia parece envolver, tradicionalmente, a investigação de problemas específicos através de metodologias e abordagens *a priori*, ou seja, aquele tipo de conhecimento

particular cuja justificação é totalmente independente da experiência ou de qualquer informação *a posteriori*, como por exemplo $T = \text{'Todos os solteiros são não casados'}$. Galen Strawson afirmou que um argumento *a priori* é aquele tipo de argumento que “podemos perceber que é verdadeiro apenas estando deitado no sofá” (cf. Sommers, 2003: s/p). Esta explicação mais ou menos humorística possui um esclarecimento bastante relevante que deve ser realçado: a presença no sofá – o pensar sem requerer ir ao mundo externo – é condição suficiente para analisar um argumento *a priori* exatamente porque qualquer outro tipo de informação – por exemplo, sair do sofá e procurar no mundo físico se de facto ‘Todos os solteiros são não casados’ – é, no mínimo, supérflua, e no máximo, epistemicamente irrelevante: bastará analisar o significado de cada conceito para perceber o valor de verdade de T .

O ponto importante a reter é o seguinte: não precisamos de fazer qualquer tipo de ciência empírica – isto é, investigação empírica ou *a posteriori* – para perceber a verdade da proposição T .

Ora, em relação à Neurociência, o oposto parece ser comumente alegado. A Neurociência, como ciência específica que estuda uma série de objetos particulares, utiliza um tipo de metodologia *a posteriori*, isto é, baseando a sua investigação a indagar o mundo físico e na experiência que podemos alcançar através da pesquisa controlada.

Como facilmente se perceberá, a Neurociência, como ciência, parece colocar-se do lado oposto da Filosofia em termos de metodologia. Se a Filosofia se pode realmente fazer simplesmente no nosso sofá, nenhum investigador em Neurociência seria bem-sucedido se fizesse o mesmo: nada teria para apresentar, e seria rapidamente despedido.

Deste modo, apontamos para uma aparente importante diferença metodológica: a Neurociência requer “descobrir” no mundo físico as respostas aos problemas que coloca. Tal parece não acontecer com a investigação filosófica tradicional.

Por isso mesmo, uma relação entre a Filosofia e a Neurociência pode causar, à primeira vista, espanto e desconfiança. A Filosofia tem as suas raízes na Grécia Antiga, já a Neurociência, como disciplina bem definida, não possui mais de cinquenta anos. Tal deve-se exatamente à suposta incompatibilidade entre os dois métodos agora introduzidos. Se a Filosofia se caracteriza unicamente por ser uma investigação *a priori*, e a Neurociência uma investigação *a posteriori*, então porque não terminamos por aqui esta investigação doutoral?

A resposta é simples, mas, como veremos, enriquecedora. O objetivo desta reflexão será exatamente recriar, sistematizar e problematizar, na medida do possível, as respostas que podem ser dadas ao problema metodológico que deriva da relação entre as duas disciplinas, isto é, ao facto de parecer, à partida, que a Filosofia e a Neurociência, por se apoiarem em metodologias diferentes, não

podem trabalhar em conjunto para investigar diversos problemas particularmente ligados à Filosofia da Mente.

É esta característica fundamental que a Filosofia parece possuir em geral: o facto de os seus profissionais fazerem Filosofia a partir da “poltrona” (*armchair*), como afirmado por dois metafilósofos contemporâneos:

[...] Assim, os naturalistas metodológicos descarados, apesar das suas doutrinas explícitas, não fizeram necessariamente filosofia de forma diferente da maneira como costumava ser feita: em geral, sem depender substancialmente de – e muito menos de coletar – dados empíricos. Na prática, se não sempre na teoria, supunha-se que a filosofia poderia e deveria, de facto, ser feita a partir de uma “poltrona! [*armchair*] (D’Oro & Overgaard, 2017: 1).¹

Esta preocupação especificamente metodológica nem sempre teve relevância na história da Filosofia, porém. É famosa, por exemplo, a tese de Gilbert Ryle que argumentava que indagações metodológicas tendem a distrair-nos da investigação filosófica em si (cf. Ryle, 2009: 331). Karl Popper também afirmou algo nessa direção: “um filósofo deveria filosofar: ele deveria tentar resolver problemas filosóficos, em vez de falar sobre filosofia” (Popper, 1968: 68).

Fazemos notar que o próprio Popper é inconsistente na sua forma de negar legitimidade filosófica à Filosofia da Filosofia. Note-se o que afirma em relação à Ciência: “[...] a investigação crítica sobre as ciências, as suas descobertas e os seus métodos”, afirma Popper, “continua a ser uma característica da investigação filosófica” (Popper, 1975: 53). Ora, a mesma coerência teórica deveria ser mantida pelo filósofo em relação à Filosofia da Filosofia, dado que uma análise crítica das descobertas e métodos de uma área, seja ela qual for, deve ser objeto de investigação filosófica, mesmo que essa área seja a própria Filosofia.

Apesar disso, o pensamento filosófico sobre o método tem vindo a desenvolver-se, como não poderia deixar de ser. Estas questões especialmente direccionadas sobre a forma como a Filosofia é praticada – ou deve ser praticada, como veremos – pertence a uma subárea da Filosofia a que podemos chamar de “Metafilosofia”, ou Filosofia da Filosofia – a segunda parte desta introdução irá detalhar mais esta área de estudo.

¹Apresenta-se um anexo final com as passagens na língua original traduzidas neste trabalho pelo Autor para língua portuguesa.

Alguns autores criticam, porém, este termo. Por exemplo, Stanley Cavell afirma: “Se nego uma distinção, é a distinção que anda na moda entre filosofia e metafilosofia, filosofia da filosofia. As observações que faço sobre filosofia [...] são [...] nada mais nada menos que observações filosóficas” (Cavell, 2002: xxxii). Esta crítica passa pela ideia de que tal termo parece sugerir que quando se está a discutir especificamente o método filosófico está-se a fazer algo diferente de Filosofia, quando o que se está efetivamente a fazer é Filosofia. Veremos, já de seguida, a forma como esta crítica impacta na nossa visão sobre o que é a Metafilosofia e sua diferença em relação à Filosofia da Filosofia.

Uma importante distinção que irá acompanhar-nos ao longo desta investigação é a distinção conceptual entre questões descritivas e questões normativas. Muito resumidamente, questões descritivas sobre método apontam para respostas sobre o que de facto os filósofos fazem (tradicionalmente ou na atualidade) quando afirmam fazer Filosofia. Diferente são as questões normativas: neste ponto, não nos interessa realmente o que os filósofos fazem *de facto*, mas o foco é sobre o que os filósofos *devem* fazer, isto é, qual é o método correto para fazer boa Filosofia, independentemente de como é que a Filosofia é realmente feita na prática.

Assim sendo, a nossa investigação irá dar uma especial atenção aos pressupostos metafilosóficos das diferentes abordagens que iremos examinar. A nossa suspeita inicial recai sobre a ideia de que grande parte dos filósofos contemporâneos não toma em consideração esta distinção como filosoficamente relevante, acabando por confundir *o que fazem* com o *que devem fazer*. Tal será especialmente analisado nas várias abordagens que procuram afirmar ou negar uma relação entre a Filosofia e a Neurociência.

Outro ponto importante que deve ser realçado é compreender que, embora seja importante ter em consideração esta distinção como fundamentalmente relevante para a investigação, há pontos de contacto óbvios: os pressupostos normativos influenciam as metodologias dos filósofos e vice-versa – estas não podem ser totalmente separadas.

A questão metodológica será, assim, central na nossa tese. Iremos investigar (iii) quatro abordagens que procuram, de algum modo, responder ao problema metodológico. As diferentes metodologias que estão presentes nos estudos da Neurociência e da Filosofia da Mente de modo a estabelecer um enquadramento teórico que resolva ou elucide os tradicionais problemas filosóficos ligados à mente humana que iremos investigar nesta investigação serão as seguintes:

- (Capítulo 2) Abordagem do Isolacionismo [*isolation approach*]: esta abordagem, mais tradicional e conservadora da Filosofia, propõe que as duas disciplinas estão metodologicamente

separadas. Se o problema metodológico procura indagar sobre existir de facto um problema, esta abordagem procura responder diretamente: não, não existe qualquer problema à partida. Nesta perspetiva, as verdades conceptuais não podem ser informadas por verdades empíricas, e vice-versa. Os conceitos filosóficos e os factos neurocientíficos ocupam-se de diferentes domínios: natural/verdade versus lógico/sentido das palavras – veremos o que tal significa com detalhe no capítulo dedicado a esta abordagem. Esta é uma posição típica tanto para neurocientistas como para filósofos, como iremos ver. Neste ponto, analisaremos em detalhe a aclamada obra *Fundamentos Filosóficos da Neurociência* (2003), do neurocientista M. R. Bennett e do filósofo P. M. Hacker.

- (Capítulo 3) Abordagem do Reduccionismo [*reductive neurophilosophy*]: esta abordagem, ao contrário do isolacionismo metodológico, afirma que de facto há um problema metodológico. Se queremos avançar no estudo da mente consciente, há apenas uma opção: a análise *a priori* tradicional da Filosofia deverá ser tendencialmente substituída por uma metodologia focada na investigação empírica. Assim, o objetivo desta metodologia é reduzir os conceitos da Filosofia da Mente ao domínio empírico. Os maiores representantes desta metodologia são Patricia Churchland (1986) e John Bickle (2003) que, diferentemente, propõe uma estrutura específica. Estudaremos as obras mais importantes de ambos, procurando reconstruir as suas argumentações em apoio de uma Neurofilosofia que se revelará não mais do que uma defesa de um materialismo eliminativista que não nos parece ser suficiente para explicar todos os fenómenos mentais.
- (Capítulo 4) Abordagem da Neurofenomenologia: esta metodologia procurará responder ao problema metodológico cruzando os princípios de base de uma escola específica da Filosofia, a Fenomenologia, isto é, a ideia de que a investigação filosófica deve passar pela procura no foco dos fenómenos como aparecem à consciência, independentemente da sua caracterização metafísica, com os avanços dos dados neurocientíficos. Descrever-se-á a importância de introduzir na equação deste problema o corpo e o ambiente de modo a desenvolver uma Neurociência que consiga explicar *todos* os fenómenos mentais, sem, contudo, reduzi-los ao meramente empírico. Reconstruiremos as suas principais características, focando a nossa atenção no fundador desta metodologia, Francisco Varela (1991, 1996), e nos posteriores trabalhos de Evan Thompson (2007, 2015).

- (Capítulo 5) Abordagem da Neurofilosofia Não-Redutiva: finalmente, nesta abordagem, procuraremos apresentar uma metodologia iterativa que procura fazer corresponder os conceitos de diferentes disciplinas pelos diferentes domínios, isto é, procura não fixar o conceito a nenhuma das disciplinas, mas, antes, ter por objetivo cruzar sucessivamente o método filosófico com o método neurocientífico para que o conceito seja o mais plausível possível. Contra uma Neurofilosofia redutiva, afirmar-se-á que a Neurofilosofia não deve ser reduzida nem ao método *a priori* da Filosofia nem ao método *a posteriori* da Neurociência. Assim, esta metodologia procurará afirmar que não há uma Neurociência livre de conceitos (afirmação da Neurofilosofia Redutiva, que procura eliminar qualquer conceito filosófico nos dados da Neurociência), mas, assumindo que os conceitos são relevantes para a Neurociência, de modo a testar a plausibilidade conceptual das descobertas neurocientíficas. Deste modo, é necessário aproximar o mundo lógico trabalhado pelos filósofos ao mundo natural da ciência. Como iremos ver, esta abordagem pode ser entendida como uma continuação do trabalho de W. O. Quine. Ademais, esta posição afirma que não há, também, uma filosofia do mundo natural independente dos factos científicos. Os factos são relevantes para uma filosofia naturalista, para testar a plausibilidade empírica dos conceitos filosóficos. Se tal for filosoficamente sustentável, estaremos perante uma nova disciplina, que não é filosófica, no seu sentido mais puro e tradicional, mas que também não é científica. A Neurofilosofia Não-Redutiva encontra-se nos limites ou na fronteira da Filosofia e da Neurociência, ganhando assim uma independência de trabalho que poderá permitir, finalmente, oferecer uma metodologia eficaz que possa trabalhar os conceitos difíceis da Filosofia da Mente.

Nesta tese, após a reconstrução crítica das quatro abordagens, iremos aplicar as mesmas a dois conceitos fundamentais: o conceito filosófico de *Qualia* e o conceito científico de *Informação* (Capítulo 6).² Como será demonstrado, cada metodologia pressuposta tem implicação direta na forma como os dois conceitos são definidos. Declarando de outro modo: as metodologias não são neutras, contribuindo, diferentemente, para uma delimitação significativa dos conceitos. Ademais, veremos a importância –

² Iremos apreciar somente a perspectiva metodológica sobre os conceitos de *Qualia* e *Informação*, e não uma discussão substancial sobre a natureza e ontologia destes dois conceitos, dado que o interesse do nosso projeto é primariamente a discussão metodológica, isto é, compreender de que forma as metodologias pressupostas influenciam a maneira como os conceitos são definidos.

muitas vezes ignorada – que os pressupostos metafilosóficos têm sobre cada metodologia. Durante a apresentação de cada metodologia, confrontaremos alguns dos conceitos centrais com resultados e investigações das ciências empíricas.

Através da análise desses conceitos, iremos estabelecer um critério que nos possa indicar qual das metodologias é mais eficaz a resolver as problemáticas obscuras. Esse critério será o da plausibilidade empírica da definição dos conceitos, cruzando-o sempre com a plausibilidade conceptual. Deste modo, procurar-se-á determinar qual das metodologias apresenta um maior poder explicativo ao definir os dois conceitos. Iremos, ainda, apresentar a aplicação da abordagem que nos parece mais plausível ao conceito de percepção no penúltimo capítulo (Capítulo 7), apresentando, para concluir esta reflexão, um capítulo final (Capítulo 8) sobre o processamento preditivo e a consciência a partir de uma perspectiva da Neurofilosofia Não-Redutiva.

Antes de passarmos para a primeira metodologia, a isolacionista, façamos uma introdução à Metafilosofia, dado que será uma área que acompanhará a nossa reflexão metodológica ao longo da nossa investigação.

I.2. A Metafilosofia: Introdução e Análise

A Metafilosofia é uma subárea de investigação da Filosofia que procura focar-se na natureza da própria Filosofia (cf. Lazerowitz, 1970: 91). Embora esta definição nos informe minimamente sobre o objeto de investigação desta área, há várias formas de interpretar a afirmação “a natureza própria da filosofia”. Esta asseveração pode ser interpretada como defendendo que a Metafilosofia é uma “Filosofia da Filosofia”, isto é, a Metafilosofia seria uma área de investigação que aplica métodos tradicionais da Filosofia à própria Filosofia. Esta leitura poderá ter duas versões:

(1) A Metafilosofia como uma análise de primeira ordem de um conjunto de temas ou assuntos (neste caso, a própria Filosofia), sendo somente mais uma instância da investigação filosófica *per se* (cf. Wittgenstein, 1953: § 121; Williamson, 2007: ix);

(2) A Metafilosofia como uma análise de segunda ordem – como uma disciplina que se analisa a si e outras disciplinas (como a Filosofia da Economia ou a Filosofia da Neurociência, como veremos seguidamente) (cf. Rescher, 2006).

Como disciplina acadêmica, esta subárea da Filosofia não tem recebido grande atenção dos próprios filósofos. Collin McGinn afirma, por exemplo, que a Metafilosofia é a área menos desenvolvida da Filosofia (cf. McGinn, 2002: 199). Rescher comenta no mesmo sentido afirmando que a Metafilosofia é uma área negligenciada da Filosofia (cf. Rescher, 2001: 1). Veja-se, por exemplo, o seguinte comentário negativo de Karl Popper em relação às questões metafilosóficas:

Acredito que a função de um cientista ou de um filósofo é resolver problemas científicos ou filosóficos, em vez de falar sobre o que ele ou outros filósofos estão a fazer ou devem fazer. Qualquer tentativa mal-sucedida de resolver um problema científico ou filosófico, se for uma tentativa honesta e dedicada, parece-me mais significativa do que uma discussão sobre perguntas como “O que é a ciência?” ou “O que é a filosofia?”. E mesmo que coloquemos esta última pergunta, como deveríamos fazê-lo, numa forma ligeiramente melhorada: “Qual é o caráter dos problemas filosóficos?”, não deveria preocupar-me muito com isso; deveria sentir que tinha pouca importância, mesmo em comparação com um problema menor da filosofia, como a questão de saber se toda discussão ou crítica deve sempre resultar de “pressuposições ou “suposições” que estão em si mesmos para além dos argumentos (Popper, 1968: 66).

Este ceticismo é comum em grande parte dos filósofos em relação à sua investigação e ao que fazem efetivamente: por exemplo, Bernard Williams declara que “a filosofia não é muito interessante quando fala sobre si mesma” (Williams, 2006: 169). Em menor número encontramos filósofos que escreveram obras explicitamente sobre esta área, como seja o caso de *The Philosophy of Philosophy* (2007) de Timothy Williamson, *Philosophical Dialectics: Essay on Metaphilosophy* (2006) de Nicholas Rescher ou *What is Philosophy?* (1994) de Gilles Deleuze e Félix Guattari.

O caso oposto, isto é, obras que implicitamente discutem o tema, é já bastante mais comum. Qualquer posição em Ética, Metafísica ou Filosofia Política assume, nem que seja implicitamente – como é a maior parte dos casos – uma definição de filosofia ou um pressuposto sobre o seu método de trabalho.³ Note-se que, muitas vezes, estes pressupostos são assumidos à partida, sem grande reflexão por parte do autor ou do investigador. Como veremos, esse será um problema relevante que levará a muitas confusões que poderiam ser evitadas na relação metodológica entre a Filosofia e a Neurociência.

³ Para uma análise sobre a distinção entre metafilosofia explícita e implícita com especial foco na segunda, ver Joll, 2010.

Por conseguinte, grande parte da metafilosofia contemporânea é implícita, e não explícita, salva raras exceções.

Uma primeira preocupação a respeito da tarefa de definir precisamente o que é a filosofia aparece com a primeira vaga da tradição analítica. Bertrand Russell e o seu discípulo Ludwig Wittgenstein esforçaram-se (em vão, dirão alguns) para pôr em prática a sua ideia fundamental: toda a (boa) Filosofia deverá começar por uma análise de proposições (cf. Russell, 1992: 9). Esta análise, para ambos os filósofos, era principalmente a aplicação de instrumentos formais – o que denominavam por “lógica” – às proposições a investigar.

Esta forma particular de fazer filosofia era baseada em dois estágios ou fases metodológicas. Primeiramente, o filósofo deveria transformar proposições ordinárias da linguagem comum ou científica na sua “verdadeira” forma. Este processo baseava-se na aplicação das formulações inovadoras da lógica de segunda ordem elaboradas por Frege que viria a influenciar a famosa “teoria das descrições” de Russell. O segundo estágio era correlacionar os elementos trabalhados do primeiro, as proposições simples, com elementos do mundo natural. Para Russell, o mundo era composto de “átomos” indivisíveis – daí defender a teoria metafísica do “atomismo lógico” – que eram estruturas limitadas pela forma lógica que, combinadas da maneira correta, formavam os objetos científicos e do quotidiano (cf. Beaney, 2007: 2-3).

Wittgenstein estava de acordo com Russell no seguinte aspeto: a estrutura lógica aparente de uma proposição poderá não ser a sua verdadeira forma e requer uma análise mais profunda (cf. Wittgenstein, 1961: secção 4.0031). O comumente apelidado de “primeiro” Wittgenstein, o autor do *Tractatus Logico-Philosophicus*, também viria a aceitar a tese metafísica do atomismo lógico. Como é sabido, Wittgenstein tinha procurado no *Tractatus* qual era a natureza da linguagem e a sua relação com o mundo. A sua doutrina principal baseava-se numa teoria pictórica do significado: a linguagem consiste em proposições que representam pictoricamente o mundo. As proposições seriam identificadas com as imagens lógicas dos factos do mundo. Toda a linguagem tinha uma estrutura comum: a sua forma lógica. O trabalho do filósofo era descobrir essa forma lógica na linguagem comum, que se encontra escondida devido à existência de objetos complexos.

A relação proposição-facto ficará clara se o objeto complexo for reduzido a objetos simples: esta é a doutrina do atomismo lógico que fará uso de uma extensão da “teoria das descrições” de Russell. Uma proposição terá sentido se exprimir a possibilidade de um facto: a verdade ou falsidade é analisada se estamos perante um facto possível (falsidade) ou um facto real (verdade) – mas ambos os valores de verdade têm sentido, por serem ambos possíveis. Para completar o esquema conceptual do atomismo

lógico, faltaria assinalar aquelas proposições que, através da redução, se compreendia serem, afinal, pseudoproposições, isto é, proposições sem sentido, evitando o que Rudolf Carnap classificou de sem sentido (der Unsinn).

Contudo, o segundo Wittgenstein, o autor das *Investigações Filosóficas*, irá criticar muitas das doutrinas agora descritas. Primeiramente, introduz-se o carácter não necessário na relação linguagem-mundo: se fosse necessária, seria única – mas o que observamos é que há muitas formas e dialetos. Deixará assim de acreditar em átomos lógicos. Ainda assim, irá manter a ideia de que a superfície da linguagem esconde a sua verdadeira natureza: mas agora o oculto será a variedade de formas a que a linguagem recorre como atividade social e interpessoal. Irá, nesta fase, introduzir o conceito de “jogos de linguagem”: as palavras não podem ser compreendidas fora do contexto em que são utilizadas. Assim, não existe nenhum carácter essencial à linguagem, mas apenas semelhanças de família.

Ademais, apesar das palavras se assemelharem todas umas às outras, pelo que parecem fazer todas o mesmo, Wittgenstein irá dar uma imagem para demonstrar que tal é apenas superficial: as palavras são como as alavancas de uma locomotiva – apesar de parecem todas iguais, umas servem para acelerar, outras para travar, outras para ligar o motor, outras para desligar, etc. Assim, o significado da palavra é dado pelo seu uso e não pelo que referem. Diz Wittgenstein numa fase de transição:

Mas, se tivéssemos de mencionar o que anima o signo, diríamos que é a sua utilização; [...] o signo (a frase) obtém o seu significado do sistema de signos, da linguagem à qual pertence. Numa palavra: compreender uma frase significa compreender uma linguagem (Wittgenstein, 1992: 30).

Seja como for, podemos facilmente observar, tanto em Russell como no primeiro Wittgenstein, que a Filosofia é definida como “crítica da linguagem”. O objetivo central da investigação filosófica é, neste paradigma, a procura da linguagem ideal ou, mais claramente, uma análise rigorosa da linguagem natural ou comum no sentido de transformá-la ou corrigi-la para uma linguagem ideal ou perfeita.

Tendo esta tese metafilosófica em mente, seguiu-se, historicamente, uma posição que iria extremar a intuição inicial de Russell e Wittgenstein e irá defender que uma particular definição de análise lógica é o único trabalho relevante da Filosofia. O Círculo de Viena, um grupo de intelectuais influenciados pelas inovações no estudo da lógica formal e, principalmente, pelos desenvolvimentos da Ciência empírica, iria defender uma posição metafilosófica bastante controversa, mas que, como veremos, ainda mantém alguma atualidade e influencia particularmente as relações entre a Filosofia e a Neurociência.

O manifesto do grupo afirmava que o trabalho filosófico residia em esclarecer problemas e afirmações em vez de engajar com problemas filosóficos “especiais” através de um método particular: o da análise lógica (cf. Neurath, Carnap, & Hahn, 1996: 321-328).

Baseados nas ideias iniciais de Auguste Comte, um conjunto de intelectuais de diversas áreas, como Rudolf Carnap, Otto Neurath ou Moritz Schlick, influenciados pelos empirismos de Hume, Russell e Mach e, principalmente, pela ideia de uma linguagem ideal que fosse alcançável através da análise lógica das proposições, iriam oferecer uma das metafisofias mais influentes do novo século. Aos ombros de Wittgenstein e avançando com uma interpretação muito particular do filósofo, irão afirmar que uma proposição apenas tem sentido se for (i) verdadeira por definição (análise analítica) ou (ii) se for verificada pela experiência (análise sintética), sendo que tudo o resto não teria qualquer valor epistêmico, afirmando que grande parte da Metafísica feita até então poderia ser considerada pseudofilosofia.

Embora o grupo não fosse filosoficamente tão uniforme como anunciado, havendo, por exemplo, alguns autores que defendiam que todas as proposições ou conhecimentos se deveriam reduzir a conhecimento científico empírico (como veremos mais à frente nesta tese), outros autores viam ainda valor na linguagem ideal que poderia ser alcançada pela análise lógica da linguagem. O ponto central é que, mais uma vez, os pressupostos metafisoficos do Círculo de Viena eram problemáticos, como defendeu, por exemplo, o segundo Wittgenstein, que irá dar origem a uma nova metafisofia.

A Filosofia da Linguagem Comum ou Ordinária, uma posição metafisofica baseada no Wittgenstein das *Investigações Filosóficas* deixa de achar relevante a linguagem lógica com objetivo ideal e passa a focar-se na análise da forma como a linguagem comum – vernacular – funciona e acontece. Se o Santo Graal do primeiro Wittgenstein é a linguagem ideal e, por isso, tendencialmente artificial, para o segundo Wittgenstein, o Santo Graal virá a ser a linguagem natural, isto é, não-artificial.

Esta tradição metafisofica germina em Oxford, particularmente com Gilbert Ryle e John L. Austin. Nesta perspectiva metafisofica, a Filosofia é igualmente definida como análise da linguagem. No entanto, esta análise é muito diferente da anterior. Uma análise das práticas linguísticas e dos significados das palavras poderá dar-nos acesso a distinções importantes que poderão reduzir as incoerências e equívocos da linguagem. Como diria famosamente Austin, esta análise não é a “última palavra” da Filosofia: somente a linguagem natural comum é a “primeira palavra” relevante (cf. Austin, 1979: 185).

Assim, para esta tradição metafisofica, os problemas filosóficos serão problemas causados pela “interpretação equivocada das nossas formas de linguagem” (Wittgenstein, 1953: § 111), isto é,

pela aplicação errônea de certos conceitos a outros conceitos. Como iremos analisar no próximo capítulo, este será o fundamento de um dos ataques mais ferozes da Filosofia à Neurociência contemporânea.

Aos ombros desta tradição e contra o movimento positivista surge William V. O. Quine que publica um influente artigo que viria a influenciar a história da relação entre a Filosofia e a Ciência. No seu “Dois Dogmas do Empiricismo” (1951), Quine atacou duas teses fundamentais:

(3) A distinção analítico-sintética entre verdades do tipo analítico – cuja verdade é alcançada pela mera análise do significado de cada termo e é independente de qualquer facto empírico; e verdades de tipo sintético – cuja verdade é fundamentada em factos empíricos;

(4) O reducionismo, a tese que defende que cada proposição com sentido é equivalente a uma construção lógica sobre termos simplificados pela análise lógica que se referem diretamente à experiência (cf. Quine, 1980: 20).

Contra (3), Quine defende que a distinção é falsa e que todas as crenças têm alguma ligação à experiência – o conceito de analiticidade é assim interpretado com ceticismo. No fundo, todas as defesas de analiticidade baseiam-se, para Quine, em explicações ou justificações circulares e, por isso, inválidas.

Já contra (4), Quine é cético em relação ao projeto reducionista dado que a conexão entre uma crença e a experiência nunca é direta. Lembre-se que um dos objetivos da análise lógica ao modo de Russell, por exemplo, era exatamente reduzir conceitos complexos em conceitos simples que teriam uma ligação direta ao mundo natural.

Assim, para Quine, quando a experiência colide com alguma crença, a crença específica que é problemática é indeterminada. Nas palavras de Quine, as crenças enfrentam “o tribunal da experiência sensorial não individualmente, mas como um órgão corporativo” (Quine, 1980: 41). Esta defesa holística do empirismo é uma marca da metafísica de Quine, que irá referir-se a esta estrutura como a “teia de crenças” (*web of beliefs*): algumas crenças, as que se encontram mais afastadas do centro da teia, estão mais expostas à experiência – e, por isso, a serem revistas – do que as que se encontram no seu centro. Mas, se uma crença for afetada, tal poderá levar à revisão de outras crenças da teia, sendo que nenhuma das crenças está livre do tribunal da experiência.

O essencial desta discussão é perceber as consequências metafísicas dos argumentos de Quine. É comumente afirmado que, com esta estratégia, Quine salvou a metafísica do positivismo que, como vimos, via com ceticismo extremo qualquer investigação metafísica de qualquer fenómeno. Dado que o critério verificacionista do positivismo cai com Quine, o conceito de “teia de crenças” implica que

até as crenças ontológicas – o que Quine chama de crenças sobre “os traços mais gerais da realidade” (Quine, 1960: 161) – são justificadas, em última instância, pela experiência.

Se tal é o caso, a diferença entre crenças do tipo ontológico e outras é apenas de grau ou generalidade. E é aqui que se dá a grande revolução metafilosófica – ou um retorno aos filósofos do passado: a metafísica passa a estar a par com a ciência natural, apenas diferindo em questão de grau, não sendo substancialmente diferente desta (cf. Quine, 1980: 45). Nesta metafilosofia, a Filosofia passa a ser definida ou entendida como uma extensão da Ciência (Glock, 2003: 30). Veremos como é que esta influência se dá em relação à Neurociência em particular.

Quine levanta, deste modo, suspeitas sobre o tipo de metodologias puramente *a priori*, sendo que “somente dentro da própria ciência, e não em alguma filosofia *a priori*, a realidade deve ser identificada e descrita” (Quine, 1981: 21). Ora, que consequências tem esta forma de ver o trabalho filosófico para a investigação filosófica em si? Para Quine, o trabalho do filósofo passa a ser o de “traduzir” as várias teorias científicas para o que chama de “notação canónica”, isto é, a lógica moderna desenvolvida por Russell, Frege, Peirce e outros (cf. Orenstein, 2002: 16). Mais importante, o filósofo deve tentar minimizar quaisquer compromissos ontológicos das teorias científicas.

Assim, embora o reducionismo positivista fosse mais radical, continuamos numa forma de prática filosófica do tipo reducionista, dada a contestação de uma das formas mais tradicionais de fazer filosofia: a procura de verdades necessárias – a analiticidade. Hilary Putnam, por exemplo, considera que Quine é o primeiro filósofo a construir uma metafilosofia sem o *a priori* como parte essencial da sua teoria (cf. Putnam, 1976: 202-213).

Um dos autores que viria a contestar a negação de analiticidade de Quine é Saul Kripke. Para Kripke, o trabalho de filósofo é diferente do anunciado por Quine. Através da renovação do conceito modal de necessidade, Kripke irá defender que uma verdade *x* sobre *y* é necessária quando *x* é verdadeira em todos os mundos possíveis que contem *y*. Num exemplo famoso neste debate, sabemos que água é H₂O neste mundo atual. Mas a questão, para o autor do *O Nomear e a Necessidade* (1980) é se tal afirmação é verdadeira em todos os mundos possíveis e não somente no mundo atual – a necessidade passa a ser um conceito modal. Assim sendo, Kripke vai defender uma tese controversa: a existência de necessidades *a posteriori*, algo que o filósofo considera estar no círculo da metafísica, e não da epistemologia.

O importante a reter, metafilosoficamente, é a seguinte ideia: com Quine e Kripke surge uma nova forma de fazer filosofia – nasceria assim o naturalismo. Mas este naturalismo assumirá distintas versões em diferentes domínios, embora haja uma correlação que se mantém. Por exemplo, o

naturalismo ontológico defende a tese de que apenas as entidades investigadas pela ciência fazem parte da realidade: qualquer que seja o fenômeno estudado, ele terá de ser explicado pelo apelo às várias leis da natureza (cf. Glock 2003: 46; Papineau, 2009: 10).

A versão metafilosófica do naturalismo irá defender uma continuidade profunda entre a Filosofia e a Ciência. Assim, os problemas filosóficos, em vez de serem resolvidos por métodos puramente *a priori*, poderão ser resolvidos ou esclarecidos através dos métodos das ciências empíricas. Dentro do naturalismo metafilosófico podemos distinguir dois tipos de posições:

- (i) O naturalismo metafilosófico fraco constituído por filósofos empíricos que fazem uso da ciência para ajudar a resolver os problemas filosóficos;
- (ii) O naturalismo metafilosófico forte, defendida por filósofos experimentais, que procuram eles próprios fazer a ciência (ou colaborar diretamente com cientistas) necessária para resolver os problemas filosóficos.

Há variados problemas com esta nova posição metafilosófica, como veremos nesta tese. Um dos mais relevantes é algo que Wittgenstein já tinha levantado em relação a uma possível estética naturalizada, isto é, uma estética puramente descritiva sem qualquer normatividade: “O tipo de explicação que se procura quando se está intrigado com uma impressão estética não é uma explicação causal, nem algo corroborado pela experiência ou por estatísticas sobre a forma como as pessoas reagem” (Wittgenstein, 1966: 17). Veremos se há uma maneira de ultrapassar esta lacuna metafilosófica.

Segundo Gary Gutting, a desconsideração da reflexão metafilosófica por parte dos filósofos dá-se por dois motivos principais: uma atitude dogmática em relação à natureza da Filosofia a partir de doutrinas filosóficas controversas – como a metafísica idealista ou a epistemologia empiricista – e uma abordagem demasiado abstrata ou geral que não oferece qualquer atenção aos detalhes da prática filosófica, ignorando assim os próprios pressupostos metafilosóficos que os filósofos usam nas suas investigações (cf. Gutting 2009: 2). Estas observações serão úteis para a investigação dos pressupostos filosóficos das diferentes abordagens metodológicas que iremos investigar mais à frente.

Moore afirmou famosamente que

Parece-me que na Ética, como em todos os outros estudos filosóficos, as dificuldades e divergências, de que está repleta a sua história, devem-se principalmente a uma causa muito

simples: a saber, a tentativa de responder a questões, sem antes descobrir precisamente que questão é que se deseja responder (Moore, 1991: vii).

Talvez esta intuição de Moore tenha real valor e nos ajude a progredir minimamente na construção de conhecimento mais sólido sobre a mente consciente.

Mas se esta breve introdução à disciplina da Metafilosofia apresenta tantas variedades do que é o trabalho filosófico, o que é que tal significa para a própria definição de Filosofia? Uma resposta pouco satisfatória é afirmar que Filosofia é uma variedade de investigações praticadas por filósofos profissionais: o eticista faz algo diferente do filósofo da mente ou do lógico – avançamos, assim, com uma definição deflacionária de Filosofia.

Nesta perspetiva, a Filosofia é tudo o que as universidades e bibliotecas catalogarem de filosofia e será difícil perceber efetivamente o que une todas as áreas que são abrangidas por este conceito geral. Se esse for o caso, encontramos-nos na seguinte situação descrita por Quine:

Filosofia [*sc.* é apenas] um dos variados termos usados por reitores e bibliotecários na sua tarefa necessária de agrupar os inúmeros tópicos e problemas da ciência e investigação académica num número manobrável de rótulos [*sc.* e as investigações de diferentes filósofos] estarem agrupados sob “filosofia” não torna nenhum dos homens responsáveis pelo tópico do outro (Quine, 1975: 228).

Mas será que esta versão deflacionária é uma boa estratégia? Repare-se que ela assenta em fundamentos puramente descritivos, isto é, observa empiricamente o que os filósofos das variadas áreas fazem e conclui que não parece haver nada que unifique áreas tão diferentes de investigação. Será que podemos encontrar outra estratégia para definir Filosofia com mais exatidão?

Essa é a tentativa do essencialismo, a posição metafilosófica que defende que podemos definir conceitos de forma explícita e exaustiva. Veja-se, por exemplo, a tentativa de definir filosofia através de um essencialismo metodológico: Filosofia é toda a atividade que fizer uso de métodos *a priori*. Assim explicitada, a definição inclui grande parte da Matemática e de outras disciplinas formais que também fazem uso de métodos *a priori*. Mas esta definição também excluirá formas de fazer filosofia atuais perfeitamente incorporadas nas universidades e centros de investigação, como é o caso da Filosofia Experimental e, por isso, esta estratégia levanta diversos obstáculos se for tida como correta.

Se, por um lado, parece que a Filosofia é demasiada inclusiva, por outro parece demasiada exclusiva. Será que podemos considerar que no meio é que está a virtude, como proposto pelo Estagirita?

O facto de não conseguirmos delimitar explicitamente o território da investigação filosófica não condena a própria disciplina à partida. Como afirmou Wittgenstein, o facto de dois países disputarem as suas fronteiras numa guerra por território não coloca em causa a cidadania de todos os seus habitantes (cf. Wittgenstein, 1967: § 556).

Ora, se a reflexão descritiva da Filosofia parece não oferecer uma resposta definitiva à questão de saber o que é de facto a Filosofia, tal não implica que não possamos fazer um melhor trabalho em relação à reflexão normativa da Filosofia, isto é, não como esta disciplina é de facto praticada, mas como deveria ser praticada.

Em relação ao problema normativo, há duas orientações extremas que será útil analisar:

- (i) A posição que defende a tese da “filosofia-como-ciência”;
- (ii) A posição que defende a tese oposta, da “filosofia-diferente-da-ciência”.

Em (i), a Filosofia é vista como sendo parte da Ciência e como tendo objetivos comuns: racionalizações sobre a verdade a respeito do mundo, sendo uma atividade cognitiva com finalidade epistémica. Já em (ii), afirma-se que a Filosofia tem objetivos distintos da Ciência, sendo um empreendimento fundamentalmente diferente desta, não sendo cognitiva e não sendo, primariamente, preocupada com a verdade.

Claro, ambas as posições metafilosóficas possuem uma clareza duvidável no que defendem devido à falta de uma maior explicitação do que significa o conceito central de (i) e (ii): “ciência”. Se, por um lado, aceitarmos que o termo “ciência” designa somente as investigações da ciência empírica, isto é, a tentativa de ler o mundo pelas leis da natureza através da observação e experimentação, então a maioria dos filósofos incluir-se-iam em (ii), sendo que apenas os filósofos experimentais fariam parte de (i).

Todavia, talvez essa forma de definir “ciência” seja demasiado restrita: muitas ciências que atualmente são totalmente aceites como tal – como a Física Teórica – seriam igualmente excluídas dessa definição. Assim, se o termo “ciência” incluir, de forma mais abrangente, métodos formais de investigação como a matemática ou a lógica formal, então grande parte dos filósofos já faria parte de (i) sem grande dificuldade.

Quiçá tal aponte para algo que já intuímos anteriormente: talvez tenhamos mais sucesso se considerarmos a Filosofia como um contínuo entre (i) e (ii), em vez de a considerar como uma extremidade descontínua de uma linha que nos parece ser tendencialmente contínua. E como veremos

nesta reflexão, esse poderá ser o principal erro de algumas metodologias, e a principal vantagem de outras.

Mas antes de começar tal empreendimento, faltará introduzir brevemente também a subárea da Filosofia da Mente que este trabalho focará especialmente: a Filosofia da Neurociência.

I.3. A Filosofia da Neurociência: Introdução e Análise

A Filosofia da Neurociência é uma área recente cujo foco se centra na análise dos fundamentos e metodologias da Neurociência como disciplina empírica. O objetivo desta área é analisar filosoficamente as questões fundacionais levantadas pelas ciências do cérebro. Historicamente, um interesse alicerçado na Neurociência pela Filosofia acontece em 1950 com o surgimento das teorias da identidade-tipo para resolver o intrincado problema mente-corpo.

Diferentes versões foram avançadas por Place (1956) e Smart (1959), sendo que ambas apontavam para a ideia de que a informação neurocientífica sobre o cérebro teria de ter um contributo fundamental para resolver o problema filosófico. Uma segunda etapa acontece com a publicação, em 1986, do impactante livro *Neurofilosofia* de Patricia Churchland, onde, pela primeira vez, a investigação em Neurociência era apresentada a filósofos e investigação em Filosofia era apresentada a neurocientistas e cientistas cognitivos.

Todavia, há alguma controvérsia em limitar a Filosofia da Neurociência à Neurofilosofia dado que a última representa, como veremos adiante, uma posição com os seus princípios fundamentais e com um foco metodológico muito preciso e definido, pelo que será mais vantajoso interpretar a Neurofilosofia como uma parte da Filosofia da Neurociência, mas não esgotando a área de investigação na sua totalidade.

Nesse sentido, devemos distinguir entre Neurofilosofia e Filosofia da Neurociência, sendo que a segunda tem o foco na investigação de problemas filosóficos internos à Neurociência: a análise dos conceitos teóricos da Neurociência, o foco na(s) metodologia(s) de investigação da Neurociência, a conexão epistémica entre os vários níveis de estudo da Neurociência e da sua relação com disciplinas adjacentes, como a Psicologia, a Inteligência Artificial ou a Psiquiatria, os pressupostos ontológicos da área, as ferramentas da análise da atividade cerebral, as questões fundacionais, entre muitos outros enigmas.

Podemos, assim, adaptando de Gold e Roskies (2008), apontar três principais focos da Filosofia da Neurociência:

- (5) Compreender a forma, estrutura e natureza da Neurociência como um todo e como ciência;
- (6) Analisar ou explicitar conceitos, princípios, métodos e questões fundamentais da Neurociência;
- (7) Elucidar questões tradicionais da Filosofia incorporando dados empíricos da Neurociência (cf. Gold & Roskies, 2008: 2).

Vejamos, brevemente, as principais questões de cada um destes focos de investigação de modo a construirmos uma visão geral da mesma. Logo de seguida, analisaremos a relação que a Filosofia da Neurociência tem com outras áreas adjacentes da Filosofia.

I.3.1. Natureza e Estrutura da Neurociência

Um primeiro problema filosófico que se levanta com a Neurociência é perceber a sua natureza fundamental como área de investigação. Tal é primariamente fundamental considerando uma característica essencial desta área: ser pobre em teoria, e rica em dados empíricos. Em termos teóricos, a Neurociência assume um pressuposto fisicalista radical e simplista com variações computacionais sem lugar para qualquer tipo de exceções aos modelos baseados nas leis da física e da química. Tal ausência de teorias fundamentais pode ser um sinal da pouca maturidade desta área de investigação que, tentando imitar metodologicamente a rainha das ciências – a Física – se encontra muito afastada desta, tanto em termos de poder explicativo dos fenómenos que pretende investigar, como em poder preditivo das suas investigações.

Embora haja uma ausência de uma teoria geral – como a teoria da relatividade geral na Física – que possa servir de meta-fundamento para explicar todos os processos mentais e/ou cerebrais, a Neurociência como disciplina é uma área riquíssima em dados empíricos que, acumulados entre si, dão origem a modelos teóricos particulares que podem servir de fundamento para a explicação de funções específicas do cérebro, como a percepção, a memória ou a consciência.

O ponto central destes modelos é que apenas se aplicam ao fenómeno que querem explicar (e.g. percepção) e não servem de explicação para outro fenómeno mental (e.g. memória) pelo que cada

fenómeno terá um modelo particular que não pode ser replicado para explicar outro fenómeno mental (cf. Gold & Roskies, 2008: 4).

Ora, apesar da limitação epistémica destes modelos particulares e da falta de uma teoria geral e unificadora (no capítulo 8 apresentaremos a mais recente tentativa de apresentar uma abordagem unificadora), a Neurociência tem produzido muito conhecimento empírico: sabemos como um neurónio processa informação recebida de outro neurónio, sabemos a estrutura anatómica e celular de vários tipos de célula nervosa, sabemos como é que vários neurónios trabalham em conjunto para criar redes neuronais que são ativadas em funções específicas, entre outras informações empíricas análogas.

Não obstante, a natureza desta ciência empírica é vaga filosoficamente: trata-se de um conjunto afluyente de dados e factos sobre o cérebro e o sistema nervoso em geral, inúmeros modelos particulares para diferentes fenómenos, mas tendo falta de uma teoria unificadora e geral que una epistemicamente este volume enorme de conhecimento experimental.

Estas características obrigam-nos a tirar ilações sobre a natureza da própria disciplina: será isso um sinal da imaturidade da Neurociência, se comparada com ciências clássicas como a Física? Será que a falta de um princípio fundamental e unificador limita a Neurociência como ciência? Será que temos algum conhecimento sustentável e causal – e não vago e meramente correlacional – sobre as funções do cérebro e a sua relação com as capacidades mentais de cada um? Será que esta ciência tem uma natureza excepcional que difere de ciências mais maduras? (cf. Hardcastle, 1999).

No mesmo sentido estão as considerações que a Filosofia da Neurociência pode levantar a respeito da Filosofia da Ciência, isto é, a forma como entendemos como uma ciência deve ser. Um exemplo que será focado com especial atenção no capítulo 3 é o da defesa do reducionismo – e da sua relação com níveis de explicação – da Psicologia à Neurociência empírica: será essa redução plausível e desejável? Ganhamos poder explicativo e preditivo com tal manobra? O que acontece aos fenómenos e conceitos que serviam de base à ciência reduzida (e.g. o inconsciente de Freud)?

Todas estas preocupações levantam considerações interessantes sobre a natureza da disciplina, sendo este primeiro problema um dos principais focos da Filosofia da Neurociência.

I.3.2. Conceitos, Princípios e Métodos da Neurociência

Um segundo foco da Filosofia da Neurociência é identificar conceitos e princípios fundamentais da Neurociência. Desde o nascimento desta disciplina, a Neurociência tem assumido, por exemplo, o

princípio de localização funcional, isto é, a ideia de que funções particulares do cérebro (e.g. percepção visual) estão relacionadas com localizações anatómicas particulares (e.g. córtex visual primário) (cf. Mundale, 2002; Hardcastle & Stewart, 2002; Lloyd, 2000). Este princípio é muito relevante, filosoficamente, porque influencia metodologicamente a própria investigação empírica em Neurociência: a forma como as experiências neurocientíficas são desenhadas e, principalmente, a forma como os dados são interpretados.

Por conseguinte, este princípio levanta inúmeras questões filosóficas. Se fizermos uma leitura forte do princípio, estaremos a afirmar que, factualmente, partes específicas do cérebro se relacionam causalmente com fenómenos mentais específicos. Esta interpretação é, como se compreende, epistemicamente exigente, pressupondo consequências ontológicas duvidáveis que se querem provar à partida e não se podem assumir com antecipação.

Porém, uma leitura fraca do princípio pode também ocorrer: podemos ler o princípio como uma metáfora metodologicamente útil para acumular informação sobre o cérebro humano, embora as consequências ontológicas de tais conhecimentos se mantenham neutras.

O próprio desenvolvimento tecnológico da área, isto é, o desenvolvimento de técnicas neuroimagéticas mais sofisticadas (e.g. fMRI) pode pôr em causa este princípio, particularmente com a descoberta de que diversas áreas distintas e altamente diferenciadas do cérebro estão envolvidas em processos mentais que se acreditava, anteriormente, necessitarem de apenas uma parte específica do cérebro (cf. Haxby *et al.*, 2001; Hanson *et al.*, 2004).

Outra questão fundamental relaciona-se com a metodologia da maioria das técnicas de investigação do cérebro: que relação podemos inferir entre os dados recolhidos pelos neurocientistas, as estruturas cerebrais ativadas durante as medições, e as funções cognitivas investigadas? (cf. Gold & Roskies, 2008: 7).

Como se percebe, estes debates pressupõem inúmeras clarificações conceptuais e metodológicas que caem no escopo da Filosofia da Neurociência, sendo estas um segundo foco relevante desta área.

I.3.3. Questões tradicionais da Filosofia e Neurociência

A Filosofia da Neurociência também se foca no estudo de conceitos e fenómenos que eram anteriormente estudados exclusivamente pela Filosofia. Um desses principais fenómenos é a existência da mente

consciente: como é que um conjunto de partículas sem consciência, combinadas entre si de forma específica, podem criar uma experiência consciente e subjetiva? Esta questão tem uma longa tradição na história da Filosofia, sendo a explicação dualista cartesiana uma das mais relevantes contribuições da investigação tradicional filosófica.

Contudo, talvez o método da Filosofia tradicional não seja suficiente para lidar com a consciência num mundo físico: talvez seja necessário combinarmos algumas das ferramentas filosóficas com as ferramentas da Neurociência ou, mais radicalmente, fazer somente uso das segundas em detrimento das primeiras.

A crença de que somente através do relato subjetivo e introspectivo podemos alcançar conhecimento sobre a consciência tem sido substituída por uma abordagem da terceira pessoa que pretende identificar com exatidão, entre outros fenómenos, os correlatos neuronais da consciência (NCC) (cf. Crick & Koch, 1998; 2003). Ora, tal progresso levanta questões sobre o próprio conceito: será que o conceito de consciência ainda faz sentido? Será que existem dois aspetos fundamentais deste conceito, como argumentado por Ned Block (2005), que distingue consciência de acesso de consciência fenomenal?

Se tal for o caso, estas questões requerem uma avaliação crítica dos variados resultados científicos que poderão ser interpretados de diferentes formas consoante os pressupostos teóricos assumidos sobre o fenómeno a investigar em questão (cf. Gold & Roskies, 2008: 9). Este trabalho, especificamente voltado para a consciência, mas também para outros fenómenos estudados tradicionalmente pela Filosofia, é um terceiro foco do tipo de investigação que um filósofo da neurociência procura fazer.

I.3.4. Filosofia da Neurociência, Relação com outras Áreas da Filosofia e Novas Áreas de Investigação

Dos três focos principais descritos brevemente nas secções anteriores, fará sentido notar que as grandes questões desta área de investigação estão em contacto direto ou indireto com outras áreas da Filosofia, como a Epistemologia, a Metafísica ou Ontologia, a Filosofia da Mente, a Filosofia da Ciência, entre outras. Vejamos algumas justaposições nesse sentido:

- **Filosofia da Mente:** esta subárea da Filosofia estuda problemas centrais ligados à mente humana. Será que a mente e o cérebro são, ontologicamente, distintos? Ou, pelo contrário, partilham a mesma substância? Este é um debate que se levanta com o tradicional problema mente-corpo. Um problema diferente desta área tem a ver com o conteúdo da nossa percepção: quando percecionamos o mundo, por exemplo, uma flor, será que temos uma relação direta com esse objeto? Ou essa relação é indireta? Ou apenas conseguimos percecionar uma ideia desse objeto, não podendo ter acesso ao objeto em si? Estas questões fazem parte do problema da percepção, outra indagação desta área de investigação. O problema do acesso a outras mentes, isto é, se é justificado atribuir mente a outras pessoas ou robots, se apenas tivermos acesso ao seu comportamento físico e/ou verbal, é também outro problema estudado na área. Ora, a relação destes problemas com a Filosofia da Neurociência torna-se clara: se a Neurociência nos demonstrar, de algum modo, que a mente não é nada mais que um produto da atividade cerebral, então este conhecimento irá resolver, por exemplo, um problema tradicional da Filosofia da Mente, nomeadamente o problema mente-corpo; um conhecimento total do cérebro humano poderá influenciar também a forma como podemos conceber uma inteligência artificial consciente; e assim sucessivamente.
- **Epistemologia:** dado que a Neurociência tem como foco compreender e explicar o funcionamento de um sistema biológico particular, o sistema nervoso, tal empreendimento pode esbarrar em variados problemas da Epistemologia, incluindo a natureza e a justificação do conhecimento, a plausibilidade das teorias e crenças científicas, a relação entre o conhecimento do cérebro e as capacidades epistémicas do mesmo, entre outros problemas (cf. Churchland, 2002a: 270; Northoff, 2004: 208). Poderá falar-se, inclusive, de uma Neuroepistemologia, como avançado por Zambrano (2012). Será que o conhecimento obtido pela Neurociência tem solidez suficiente para ser interpretado com um facto do mundo? Será que esse conhecimento é justificado pelo apelo à mera experimentação, ou é necessário algo mais? Estas e outras perguntas entram em contacto com a Epistemologia.
- **Metafísica:** um dos principais focos da Filosofia da Neurociência é o estudo do estatuto ontológico da mente humana. Tal foco requer considerações de carácter metafísicas que são influenciadas pelo conhecimento da Neurociência. Por exemplo, se o materialismo eliminativista estiver correto, então estados mentais reduzem-se a estados cerebrais, sendo que tal redução

clarifica ontologicamente o estatuto da mente humana como sendo material e como sendo parte do escopo explicativo das leis da natureza (Churchland, 1986). Outros problemas podem ser esclarecidos pela investigação empírica: por exemplo, se se demonstrar que as nossas decisões não são livres, mas que o nosso cérebro já pré-definiu a ação que acreditamos ser livre, tal conhecimento irá ter impacto direto no problema metafísico do determinismo – a ideia de que tudo no universo está já determinado pelas leis da natureza, inclusive a ação humana – *versus* a existência do livre-arbítrio – a tese de que podemos escolher entre variadas ações e que essa escolha não é pré-determinada por outros fatores externos à própria escolha (cf. Libet, 1985).

- Filosofia da Ciência: um dos principais objetivos da Filosofia da Ciência é oferecer uma definição precisa de explicação. Uma das abordagens principais deste conceito argumenta que a explicação de um fenómeno requer uma dedução lógica de leis específicas em relação ao fenómeno que se quer explicar em primeira instância (cf. Mandik *et al.*, 2001: 8). Esta abordagem, conhecida como modelo dedutivo-nomológico, relaciona leis gerais com eventos particulares. Além de explicar fenómenos ou eventos, esta abordagem pode também ser aplicada à relação entre uma ciência particular e uma ciência mais básica onde se relacionam os diversos vocabulários conceptuais com o objetivo de se alcançar uma derivação entre a ciência reduzida e a ciência mais básica. Ora, esta abordagem foi criticada historicamente pela defesa de que o raciocínio lógico que serve de fundamento entre as relações indicadas é falacioso, cometendo a falácia da afirmação do consequente (cf. Popper, 1959). Outras críticas foram levantadas da abordagem dedutiva-nomológica, como em Kuhn (1970) ou Betchel e Richardson (1993). Ora, o ponto que queremos focar é que este debate afeta diretamente a forma como podemos entender a disciplina da Neurociência e a sua relação com as outras ciências, como a Psicologia ou a Biologia Celular. Assim sendo, os conhecimentos e metodologias da Neurociência levantam considerações filosóficas relevantes que devem ser pensadas em relação à Filosofia da Neurociência.

Poderá dar-se o caso, também, de existir a renovação de áreas tradicionais da Filosofia e a criação de novas áreas multidisciplinares, como a Neuroética ou Neuroestética. Vejamos alguns exemplos ilustrativos:

- Neuroética: esta área recente assume a ideia fundamental de que a investigação e o progresso em Neurociência podem elucidar ou iluminar alguns dos problemas éticos tradicionais investigados pela Filosofia (cf. Roskies, 2002). O objetivo central desta área é tentar relacionar aspetos da neurobiologia do comportamento moral com problemas e soluções éticas. Veja-se a famosa experiência mental avançada por Philipa Foot, o dilema do elétrico. Os neurocientistas pegaram neste exercício filosófico e tentaram analisar o que se passa nos cérebros dos agentes quando estão a fazer a escolha – mudar o botão ou empurrar o homem ou não o fazer: estudos em Neurobiologia demonstraram o que se estava a passar no cérebro durante a apresentação do dilema e a escolha consequente – mas não o que *deveriam* fazer. Através do mapeamento cerebral, a investigação descobriu que na primeira versão do exercício as pessoas ‘usam’ mais a parte orbitofrontal e ventromedial do cérebro, onde estão “localizados” o sentido de equidade, do pensamento lógico e racionalidade; já na segunda versão do dilema, a maior parte da atividade cerebral centraliza-se na amígdala, área responsável pelas respostas emocionais. O que aconteceu com a maior parte das pessoas ao dizerem que “sim”, na primeira versão, e “não”, na segunda versão, é que na primeira fizeram uma análise fria e distanciada do cenário, usando o pensamento lógico; já no segundo caso, imaginámo-nos naquela situação de ter de empurrar, literalmente, uma pessoa, o que causa uma resposta emocional (cf. Green *et al.*, 2004). Ora, este é um exemplo do tipo de conhecimento que podemos originar cruzando a Neurociência com a Ética, sendo que este cruzamento parece legitimar que se fale numa disciplina independente tanto da Neurociência propriamente dita, como da Ética ou Filosofia Moral.
- Neuroestética: algo semelhante acontece com a recente disciplina que cruza a Filosofia da Arte ou Estética com a Neurociência. Dado que a experiência estética é pelo menos parcialmente uma experiência que envolve os mecanismos da percepção – principalmente visual, mas também auditiva ou sensitiva – esta área procura investigar a neurobiologia associada à percepção especificamente estética. Publicações fundamentais desta área são *Inner Vision* (1999) de Zeki, onde o autor argumenta que a arte moderna é influenciada pela forma como o sistema visual funciona, ou a análise do sorriso da Mona Lisa apresentada em *Vision and Art: The Biology of Seeing* (2002), de Margaret Livingstone, que defende que o aspeto misterioso do sorriso da Mona Lisa é resultado da anatomia da retina de cada observador. Este é um exemplo do tipo de conhecimento que é gerado nesta nova área de investigação.

Percebemos, desta forma, que estamos a lidar com uma disciplina complexa com muitas e diversas implicações para outras áreas da Filosofia ou até de novas disciplinas. Esta complexidade será um desafio interessante dado que obrigará a investigação a saltar de área em área, cruzando conceitos, métodos e teses de cada uma, originando novas ou contestando concepções mais tradicionais.

Restará apresentar nesta introdução, muito brevemente, algum conhecimento básico da Neurociência como disciplina empírica de modo a servir de auxílio para vários debates que acontecerão nos restantes capítulos e que envolverão, obrigatoriamente, a discussão de conhecimentos das ciências do cérebro.

I.4. Breve Introdução à Neurociência

Pode ser argumentado que as primeiras reflexões sobre o cérebro acontecem com os Gregos Antigos que acreditavam que o coração era a sede da inteligência, dos sentimentos e da consciência.⁴ O pai da Medicina, Hipócrates, defendia, pelo contrário, que essas capacidades “mentais” estariam relacionadas com funções do cérebro, tendo observado que lesões particulares no mesmo causavam paralisias, perdas de movimento, demência ou perdas de memórias, associando, assim, essas capacidades ao cérebro. Mais tarde, Aristóteles volta a considerar que as funções mentais “acontecem” no coração, relegando o cérebro para um mero sistema de arrefecimento do sangue.

Teríamos de esperar pelo Império Romano para que a sede das funções mentais voltasse ao cérebro, aos ombros do médico Galeno e, um milénio mais tarde, Vesalius, considerado o pai da anatomia moderna que defendia que os ventrículos do cérebro e os seus fluidos estariam relacionados com a percepção, o pensamento, a consciência ou inteligência.

Ademais, argumentava que os tecidos envolventes – a matéria branca e cinzenta, hoje consideradas partes fundamentais – teriam apenas um papel de suporte aos ventrículos e fluidos. Curiosamente, devido a esta associação muito particular, Vesalius viria a defender que tal facto explicava por que razão a nossa inteligência e experiência consciente era igualmente fluida e oscilante – literalmente (cf. Northoff, 2014a: 73-74).

⁴Secção baseado em Northoff, 2014a: 69-90. Todas as referências poderão ser aí encontradas.

Embora a Neurociência integralmente formalizada apenas exista há cerca de 50 anos, os primeiros passos mais “científicos” e menos especulativos desta jovem ciência começaram em 1810 com a Frenologia de Franz-Joseph Gall e J. G. Spurzheim, que procuraram atribuir funções específicas a diferentes partes do cérebro. Essas funções variavam desde a linguagem à esperança e à autoestima. De imediato surgiram reações críticas a este método localizacionista.

Em 1824, o fisiologista experimental Jean-Pierre Flourens rejeitava a ideia fundamental de que certas funções estavam confinadas a certas regiões específicas do cérebro; em Inglaterra, o neurologista Hughlings Jackson, com os seus trabalhos com doentes epiléticos, notou que muitas regiões diferentes do encéfalo contribuíam para a mesma função; finalmente, em 1861 surgia o famoso caso relatado por Paul-Pierre Broca, que parecia confirmar que uma lesão numa determinada zona do cérebro – o lobo esquerdo frontal, que viria a ser conhecido posteriormente como Área de Broca – originava uma lesão específica.

Este foi o tema escolhido pelo neurologista alemão Karl Wernicke para o seu doutoramento. Em 1876, Wernicke estudou uma vítima de um acidente vascular cerebral que conseguia falar quase normalmente, mas tudo o que dizia não fazia qualquer sentido. Mais tarde, os trabalhos de Camillo Golgi, que tinha desenvolvido uma técnica de coloração que permitia a identificação de neurónios individuais através de prata líquida, permitiram a visualização de um único neurónio.

Aplicando este método, o médico espanhol Santiago Ramón y Cajal viria a descobrir que, ao contrário do que se pensava até então, os neurónios eram entidades únicas, com uma natureza unitária determinada, além de ter descoberto que a transmissão da informação elétrica se dava apenas das dendrites para os axónios e não bidirecionalmente, como se defendia até então.

Estas descobertas causaram enorme espanto aos pensadores da época, que começavam a conseguir ligar aspetos mentais a partes específicas do cérebro, o que indicaria alguma relação necessária entre elas. Tal poderá ser identificado nas obras do neurocientista inglês Charles Sherrington que, apesar de defender um dualismo de tipo cartesiano (como veremos no capítulo seguinte), começava a identificar a verdadeira natureza da espinal medula como centro dos reflexos, tal como o papel do córtex cerebral para a produção da resposta reflexa.

Igualmente relevante para o desenvolvimento da Neurociência foi o estudo de doenças neurológicas. Erwin Strauss, Kurt Goldstein e Hartmut Kuhlenbeck investigaram o cérebro com lesão de diversos soldados da Grande Guerra. Primeiramente, focaram-se na localização particular dessas lesões; de seguida, observaram o seu comportamento e funções cognitivas; finalmente, e de forma interessante, perguntaram aos pacientes de que forma as suas lesões afetaram as experiências do seu corpo e da sua

relação com o mundo, conectando, indiretamente, descrições fenomenais da consciência a características neuronais do cérebro – veremos, no quarto capítulo, a importância de conectar a Fenomenologia com a Neurociência (cf. Northoff, 2014a: 75-76).

O progresso da Neurociência está também associado ao desenvolvimento tecnológico de ferramentas para estudar o cérebro de forma mais direta. Um dos primeiros momentos relevantes desta relação acontece com o investigador alemão Hans Berger que registou, pela primeira vez, atividade elétrica no crânio de seres humanos. Tendo refinado esta tecnologia ao longo dos anos, alcança nos anos 30 a técnica da eletroencefalografia (EEG), conseguindo identificar, através da sua aplicação, ritmos e flutuações particulares da atividade elétrica no cérebro, como a famosa “onda alfa”, com frequências entre os 8 e 12 Hz.

Outras tecnologias relevantes que surgiram foram a tomografia por emissão de positrões (PET) e a imagem por ressonância magnética funcional (fMRI) que não medem, diretamente, a atividade elétrica, mas, antes, medem a atividade ou concentração bioquímica e metabólica: por exemplo, na fMRI podemos medir a libertação exata do número de moléculas de oxigénio no neurónio e inferir, posteriormente, o valor da atividade elétrica. Já a PET faz uso de substâncias radioativas para identificar recetores específicos ou diferentes balanços bioquímicos.

A grande diferença entre as tecnologias EEG e fMRI/PET é o tipo de resolução que conseguimos alcançar: com a primeira, temos uma medida temporal de alta resolução (na ordem dos milissegundos) mas baixa resolução espacial. Pelo contrário, com fMRI/PET alcançamos elevada (muito detalhada) resolução espacial, e pouca resolução temporal (na ordem dos segundos) (cf. Northoff, 2014a: 81-83).

Ficamos, assim, com uma introdução brevíssima à Neurociência, ficando claro a importância da evidência empírica – e menos das ferramentas apriorísticas da Filosofia – no desenvolver desta área de investigação. De seguida, começaremos a nossa reflexão doutoral analisando a abordagem isolacionista ao problema metodológico entre a Neurociência e a Filosofia.⁵

⁵ Uma introdução muito útil a esta área de estudo pode ser encontrada em Churchland (1986: 35-235). Uma análise mais detalhada e pormenorizada da mesma pode ser encontrada em Kandel & Schwartz (1981).

II. ISOLACIONISMO METODOLÓGICO

Neste segundo capítulo procuraremos introduzir a primeira abordagem que tenta lidar com o problema apresentado no capítulo introdutório. Esta abordagem, a que chamaremos de abordagem isolacionista ou isolacionismo metodológico, procura dar uma resposta direta ao problema: este, de facto, não existe, sendo apenas uma obscura confusão teórica. Para tal, serão apresentados os principais argumentos presentes na obra *Philosophical Foundations of Neuroscience* (2003) do filósofo Peter Hacker e do neurocientista Max Bennett, trabalho que exemplifica perfeitamente este tipo de abordagem. Ademais, procuraremos demonstrar por que razão Hacker e Bennett poderão estar equivocados nos seus pressupostos metafilosóficos, especificamente sobre as definições de Filosofia e Neurociência. Concluiremos que, embora as suas teses não pareçam plausíveis de ser sustentadas, apontam para ideias interessantes que devem merecer mais reflexões e estudos.

As discussões mais recentes da Filosofia da Mente contemporânea têm levantado questões interessantes sobre o método que a mesma deve adotar para lidar com os grandes problemas do seu escopo de trabalho. A Filosofia solitária *a priori* tem vindo a incorporar os resultados das investigações em Neurociência e Ciência Cognitiva em geral.

Contudo, tal abordagem poderá ser ineficaz: afinal, a Filosofia existe desde, pelo menos, a Antiga Grécia, e o seu método nada parece ter a ver com a disciplina recente da Neurociência, que não deverá ter mais de cinquenta anos como área de estudo formalizada e sistematizada.

De um lado, muito sucintamente, possuímos o método lógico-argumentativo da Filosofia, que procura analisar proposições e argumentos; do outro, temos o método baseado na observação e experimentação da Neurociência.

A literatura científica da área tem equacionado vários problemas de resolução difícil (senão impossível) tanto com o uso do método filosófico, como com o uso do método científico. Haverá, então, espaço para uma fusão metodológica que consiga orientar as investigações para um progresso tão esperado?

Analisaremos de seguida uma resposta negativa a esta questão, que procura defender que as duas disciplinas são essencialmente diferentes e, por isso, não podem trabalhar em conjunto.

II.1. Abordagem do Isolacionismo

Uma possível resposta ao problema metodológico é a abordagem isolacionista, que também pode ser apelidada de “Paralelismo Metodológico” (cf. Northoff, 2014a: 144). O seu nome oferece-nos algumas pistas: é “isolacionista” porque a sua principal afirmação é assertar que a Filosofia e a Neurociência não podem trabalhar em conjunto, visto que são duas disciplinas dissemelhantes e, por isso, isoladas. A primeira lida com possibilidades lógicas e com a análise rigorosa de argumentos e proposições – uma metodologia essencialmente *a priori*. A segunda lida com investigação empírica e dados – uma metodologia essencialmente *a posteriori*.

Assim, é simplesmente incorreto, segundo esta abordagem, aceitar que duas disciplinas tão distintas quanto ao seu método e objetivos possam trabalhar em conjunto para lidar com os grandes problemas que a mente consciente causa à ciência contemporânea. Ambas são mutuamente independentes e não se podem confundir.

Veja-se o seguinte comentário que aponta nesse sentido:

Os projetos isolacionistas são baseados no pressuposto de que as descobertas das ciências naturais, incluindo a neurociência, não influenciam diretamente a prática da filosofia, sendo que pelo menos alguns aspetos da reflexão filosófica são independentes dos factos empíricos investigados pelos cientistas (Brozek *et al.*, 2013: 16).

Para demonstrar este tipo de perspetiva sobre o problema metodológico, iremos apresentar um exemplo importante que ilustra a resposta isolacionista na perfeição. Na aclamada obra *Fundamentos Filosóficos da Neurociência* (2003), o neurocientista M.R. Bennett e o filósofo P. Hacker apresentaram uma análise da linguagem usada pelos vários neurocientistas cognitivos para descreverem as suas investigações e experimentações.

Esta obra, considerada por muitos como “a mais significativa contribuição para a filosofia da mente nos últimos anos” tendo examinado “minuciosa e cuidadosamente as pretensões da ciência cognitiva de ter tornado obsoleta a psicologia filosófica” (Kenny, 2013: xvii), procurará ser uma defesa explícita desta primeira resposta ao problema metodológico.

O seu principal objetivo é demonstrar as inúmeras confusões conceptuais que têm influenciado a Neurociência, desde a sua fundação até à atualidade. Sucintamente, é argumentado que os conceitos neurocientíficos são, por definição, diferentes dos conceitos filosóficos. Os últimos são puramente

lógicos, são sobre sentido e significado. Já os conceitos neurocientíficos, em contraste, são sobre a verdade e a realidade.

Por exemplo, o conceito de atividade do córtex pré-frontal não possui qualquer sentido ou significado em si mesmo, é um facto puro que pode ser observado e medido. Tal é fundamentalmente diferente dos conceitos filosóficos: o conceito de livre-arbitrio possui um significado e um contexto semântico. Assim, há uma diferença de princípio, e por isso não podemos unir os dois conceitos. A interação entre a Filosofia e a Neurociência é assim tida como impossível.

Veja-se a seguinte citação dos autores que aponta nesse sentido:

As questões empíricas acerca do sistema nervoso são do domínio da neurociência. É um assunto seu estabelecer a matéria de facto no que respeita às estruturas e operações neurais. É tarefa da neurociência cognitiva explicar as condições neurais que tornam possíveis as funções perceptivas, cognitivas, cogitativas, afetivas, volitivas. Essas teorias explicativas são confirmadas ou infirmadas pelas investigações experimentais. Em contrapartida, as questões conceptuais (respeitantes, por exemplo, aos conceitos de mente ou memória, pensamento ou imaginação), a descrição das relações lógicas entre conceitos (tais como entre os conceitos de percepção e sensação, ou os conceitos de consciência e autoconsciência) e o exame das relações estruturais entre diferentes campos conceptuais (tais como entre o psicológico e o neural, ou o mental e comportamental) são domínio próprio da filosofia (Bennett & Hacker, 2003: 1).

Há, portanto, uma delimitação clara por parte dos autores entre as duas disciplinas. O seu foco de estudo é claramente identificado e não problemático: a Neurociência deverá lidar com problemas de apurar se uma determinada afirmação sobre a mente humana é verdadeira ou falsa; já a Filosofia deverá focar-se no sentido (ou na falta de sentido) dessa mesma proposição.

Tal posição deve-se ao facto de os autores defenderem o pressuposto metafilosófico seguinte: as questões conceptuais sobre x são anteriores (metodologicamente) à verdade ou falsidade de x . A análise conceptual determina não o que é verdadeiro ou falso, mas aquilo que faz ou não sentido (cf. Bennett & Hacker, 2003: 2).

Por isso mesmo, é falso que a Filosofia possa contribuir positivamente, pelo menos diretamente, para a investigação empírica da Neurociência dado que as questões conceptuais não podem, por definição, serem informadas ou resolvidas por investigação empírica proveniente da Neurociência.

Vejamos o argumento que resume a tese isolacionista:

- (1) A Filosofia é metodologicamente baseada no *a priori*;
- (2) A Neurociência é metodologicamente baseada no *a posteriori*;
- (3) Metodologias *a priori* e *a posteriori* são incompatíveis;
- (4) Logo, a Filosofia e a Neurociência são metodologicamente incompatíveis.

Deste modo, percebemos a estrutura fundamental da abordagem isolacionista ao problema metodológico: como se assume (3), isto é, que metodologias de diferentes fundamentos são inconciliáveis entre si, e como (1) a Filosofia é baseada num tipo de metodologia, o *a priori*, e (2) a Neurociência é baseada no *a posteriori*, então conclui-se a tese isolacionista (4): a Filosofia e a Neurociência são metodologicamente inconciliáveis.

II.1.1. Os Conceitos antecedem os Factos

Um segundo passo teórico desta abordagem passa pela afirmação de que os próprios conceitos e relações conceptuais trabalhados pela Filosofia estão já pressupostos pelas investigações empíricas da Neurociência, pelo que se segue também que a Neurociência nada possa contribuir para resolver os tradicionais problemas da Filosofia.

Deste modo, Bennett e Hacker irão demonstrar a Filosofia como forma de elucidar a nossa linguagem, sendo assim distinta do empreendimento científico. O objetivo desta elucidação é distinguir o que faz sentido do que não faz sentido, logicamente considerado.

Tome-se em consideração as seguintes afirmações de três influentes neurocientistas:

- (5) F. Crick: “O que você vê não é o que está realmente aí; é o que o seu cérebro acredita que está aí... O seu cérebro faz a melhor interpretação que pode [...] O cérebro combina a informação fornecida [...] e decide qual é a interpretação mais favorável” (Crick, 1995: 30-57);
- (6) G. Edelman: “[...] as estruturas do cérebro categorizam, discriminam e recombina as diversas atividades cerebrais que ocorrem em diferentes tipos de mapeamentos globais” (Edelman, 1994: 109);

(7) A. Damásio: “[...] os nossos cérebros podem muitas vezes decidir corretamente, em segundos ou minutos [...] têm de fazer esse trabalho maravilhoso com mais alguma coisa do que apenas a razão” (Damásio, 1996: 173).

Haverá algum problema com (5), (6) e (7)? À primeira vista, parecem afirmações perfeitamente plausíveis declaradas por prestigiados neurocientistas. Afinal, porque haveriam reputados cientistas de fazer afirmações erradas? De facto, não parece haver qualquer problema em expor que o cérebro categoriza, discrimina, decide, vê ou faz interpretações, sendo tal afirmado por qualquer cientista regularmente, seja em palestras informais ou publicações com revisão por pares.

Contudo, para os autores desta primeira abordagem, as três afirmações cometem erros conceptuais flagrantes que evidenciam perfeitamente o papel que a Filosofia deve ter neste debate.

Especificamente, argumentam que muitos neurocientistas atuais cometem uma falácia – embora os detalhes sejam diferentes de autor para autor, todos eles cometem o mesmo erro lógico – ao confundir os predicados aplicados ao “todo”, aplicando-os indevidamente às “partes”. Esta falácia mereológica, como se poderia denominar, é descrita da seguinte forma:

(FM) Falácia Mereológica: “O princípio de que os predicados psicológicos que só se aplicam aos seres humanos (ou a outros animais) como um todo não se podem aplicar inteligivelmente às suas partes” (Bennett & Hacker, 2003: 73).

Ora, os predicados psicológicos de (5), (6) e (7) não podem ser aplicados a apenas uma parte do ser humano (o seu cérebro), mas somente podem ser aplicados ao seu todo: à pessoa inteira. Não seguir esta regra será cometer a (FM), o que para os autores é visto como um erro que determina a qualidade epistémica do projeto científico do estudo do cérebro.

Curiosamente, (FM) é a aplicação de uma forma mais geral de verdade conceptual que Aristóteles apresentou no passado: “dizer que a alma está zangada é como se alguém dissesse que a alma tece ou constrói. Pois é certamente melhor não dizer que a alma tem pena, aprende ou pensa, mas que um homem faz isso com sua alma” (*apud* Hacker *et al.*, 2007: 131; Aristóteles, DA 408^b 2-15).

A segunda fonte de inspiração, mais direta, é um parágrafo de Wittgenstein:

(8) “Somente de um ser humano e do que se assemelha (se comporta como) a um ser humano vivo se pode dizer: tem sensação; vê, é cego; ouve, é surdo; é consciente ou inconsciente” (Wittgenstein, 1953: § 281).

Deixaremos a análise de (8) para depois. Primeiramente, iremos demonstrar como é que, segundo os autores, (FM) existe atualmente na Neurociência. De forma interessante, os autores não apresentam apenas a maleita conceptual, mas oferecem também o seu diagnóstico.

Segundo Bennett & Hacker, podemos aprender muito de uma área observando a sua história. Com o aparecimento da Neurociência surgiram paralelamente problemas conceptuais e lógicos – tal como ocorreu em muitas outras áreas científicas – que têm criado uma ciência com uma linguagem confusa e ambígua de elucidação: e é aí que, para os autores, o papel da Filosofia deve ser fundamental.

Charles Sherrington pode ser considerado como sendo parte da primeira geração de neurocientistas. O cientista ajudou a esclarecer a verdadeira natureza da espinal medula como centro de reflexo, tal como o papel do córtex na produção dos reflexos (cf. Bennett & Hacker, 2013: 231-236). O seu contributo científico é claro.

Apesar disso, defendeu um dualismo de espécie cartesiana: a mente e o cérebro são duas substâncias distintas. Nas palavras do autor: “[...] No entanto, a mente mostra-se esquiva à descrição da Física e da Química. Pertence a uma categoria que não pode ser sentida, ao insensível, onde o físico e o químico não têm lugar (Sherrington, 2009: 339)”. Para o neurocientista inglês, a categoria do mental ou da *psyché* imaterial está em oposição a todos os outros fenómenos da natureza, estando, assim, perante dois “fenómenos de duas categorias [diferentes]” (Sherrington, 2009: 318).

Segundo Bennett e Hacker, Sherrington não compreendeu que o problema da relação entre a mente e o corpo ou o cérebro não era um problema empírico, mas, antes, um problema conceptual (cf. Hacker & Bennett, 2003: 45). Assim, afirmou variadas ideias que, para os isolacionistas, não fazem qualquer sentido, tais como:

- “O pensamento do corpo parece ser feito para ele, nomeadamente por intermédio do cérebro” (Sherrington, 2009: 237).
- “Tenho visto posta a questão: ‘porque deveria a mente ter um corpo?’ A resposta pode bem ser: ‘para mediar entre ela e outras mentes’” (Sherrington, 2009: 261).

Ora, segundo Hacker, todas estas ideias são confusas: de facto, não é o corpo nem o cérebro que “possui” uma mente; antes, são as pessoas que têm mentes e corpos, seguindo (FM). O cérebro também não “tem” sensações, mas apenas nós, seres humanos (e alguns animais), temos sensações. Similarmente, o corpo não produz qualquer pensamento, nem o cérebro: apenas pessoas pensam, porque o cérebro não pensa em nada. O facto de apenas seres humanos que tenham um sistema nervoso funcional e saudável possam pensar em nada parece implicar que pensamos com o cérebro, tal como não caminhamos com as pernas (mas com o corpo inteiro) e não vemos com os olhos (sendo o olho apenas uma parte do processo de ver que se estende a muitas outras partes do corpo) (cf. Bennett & Hacker, 2003: 45).

Sherrington, ao abordar o problema mente-corpo através da lupa cartesiana, herdou também o problema de explicar como é que pode existir uma interação entre duas substâncias tão diferentes entre si. Afirmou que:

A ignorância do "como" da ligação entre o cérebro e a mente [...] não é culpa de quem estuda a mente nem de quem estuda o cérebro. Constitui uma deficiência comum a ambos. Uma ligação entre eles é o que cada um procura esclarecer. Que existe uma ligação [*sc. liasion*] nenhum deles duvida. O "como" de conceber tal permanece para a ciência como para a filosofia um enigma que precisa de ser equacionado. (Sherrington, 2009: 240)

Assim, ao herdar-se a concepção cartesiana (como veremos), herda-se o problema. Ora, o problema da relação mente-corpo não poderia colocar-se numa concepção aristotélica de mente ou *psyché*. É sabido que Aristóteles foi um dos primeiros pensadores que tentou detalhar o funcionamento do corpo humano e explicar de que modo o corpo poderia causar movimento. É com o pensamento aristotélico que surgem as primeiras indagações neurocientíficas rudimentares, complementadas por Galeno no início do primeiro milénio, e desenvolvidas por Nemésio, que se iria manter por mais de mil anos com Avicena (980-1037), Antonio Guainerio (1440) e Leonardo da Vinci (1452-1519), todos estes aceitando versões diferentes da doutrina ventricular (cf. Hacker e Bennett, 2003: 12-16).

Aristóteles concebia a alma, ou a *psyché*, como a relação das capacidades humanas e os diferentes órgãos que constituíam o ser humano. Para o Estagirita, a alma não é algo que se possui, algo extra que complementa o ser vivo. Pelo contrário, a alma, que será tripartida em nutritiva, racional e sensorial, abrange as funções essenciais do ser humano. A alma e o corpo são, na concepção aristotélica, algo que existe em paralelo: assim, pensar, sentir, perceber, memorizar, etc., são funções psicológicas

do ser humano no seu todo, e não algo que a alma ou o corpo faça individualmente (cf. Hacker e Bennett, 2003: 13-15).

Ora, a doutrina ventricular irá ser colocada em causa com a publicação, em 1542, da obra *De Naturali parte Medicinae* de Jean Fernel, onde o termo “fisiologia” apareceria pela primeira vez. Fernel, influenciado por São Tomás de Aquino, irá alterar a concepção aristotélica concebendo a alma racional do ser humano como algo que é separável e independente do corpo, sendo que o último poderia morrer, mas a alma seria imortal. Relembre-se que São Tomás tentaria conciliar duas perspetivas profundamente difíceis de harmonizar: por um lado, tentava compatibilizar os ensinamentos daquele que considerava ser “o Filósofo”, Aristóteles, com os princípios do Cristianismo.

Assim, Fernel, baseando-se em experiências e observações mais ou menos rigorosas, concluiu que alguns dos nossos atos são involuntários, sem qualquer influência da alma, como o movimento dos olhos ou da respiração. Mas estas observações eram lidas num contexto aristotélico de explicações teleológicas que deixariam de fazer sentido com os sucessos das explicações mecanicistas que aparecem com o advento da astronomia de Kepler e da física de Galileu (cf. Bennett & Hacker, 2003: 25).

Descartes, apoiado na ideia de que toda a fisiologia humana poderia ser explicada em termos mecânicos, iria alterar substancialmente a visão que iria receber da escolástica aristotélica em quatro pontos fundamentais:

- (i) A *psyché* ou mente que, em Aristóteles, era apenas uma parte da alma (tripartida) no seu todo, passa, com Descartes, a ser a alma na sua totalidade. As restantes partes da alma descritas pelo Estagirita deixam de ser, com o filósofo francês, uma função fundamental da mente, sendo essa função deslocada para o corpo, restando apenas a alma como princípio do Eu, do pensamento, da consciência, em oposição ao princípio da vida aristotélico;
- (ii) A mente escolástico-aristotélica é reconceptualizada por Descartes, deixando de ser somente focada nas funções de raciocínio e da vontade, mas passando a incluir as sensações, percepções, a imaginação e o apetite animal – todas estas funções farão parte do que é *pensar* (cf. Descartes, 1978: 60): a mente deixa de ser um mero intelecto, e passa a ser *res cogitans*;
- (iii) Descartes defendeu, ao contrário da visão unitária da escolástica – que via o ser humano como um ser uno – que o ser humano é uma entidade composta por duas substâncias fundamentalmente diferentes, embora relacionadas intimamente;
- (iv) Finalmente, o cartesianismo irá assinalar a visão dualista do mundo: a matéria teria como propriedade fundamental a extensão, sendo que os processos físicos e biológicos seriam todos

explicados mecanicamente; os seres humanos, contudo, teriam uma outra propriedade fundamental: a mente passaria a ser pensamento, no sentido cartesiano ampliado indicado em (ii) (cf. Bennett & Hacker, 2003: 25-27).

Ora, é este o quadro conceptual que Sherrington irá receber e aplicar na sua investigação empírica. Curiosamente, fez notar que a ciência era impotente para resolver o problema mente-corpo:

Mas entre energia⁶ e mente a ciência não encontrou nenhum ‘como’ de toma-lá-dá-cá [...] O mundo permaneceu obstinadamente duplo à compreensão do homem [...] De facto, a ciência natural repudia-a como sendo algo fora do seu alcance (Sherrington, 2009: 291).

Neste sentido, o Nobel da Medicina de 1932 vai continuar uma longa linhagem de pensamento que identifica no reducionismo uma tarefa filosófica impossível de alcançar: “Os dois [*sc.* o cérebro e a mente] (...) permanecem teimosamente separados. Parecem-me diferentes; não convertíveis mutuamente; intraduzíveis um no outro” (Sherrington, 2009: 312).

Embora alterando o papel da glândula pineal cartesiana, que serviria como ponto central de interação psicofísica, Sherrington, segundo Hacker, não percebeu que o fundamento do problema mente-corpo não se devia à ignorância empírica ou científica, mas a uma confusão conceptual intrínseca aos fundamentos teóricos da visão da mente e do corpo dualista (cf. Hacker & Bennett, 2003: 47).

Um discípulo de Sherrington, Wilder Penfield, que pode ser considerado como pertencente à segunda geração de neurocientistas, disse o seguinte sobre o problema:

[...] após muitos anos de empenho para explicar a mente com base apenas na ação cerebral, cheguei à conclusão que é mais simples (e muito mais lógico) adotar a hipótese de que o nosso ser consiste em dois elementos fundamentais [...] porque me parece ser impossível explicar a mente com base na ação neuronal do cérebro [...] sou forçado a escolher a proposição de que o nosso ser deve ser explicado com base em dois elementos fundamentais (Penfield, 1975: 80).

⁶ Sherrington usa o termo “energia” para designar a matéria em conjugação com a energia no sentido físico e material que está em oposição à mente imaterial (cf. Sherrington, 2005: 251).

Embora tivesse tentado explicar a mente apelando somente a mecanismos do cérebro, Penfield viria a abraçar o dualismo herdado do seu mestre.

Tal como Sherrington, Penfield baseava a sua crença em provas científicas rudimentares cujos resultados eram erroneamente interpretados segundo uma abordagem conceptual cartesiana. Os fenómenos que mais impressionaram o neurocientista canadiano eram os seguintes: (a) o automatismo epilético; e (b) a estimulação cortical por elétrodos. Em relação a (a), quando um paciente entra em estado epilético irá entrar num estado automático de ação, não registando conscientemente qualquer movimento, pelo que estará numa condição de ausência temporal sem memória do que aconteceu durante o ataque:

Durante a interferência epilética com a função desta matéria cinzenta [...] a consciência extingue-se, e com ela desaparece a direção e o planeamento do comportamento. Ou seja, a mente deixa de estar em ação e entra em ação com o funcionamento normal deste mecanismo (Penfield, 1975: 47).

Assim, para Penfield, (a) causaria um “desligar da mente”, dado que, durante um ataque epilético, o cérebro passaria a controlar a mente que antes teria o controlo.

Em relação a (b), utilizando técnicas de estimulação elétrica em zonas específicas do cérebro, Penfield notou que poderia influenciar o comportamento e ação dos pacientes com essas estimulações, fazendo com que, por exemplo, um paciente não conseguisse descrever uma imagem (e.g. de uma borboleta) dada a interferência elétrica em causa ou levantasse um braço sem conscientemente dar-se conta desse movimento. Concluiria destes factos que o papel da mente seria selecionar os conceitos que, depois, iriam influenciar os mecanismos superiores do cérebro que, por sua vez, iriam influenciar causalmente o mecanismo da fala (cf. Penfield, 1975: 52).

Deste modo, vai concluir que é a mente que “raciocina, compreende e decide” (Penfield, 1975: 75-77), sendo que o mecanismo cerebral superior seria o lugar onde o mental encontraria o cerebral (cf. Penfield, 1975: 63).

Ora, mais uma vez, Bennett e Hacker apontam os erros de Penfield. Em primeiro lugar, este aceitou a visão do seu mestre Sherrington de que a mente era uma substância independente, o que chamaria de “elemento fundamental”, tendo identificado o Eu na mente, atribuindo a este as capacidades psicológicas humanas e o papel de agente causal que influencia as ações do corpo (cf. Bennett & Hacker, 2003: 62-64). Mas como já fizemos observar, para os autores, a mente não é uma

substância, mas as capacidades psicológicas dos seres humanos. Todavia, o sujeito dessas capacidades não é a mente, nem sequer o cérebro, mas a pessoa no seu todo: somos nós que decidimos, pensamos, compreendemos, e não a nossa mente, ou partes do nosso cérebro (cf. Bennett & Hacker, 2003: 63).

Em segundo lugar, Penfield abraçou a pressuposição metodológica de que é uma questão empírica saber se o cérebro é responsável pela mente ou se são duas substâncias diferentes, sendo estas duas hipóteses de natureza científica a necessitarem de esclarecimento através de dados empíricos subsequentes. Ora, seguindo a mesma linha de crítica, nem o cérebro, nem a mente são sujeitos das capacidades ou funções psicológicas, sendo que não precisamos de aceitar o que os autores chamam de “dilema de Penfield”: no fundo, Penfield estava certo que não se poderia explicar o comportamento, por um lado, e a consciência, por outro, apelando somente à atividade cerebral. Contudo, Penfield estava enganado ao defender que tal se devia à falta de uma hipótese empírica por descobrir, e não a uma mera confusão conceptual, como os autores defendem (cf. Bennett & Hacker, 2003: 51).

Portanto, se aceitarmos estas considerações conceptuais, nem (a) nem (b) funcionam como prova empírica da hipótese dualista: o que acontece, apenas, é que (a) e (b) provam que esses fenómenos diminuem algumas capacidades psicológicas humanas temporariamente.

O relevante desta breve análise da primeira e segunda geração de neurocientistas é, fundamentalmente, a afirmação do cartesianismo tradicional e, por consequência, a continuação das obscuridades conceptuais.

Chegamos, finalmente, à terceira geração, a atual: os neurocientistas presumem que o cérebro tem uma vasta gama de capacidades cognitivas, percetivas e volitivas. Por exemplo, como vimos, Francis Crick afirma: “o que vê não é o que está realmente aí; é o que o seu cérebro acredita que está aí” (Crick, 1995: 31) – de forma que o cérebro tem experiências, acredita em coisas, interpreta pistas com base em informação que foi posta à sua disposição, faz conjecturas, etc. (cf. Bennett & Hacker, 2003: 68).

Já Colin Blakemore afirma que os “os neurónios [...] têm conhecimento [*sc.* e são] inteligentes dado que são capazes de estimar a probabilidade de eventos externos. [...] Os neurónios apresentam argumentos ao cérebro baseados em características específicas que detetam, argumentos esses em que o cérebro se baseia para construir as hipóteses da sua perceção” (Blakemore, 1997: 91) – dado que os neurónios são sujeitos lógicos do conhecimento, são inteligentes, apresentam argumentos, etc.

António Damásio também afirma algo neste sentido: “os nossos cérebros podem muitas vezes decidir corretamente” (Damásio, 1996: 173). Estas e outras afirmações são comuns na Neurociência contemporânea, como indicado por Bennett e Hacker ao longo da sua obra *Fundamentos Filosóficos da Neurociência*.

Surge agora, para os autores, o trabalho filosófico: tudo isto são atributos dos seres humanos. Porém, será uma nova descoberta que o nosso cérebro se envolva neste tipo de atividades? Ou será uma novidade linguística introduzida por razões teóricas apresentada pelos neurocientistas, ampliando o uso comum das expressões? Ou será apenas uma confusão conceptual baseada em pressupostos teóricos equivocados (cf. Bennett & Hacker, 2003: 71)?

Por que razão é que tão reputados cientistas usam estas expressões sem qualquer reflexão sobre elas? A resposta está, segundo os autores, na história: como vimos, a primeira e segunda geração de neurocientistas afirmaram o dualismo de substâncias, que aplicava os atributos psicológicos à parte mental ou imaterial, mas a terceira geração negou o dualismo, afirmando o monismo fisicalista, mas desviando a aplicação desses atributos da mente – que deixa de existir como sendo imaterial e invisível – para o cérebro. Assim se explica, segundo Bennett e Hacker, porque é que (FM) é tão comum na Neurociência atual.

É por isso que, segundo os autores, é muitíssimo importante saber distinguir entre questões conceptuais ou *a priori* e questões empíricas ou *a posteriori*. confundir ambas é criar um nó conceptual obscuro que nenhuma investigação empírica, por mais dispendiosa ou reputada que seja, poderá resolver.

Mas que pressupostos são utilizados pelos autores para sustentar a defesa de (FM) e as consequências que tal implica no trabalho da Filosofia e da sua relação com a Neurociência?

Os autores irão apresentar a sua posição isolacionista nas seguintes afirmações:

(9) “As questões conceptuais antecedem questões de verdade e falsidade. [...] Elas determinam não o que é empiricamente verdadeiro ou falso, mas o que faz e o que não faz sentido. Portanto, questões conceptuais não são passíveis de investigação e experimentação científica [...] Distinguir questões conceptuais de questões empíricas é de enorme importância” (Bennett & Hacker, 2003: 2);

(10) “Não achamos que a investigação empírica possa resolver problemas filosóficos, assim como não pode resolver problemas em matemática” (Bennett & Hacker, 2003: 414).

As afirmações (9) e (10) ilustram perfeitamente a posição dos autores quanto ao problema metodológico. É proposto que as duas disciplinas estão metodologicamente separadas. Deste ponto de vista, verdades conceptuais não podem ser informadas por verdades empíricas, e vice-versa.

Segundo os autores, a separação metodológica é justificada pelo objeto de estudo de cada disciplina: segundo os autores, a Filosofia trabalha o mundo lógico das possibilidades e o sentido das palavras. Já a (Neuro)ciência trabalha o mundo natural e a verdade. Curiosamente, esta é uma forma tradicional de definir o trabalho filosófico pelos filósofos, e muitos cientistas confirmam esta concepção, embora o façam com algum desdém.

Assim, (9) e (10) pressupõem que os conceitos são *a priori* de qualquer investigação científica, e que não tomar isto em consideração levará necessariamente a erros do tipo demonstrado em (FM). Nesta concepção isolacionista, o trabalho do filósofo não é apontar se tal e tal experiência nos dá conhecimento verdadeiro ou falso. Antes, analisando a sua linguagem, cabe-lhe perceber se a sua estrutura lógica faz ou não sentido.

II.1.2. A Influência de Wittgenstein

Wittgenstein afirmou algo que poderá ser visto como semelhante a esta abordagem metodológica. Podemos afirmar até que Bennett e Hacker estão a aplicar algumas ideias de Wittgenstein ao problema da relação entre a Filosofia e a Neurociência. Mas o problema, como tentaremos argumentar, é que os autores estão a ser influenciados por uma ideia presente no trabalho de Wittgenstein, nomeadamente (8) (ver acima, secção II.1.1, pág. 33).

Esta é a interpretação que os autores fazem das palavras de Wittgenstein: a manifestação comportamental funciona como critério lógico para a aplicação dos conceitos mentais. Somente de um ser que se comporta de certa forma podemos dizer que possui certo conceito mental.

Quando observamos uma criança a chorar com uma cara triste, saberemos que ela está triste devido a algo. E apenas este tipo⁷ de comportamento nos pode oferecer conhecimento sobre os estados mentais das outras pessoas. Assim, é fácil perceber porque é que o cérebro não pensa ou percebe algo: o cérebro não se parece como um ser que pensa ou percebe algo – um cérebro não pode exibir o comportamento apropriado de pensar, isto é, as condições necessárias para que a proposição seja conceptualmente sólida e faça sentido.

Influenciados por esta ideia, Bennett e Hacker pensam que falar desta maneira sobre o cérebro é um erro particular: uma confusão. Não é que seja errado empiricamente que o cérebro pense ou

⁷Não este especificamente, mas todos os comportamentos ou disposições associadas ao sentimento de tristeza.

decida. Mas é, antes, uma confusão conceptual: o que é conceptualmente verdadeiro é que a percepção, os sentimentos, etc., são atributos de seres humanos (ou semelhantes) e não das suas partes – não dos seus cérebros.

Assim, a grande contribuição dos autores para o debate é afirmar que a ciência pode estar equivocada conceptualmente como qualquer outro empreendimento intelectual humano. Veremos já de seguida um exemplo que demonstra de que forma a própria história da ciência se encontra cheia de confusões conceptuais que não estavam apenas empiricamente equivocadas, mas conceptualmente incorretas.

A Neurociência, como ciência, poderá sofrer dos mesmos erros que os autores tentam demonstrar na sua muita conhecida obra. Em alguns casos, esses erros conceptuais podem não afetar decisivamente a investigação empírica *per se*, mas podem, de facto, pôr em causa a interpretação e compreensão dos resultados dessas investigações empíricas.

Como afirmou Heinrich Hertz na introdução à sua obra *Princípios de Mecânica*: “Quando estas dolorosas contradições [*sc.* conceptuais] são eliminadas [...] as nossas mentes, deixando de ser contrariadas, também deixarão de colocar perguntas ilegítimas” (*apud* Bennett & Hacker, 2003: 5). É neste espírito de “limpeza conceptual” que estes autores⁸ pretendem inserir-se focando-se nas investigações mais recentes da Neurociência contemporânea.

II.1.3. Exemplo de Investigação Isolacionista: o Determinismo

Vejamos, como exemplo desta tese geral, o problema de determinismo versus livre-arbitrio examinado pelo neurocientista Benjamim Libet. Segundo este neurocientista, todas as nossas ações voluntárias são, ao contrário do que poderíamos conceber, influenciadas pelo nosso cérebro numa primeira fase, pelo que a crença de que temos controlo sobre as nossas ações e que estas são livres no sentido de não serem influenciadas por fatores externos ao sujeito revelar-se-ia falsa. Segundo Libet, esta não é uma simples tese filosófica declarada *a priori*, mas antes, pelo contrário, trata-se de uma conclusão que é demonstrada pela investigação empírica em Neurociência.

⁸ Este projeto aproxima-se da investigação avançada por Gilbert Ryle no seu *The Concept of Mind*. A grande novidade é uma tentativa de defender detalhadamente uma rede lógica conceptual que estava em falta nesse projeto. Ademais, embora Ryle siga Wittgenstein em muitos aspetos, difere do mesmo em muitos outros, como por exemplo o valor das expressões de categoria. Seguindo a explicação de Hacker, Wittgenstein via essas expressões como indeterminadas. Pelo contrário, Ryle assumia-as como sólidas, cujo problema era apenas serem aplicadas de forma errónea (cf. Hacker, 2019: 201).

A grande descoberta de Libet foi ter identificado que o cérebro “decide” um ato particular num determinado tempo *antes* do sujeito decidir esse mesmo ato conscientemente. Tal é empiricamente demonstrável com a identificação de neurónios específicos presentes no córtex motor suplementar que se excitam 500ms *antes* dos impulsos chegarem, por exemplo, aos músculos da mão (cf. Libet, 1993a: 270).

Ora, essa identificação não é ainda suficiente para colocar em causa, supostamente, a tese do livre-arbitrio. O que é necessário mostrar é que há uma diferença entre o momento de excitação neuronal (i.e. 500ms) e o tempo de decisão subjetiva humana. Foi exatamente isso que Libet tentou demonstrar: segundo a sua investigação, o ato de querer ou o impulso para realizar o movimento da mão foi descrito pelos sujeitos como acontecendo somente 150ms *antes* de a ação ser realmente efetuada.

Assim, fazendo as contas, o que Libet concluiu é que um ato supostamente livre e voluntário de um sujeito é iniciado, ao contrário da percepção subjetiva do mesmo, cerca de 350ms antes do impulso do movimento (cf. Libet, 1993b: 398 e seg.). Para Libet, um ato é voluntário quando os sujeitos sentem introspectivamente que a ação que estão a realizar naquele preciso momento acontece por sua iniciativa, desejo e intenção.

Contudo, segundo Bennett e Hacker, estes resultados empíricos baseiam-se em confusões conceptuais: não é condição nem necessária nem suficiente que algum sentimento de desejo ou intenção de realizar um movimento tenha de anteceder uma ação voluntária para acontecer (cf. Bennett & Hacker, 2003: 229).

Segundo os autores, há vários erros conceptuais cometidos na abordagem de Libet ao livre-arbitrio:

- (i) Quando x provoca uma ação voluntária y , x não tem qualquer propriedade antecedente que Libet acredita que x possui, seja um impulso, uma intenção ou um desejo. Diferentemente será a ideia de que podemos, de facto, *dizer* se tal e tal y ou $\neg y$, mas esse ato nunca é influenciado por algo sentido anteriormente por x de modo a provocar y ;
- (ii) Os impulsos, intenções ou desejos não são suficientes, como Libet assume, para que x cause uma ação voluntária y , pelo contrário. Por exemplo: x pode sentir o impulso de bocejar e, logo depois, bocejar efetivamente. Ora, esta precedência volitiva não torna o ato de bocejar uma ação voluntária;
- (iii) Contra o que pressupõe Libet, o facto de x causar uma ação voluntária y por causa de um impulso prévio que sentiu não torna essa ação voluntária à partida. Por exemplo: se um impulso

de bocejar afeta x , este impulso é a causa de x bocejar, e todo este processo é somente involuntário, mesmo que baseado em algo que x sentiu antes de efetivar uma ação;

(iv) No fundo, Libet faz uso de uma equivocada teoria da natureza da ação: não é qualquer impulso ou desejo prévio a x que torna uma ação voluntária ou não, nem a ausência total desse desejo prévio ou impulso o tornaria involuntário. Como afirmam os autores: “(...) um movimento voluntário não é um movimento causado por um impulso sentido, tal como abster-se voluntariamente de mexer não é sentir um impulso para *não* mexer que nos *impede* de mexer” (cf. Bennett & Hacker, 2003: 208-209).

Mas como é que um reputado cientista como Libet se deixa levar por concepções teóricas errôneas sobre atos voluntários e involuntários? Além de identificarem os erros, os autores identificam também a razão por que neurocientistas como Libet cometem este tipo de equívocos teóricos.

Esta influência conceptual é traçada naquilo a que apelidam de “criptocartesianismo” (cf. Bennett & Hacker, 2003: 233). De acordo com Descartes, as volições seriam ações da alma imaterial, sendo que haveria uma classe particular de volições que teriam impacto no corpo material através da famosa glândula pineal. Veja-se as palavras do filósofo francês:

A atividade da alma consiste inteiramente no facto de que, simplesmente querendo uma coisa, faz com que a pequena glândula pineal à qual está estreitamente ligada se mova da maneira requerida para produzir o efeito correspondente a esta volição (Descartes, 1964: I-41).

Ora, segundo Descartes, se queremos dar um passo para iniciarmos uma caminhada, por exemplo, o que vai acontecer é que esse impulso (volição) vai ter um impacto causal na glândula pineal que irá influenciar os músculos específicos para o movimento corporal pretendido (cf. Bennett & Hacker, 2003: 212).

Como os autores fizeram notar, Libet concebe a mesma estrutura de ação voluntária: um impulso (volição) anterior causa o movimento corporal consequente. O que este investigador descobriu, afinal, foi que a volição em si mesma é um ato do cérebro, que acontece inconscientemente antes de o sujeito se dar conta, conscientemente, de qualquer volição para, por exemplo, iniciar uma caminhada e começar a caminhar (cf. Nachev & Hacker, 2014: s/p).

Assim, o que Libet faz é aplicar o (erróneo) esquema conceptual cartesiano na sua totalidade, apenas efetuando uma alteração particular: atribuir ao cérebro o que Descartes atribuía à alma, isto é, conceber que os impulsos, desejos e intenções – as volições – são atos do cérebro, em vez da alma cartesiana (cf. Bennett & Hacker, 2003: 235).

A investigação de Libet antecede uma investigação empírica que pressupõe o mesmo tipo de erro. John Eccles, um dualista assumido e também Nobel da Medicina em 1963, já teria concluído algo semelhante através da sua investigação empírica. Eccles viria a seguir a conceção dualista de Sherrington, acreditando que esta seria comprovada por duas linhas de investigação particulares: (a) pelo trabalho realizado por Roger W. Sperry com doentes com a divisão do cérebro em dois hemisférios através da comissurotomia, cada um com funções específicas (cf. Sperry, 1961) e (b) o trabalho realizado por Hans Helmut Kornhuber sobre a relação entre uma ação intencional e o potencial elétrico produzido especificamente no córtex cerebral (cf. Kornhuber, 1984).

Por exemplo, em (a), Eccles observa que as investigações de Sperry, que relacionavam as capacidades ou funções com cada hemisfério, sendo que apenas uma parte específica do cérebro servia de ponto de contacto com a mente consciente, provavam um dualismo mente-cérebro:

[...] só uma zona especializada dos hemisférios centrais está em ligação com a mente autoconsciente. O termo cérebro de ligação denota todas essas áreas do córtex cerebral que são potencialmente capazes de estar em ligação direta com a mente autoconsciente (Popper & Eccles, 1977: 358).

Contudo, a divisão de funções por hemisférios avançada por Sperry e aceite por Eccles é somente uma confusão teórica: nem as atividades neuronais no hemisfério direito “são desconhecidas do sujeito falante”, não podendo “se exprimir na linguagem”; nem o hemisfério esquerdo é predominante devido às suas aptidões verbais e ideativas e a sua ligação à autoconsciência (cf. Eccles, 1979: 220), porque, segundo Bennett e Hacker, nenhum hemisfério se pode “exprimir na linguagem”, tal como o braço esquerdo não se pode exprimir na linguagem, dado que nenhuma parte particular do ser humano se pode “exprimir na linguagem” (cf. Bennett & Hacker, 2003: 56).

No mesmo sentido, “os hemisférios centrais” não estão em contacto com a mente consciente porque nenhuma parte do ser humano, quanto mais do cérebro humano, se relaciona com a consciência, mas somente o ser humano por inteiro como vimos em (FM).

Já em relação a (b), Kornhuber investigou as alterações do potencial elétrico que antecederiam uma ação voluntária: descobriu que o chamado potencial de prontidão iniciava-se 800 milissegundos antes do início do potencial de ação do músculo, isto é, do próprio movimento (cf. Eccles, 1979: 217).

A interpretação de Eccles dos resultados empíricos observados foi a seguinte: os dados provam que a mente interage causalmente não na glândula pineal cartesiana, mas nas células piramidais do córtex motor, e que essa interação não é instantânea como assumia Descartes, mas gradual – daí a diferença de tempo entre o ato voluntário e o próprio movimento ocorrido.

Ora, o erro de Eccles é o mesmo de Libet, que se basearam numa conceção cartesiana de movimento voluntário cuja ação seria causada por atos prévios de vontade. Mas é duvidoso, segundo os isolacionistas, que a maior parte das nossas ações seja antecida por atos de vontade; e mesmo que tal fosse o caso, essa informação em nada nos informaria sobre se essa ação seria voluntária ou involuntária (cf. Bennett & Hacker, 2003: 53).

Como vemos neste exemplo sobre livre-arbítrio, de facto, uma ciência conceptualmente pobre pode coexistir com uma ciência que floresce e acrescenta conhecimento ao mundo (cf. Bennett & Hacker, 2003: 5).

Mas por que razão deve a Filosofia reivindicar a correção da investigação em Neurociência? Por que é que a Filosofia deve julgar a clareza e coerência de conceitos usados por neurocientistas reputados? Como é que a Filosofia pode afirmar que muitas afirmações feitas por esses mesmos neurocientistas não têm qualquer sentido?

A resposta a estas questões será então a defesa da abordagem isolacionista. Para os autores, aquilo que é verdadeiro ou falso para a Ciência é o sentido (ou a sua falta) para a Filosofia. Assim, o erro empírico resulta em falsidade. Mas o erro conceptual resulta em falta de sentido – sentido será, como iremos ver, o carácter lógico-gramatical dos conceitos.

Quem deve investigar os limites do sentido? Será tal uma tarefa a ser executada pela Neurociência? Ou tal deve ser feita pela Filosofia? Segundos os autores, a resposta é óbvia: embora a Filosofia tenha vários empreendimentos, a principal que tem vindo a definir uma forma específica de fazer Filosofia é exatamente a análise do uso das palavras, isto é, a análise conceptual.

Um problema conceptual será então quando uma expressão ou conceito é aplicado contrariamente às regras do seu uso (cf. Bennett & Hacker, 2003: 6). Mas que regras são estas a que os autores se referem?

Segundo os autores, há dois tipos fundamentais de conceitos: (a) os conceitos comuns, não técnicos, cujas regras são extraídas do uso normalizado dos mesmos e das descrições dos seus

significados; (b) os conceitos técnicos e especializados, cujas regras são extraídas pela introdução do conceito pelo teórico e das explicações que este apresenta para o seu uso (cf. Bennett & Hacker, 2003: 7).

Quando os conceitos são usados contra as regras do seu uso, existe o problema conceptual ou, nas palavras dos autores, o “contrassenso”: uma formulação específica que não é estipulada por qualquer regra de uso e que, por isso, não faz sentido e deve ser excluída da linguagem.

Uma segunda forma de contrassenso é o uso de uma expressão ou conceito com uma definição própria dando origem a uma nova definição que contradiz a definição original. Segundo os autores, o trabalho da Filosofia, isto é, do filósofo, é identificar essas “transgressões dos limites do sentido” (Bennett & Hacker, 2003: 6).

O pressuposto metafilosófico que está a ser intuído é definir o trabalho filosófico como “crítica conceptual”. A que se reduz esse trabalho, especificamente? A Filosofia, nesta posição isolacionista, define-se por um empreendimento intelectual cuja função é mostrar que, por exemplo, o neurocientista usou um conceito que vai contra o seu uso legítimo ou que, fazendo uso do seu sentido tradicional lhe assumiu um novo significado (cf. Bennett & Hacker, 2003: 6).

Mas então, em que ponto a Filosofia se poderá relacionar com a Neurociência? Os autores, afirmando claramente o seu isolacionismo, não têm qualquer dúvida quanto ao papel da Filosofia na investigação empírica:

[...] as únicas maneiras pelas quais uma investigação conceptual [*sc.* filosofia] consegue auxiliar um tema empírico [*sc.* neurociência] é identificar o erro conceptual e proporcionar um mapa para evitar que os investigadores empíricos [*sc.* neurocientistas] vagueiem fora das principais estradas do sentido (Bennett & Hacker, 2003: 7)

Repare-se que estamos perante uma metafilosofia prescritiva (ou normativa) clara e que poderá, como iremos propor mais à frente, estar equivocada.

Ora, será que esta posição metafilosófica isolacionista atribui à Filosofia um papel meramente negativo ou deflacionário, isto é, a de identificar problemas conceptuais e falácias das investigações contemporâneas em Neurociência? Segundo os autores, esta é, de facto, uma das tarefas. Mas o trabalho filosófico não se reduz meramente a esta empreitada deflacionária: a tarefa de representar claramente o mapa conceptual de cada um dos conceitos e expressões usadas pelos neurocientistas é uma tarefa construtiva que deve ser fornecida pelo filósofo e pela sua investigação particular.

Diferentemente do que acabamos de afirmar, é avançar-se com a ideia de muitos neurocientistas contemporâneos de que “a filosofia é inútil para a ciência e que nada contribui para a mesma”.⁹ Ora, segundo a abordagem isolacionista, esta crítica não faz qualquer sentido: exigir que a investigação conceptual *a priori* proponha hipóteses empíricas que possam resolver os problemas empíricos dos cientistas é uma confusão conceptual enorme que confunde o trabalho filosófico com o trabalho científico. Nas palavras dos autores:

[...] não pode ser a tarefa da investigação conceptual propor hipóteses empíricas que consigam resolver os problemas empíricos enfrentados pelos cientistas. Lamentar-se que a investigação filosófica na neurociência não tenha contribuído com uma nova teoria neurocientífica é como lamentar-se a um matemático que um novo teorema que ele demonstrou não seja uma nova teoria física (Bennett & Hacker, 2003: 7).

E novamente percebemos claramente a posição metafilosófica da abordagem isolacionista: a Filosofia tem uma tarefa particular, a Neurociência tem outra, e ambas não se confundem (cf. (4)).

Mas qual é essa tarefa, exatamente? Para os isolacionistas, e seguindo uma interpretação muito particular de Wittgenstein, defender-se-á a tese metafilosófica que a Filosofia deve ser vista como uma ferramenta linguístico-terapêutica cuja função é analisar a forma como as palavras e os conceitos são usados numa comunidade linguística particular.

Nesta conceção de Filosofia, o papel da mesma não é oferecer soluções aos problemas filosóficos, mas, antes, dissolvê-los através da análise do uso de determinados conceitos e da forma como estão relacionados com outros conceitos dado que os problemas filosóficos são, na sua maioria, problemas causados pelo uso errôneo de determinados conceitos.

Desta forma, seguindo esta metafilosofia específica, a investigação filosófica é profundamente dissemelhante da investigação científica, sendo que há uma divisão clara entre ambas, como é afirmado por Hacker: “A filosofia não contribui para o conhecimento humano, mas para a compreensão humana” (Hacker, 2001: 139).

Nesta conceção, por mais factos empíricos que tenhamos sobre um problema filosófico específico, tal informação não nos ajudará minimamente a resolver esses problemas, dado que os

⁹ Por exemplo, Stephen Hawking afirma no seu *The Grand Design* que “[...] a filosofia está morta” (Hawking, 2010: 5).

problemas da Filosofia são problemas conceptuais por definição: requerem uma elucidação teórica dos conceitos, e não evidências empíricas dos conceitos problemáticos à partida.

Ora, não será espantoso que esta conceção tenha levantado várias críticas por parte dos próprios neurocientistas, que podem ser reduzidas a quatro tipos:

- (i) A Filosofia é irrelevante para os interesses da Neurociência e dos neurocientistas;
- (ii) Os métodos da Filosofia *a priori* são inúteis para investigar a natureza da mente;
- (iii) Os temas e problemas da Filosofia são os mesmos que os da Neurociência, pelo que o problema da mente consciente não é exclusivo da Filosofia, sendo que esta já não serve para explorar esse tema;
- (iv) A realização dos filósofos durante 2000 anos é insignificante em relação ao problema da mente ou da consciência (cf. Bennett & Hacker, 2003: 397).

Todas estas críticas são realmente defendidas por neurocientistas contemporâneos: em relação a (i), Ian Glynn afirma “A maior parte dos meus colegas cientistas e médicos tende a depreciar a Filosofia. É evidente que a sua atitude reflete uma avaliação realista da provável importância da Filosofia para as suas preocupações imediatas” (Glynn, 1999: 367). Em relação a (ii), Gerald Edelman declara: “Uma das tentações de ter uma mente é tentar usá-la isoladamente para resolver o mistério da sua própria natureza. Os filósofos tentaram fazer isto desde tempos imemoriais [...] como método geral para explorar a mente, simplesmente não serve (Edelman, 1994: 31).

Já em relação a (iii), Edelman e Tononi afirmam “sugerimos que a consciência possa ser considerada um tema científico e não somente um domínio específico da Filosofia” (Edelman & Tononi, 2000: 3). Finalmente, em relação a (iv), Francis Crick afirma que “os filósofos tiveram um registo tão insignificante nos últimos dois mil anos que fariam melhor em mostrar uma certa modéstia em vez da pomposa superioridade que revelam habitualmente” (Crick, 1995: 258).

Todavia, farão sentido estas acusações? Hacker, defendendo então uma metafilosofia específica, irá mostrar que não, porque as mesmas se encontram equivocadas. Por um lado, a Filosofia tem a tarefa construtiva de investigar e descrever os limites do sentido, pelo que não se interessa diretamente por questões de facto ou empíricas, mas por questões de sentido. Por outro, tem a tarefa deflacionária de crítica das “transgressões dos limites do sentido” quer na formulação dos problemas quer nas soluções propostas a esses problemas.

Note-se que tanto os problemas filosóficos *a priori* quanto os problemas científicos *a posteriori*, assim como as soluções filosóficas ou as descobertas empíricas podem transgredir os limites do sentido e essa é a tarefa do investigador de Filosofia (cf. Bennett & Hacker, 2003: 399).

Nesta posição isolacionista, seguindo uma tradição de conceber a Filosofia como análise da linguagem (cf. Capítulo 1, I.2), transmitimos conhecimento do mundo através da nossa linguagem composta de proposições que são, por sua vez, compostas de palavras que podem ser combinadas de formas específicas. São estas combinações que podem ser lícitas ou ilícitas, que o filósofo deve analisar, quer seja feita esta análise sobre problemas ou teorias *a posteriori*, ou sobre problemas filosóficos tradicionais.

A tarefa da Filosofia, na abordagem isolacionista, será então o exame e descrição cuidadosa da forma como determinadas palavras são usadas em determinado contexto. Nas palavras de Bennett e Hacker:

[sc. os conceitos e palavras] são elucidados pela análise conectiva, que reconstitui, na medida em que é necessário para o esclarecimento e a solução ou dissolução dos problemas e dúvidas em questão, as ramificações da teia lógico-gramatical de conexões entre o conceito problemático e os conceitos adjacentes (Bennett & Hacker, 2003: 400).

Consequentemente, as acusações (i)-(iv) não fazem sentido para um isolacionista, porque não é tarefa da Filosofia oferecer hipóteses e teorias empíricas, mas analisar as possibilidades lógicas dos conceitos. Assim, o desdém dos neurocientistas pela Filosofia como uma investigação de “simples palavras” seria como desdenhar os óculos que alguns de nós usam para ver o mundo apenas porque são compostos de vidro (cf. Bennett & Hacker, 2003: 401).

A Filosofia será, antes, a procura da elucidação dos conceitos pela análise da aplicação dos mesmos aquando do seu uso, sendo que esses conceitos antecedem a experiência e pressupõem qualquer teoria ou experimento científico.

Seguindo a abordagem isolacionista, declaremos o que significa para a Filosofia, de um lado, e para a Neurociência, do outro, estudar a natureza da mente (cf. Bennett & Hacker, 2003: 402-403):

- Filosofia: a investigação filosófica sobre a mente procura esclarecer o conceito “mente” e as ramificações das suas conexões lógico-gramaticais com conceitos relacionados (e.g. pessoa, consciência, cérebro, corpo, etc.);

- Neurociência: a investigação empírica sobre a mente procura descobrir quais são as bases neuronais de determinadas capacidades ou funções psicológicas (e.g. percepção, memória, consciência, etc.).

Assim, além de concluir que “os empreendimentos filosófico e neurocientífico são muito diferentes” (Bennett & Hacker, 2003: 403), os autores ainda afirmam que as investigações do segundo empreendimento, a Neurociência, pressupõe as investigações do primeiro empreendimento, da Filosofia:

[...] na medida em que a falta de clareza ou a confusão em relação ao conceito de mente e aos conceitos psicológicos com ele relacionados vão contaminar a descrição e a compreensão dos problemas e soluções neurocientíficas [...] a filosofia não pode dar qualquer contributo às teorias científicas acerca da base neural da consciência [...] tal como a neurociência não pode dar qualquer contributo para o esclarecimento do conceito de consciência (Bennett & Hacker, 2003: 403).

Assim, para a abordagem isolacionista, nenhum argumento filosófico pode resolver qualquer problema experimental, seja este qual for.

II.1.4. Várias Defesas da Abordagem Isolacionista

Ora, mas será que (4) é uma tese exclusivamente defendida por Bennett e Hacker? A resposta é claramente negativa.¹⁰ Existe uma multiplicidade de filósofos que define Filosofia como uma investigação dos limites do sentido. Na introdução do seu *Linguistic Turn*, Richard Rorty afirmou:

O objetivo do presente volume é fornecer materiais para reflexão sobre a mais recente revolução filosófica, a da filosofia linguística. Por “filosofia linguística”, entenderei que os problemas filosóficos são problemas que podem ser resolvidos (ou dissolvidos) pela reforma

¹⁰Uma defesa da aprioridade metodológica da filosofia pode ser oferecida pelos defensores contemporâneos do Tomismo, uma abordagem filosófica baseada no pensamento de São Tomás de Aquino cujo objetivo principal era conciliar o pensamento aristotélico, versões do Neoplatonismo e os textos sagrados da Bíblia (cf. Kerr, 2002; Cessario, 2003): o Tomismo argumenta a favor da autonomia da reflexão filosófica e pela ideia de que nenhum dado empírico pode falsificar ou ser relevante para rejeitar uma teoria filosófica dado que se encontram em dois planos de reflexão totalmente separados (cf. Brozek, 2013: 17).

da linguagem ou pela compreensão [sc. elucidação] da forma como usamos a nossa linguagem (Rorty, 1967: 3).

A. J. Ayer, uns anos antes, declara algo semelhante:

[sc. O] filósofo, como analista, não está diretamente preocupado com as propriedades físicas das coisas. Ele está preocupado apenas com a maneira como falamos sobre elas. Noutras palavras, as proposições da filosofia não são factuais, mas de carácter linguístico – isto é, elas não descreverem o comportamento de objetos físicos, ou mesmo mentais; antes, elas expressam definições ou consequências formais das definições (Ayer, 1936: 61-62).

Para Ayer, a forma de definir o trabalho filosófico é baseada na história da própria disciplina: podemos notar a influência de David Hume no contraste entre as “definições” de Ayer e “relações de ideias” de Hume; no mesmo sentido, a “descrição de objetos” relaciona-se com as “matérias de facto” do filósofo escocês. Mais importante, a concepção de Filosofia como método *a priori* não nos pode oferecer conhecimento empírico, nem este pode influenciar o conhecimento analítico.

No mesmo sentido, Michael Dummett atribuiu a Gottlob Frege a seguinte pressuposição metafilosófica:

Somente com Frege foi finalmente estabelecido o objeto apropriado da Filosofia: a saber, primeiro, que o objetivo da Filosofia é a análise da estrutura do *pensamento*; segundo, que o estudo do *pensamento* deve ser nitidamente distinguido do estudo do processo psicológico de *pensar*; e, finalmente, que o único método adequado para analisar o pensamento consiste na análise da *linguagem* [...] [sc. A] aceitação desses três princípios é comum a toda a escola analítica¹¹ (Dummett, 1978: 458).

Bede Rundle, um filósofo do *Trinity College* da Universidade de Oxford, já tinha antecipado de certa forma a crítica mereológica avançada pelos isolacionistas aos ombros da mesma concepção de Filosofia:

[...] versões mais cruas que creditam ao cérebro propriedades que podem ser significativamente atribuídas apenas a pessoas ou aos seus estados mentais, obrigando esse

¹¹Embora muitos filósofos que defendem a concepção isolacionista nega a prioridade da terceira condição de Frege, isto é, a redução da análise do pensamento à linguagem, e não ao pensamento diretamente.

órgão a tornar-se um *locus* de significado e intencionalidade, um repositório de regras, um lar para estados internos caracterizados gramaticalmente e assim por diante (Rundle, 1997: 25).

Para Rundle, há uma inconsistência lógica entre usar conceitos psicológicos ou mentais, como “acreditar”, e conceitos da Neurociência, como “sinapses”.

Muitos outros autores poderiam aqui ser indicados e citados. Mas mais relevante que nomear esses investigadores, será mais importante analisar as críticas que podem ser levantadas a esta concepção isolacionista que, pelo menos no seu sentido mais forte, condena qualquer intercâmbio entre a Filosofia e a Neurociência.

II.2. Reflexão Crítica:

Depois de termos refletido sobre as várias dimensões da Abordagem do Isolacionismo, focaremos agora a nossa atenção nas principais críticas e problemas desta abordagem metodológica.

II.2.1. A Objeção da “Invariância Conceptual”

Se tomarmos em consideração como exemplo a ideia de que “o cérebro decide o que quer fazer”, a resposta de Bennett e Hacker seria que tal afirmação nem sequer está errada, mas antes é uma confusão nos termos. Como já vimos, estaríamos a cometer (FM), e tal é logicamente inadmissível, pois compromete conceptualmente a investigação científica em causa.

Apesar de esta posição ser bastante interessante, e descrever muito do trabalho filosófico atual, achamos, seguindo vários autores, que possui pressupostos duvidosos.

Primeiramente, a posição apresentada pelos autores parece implicar que não há evolução do sentido das palavras: os significados estão fixados numa rede lógica particular – só assim podemos assinalar falácias do tipo mereológico. Existe, assim, um limite conceptual onde as palavras podem numa mesma proposição estar relacionadas.

Essa limitação dá origem ao seguinte esquema conceptual: as proposições podem ter ou não sentido; se tiverem sentido, podem ser verdadeiras ou falsas – é aí que o trabalho científico entra em ação. Note-se o seguinte exemplo que pode ilustrar este ponto.

A afirmação: “A mesa é vermelha” tem sentido – é logicamente compatível atribuir um predicado de cor a um objeto; além de ter sentido, esta frase pode ser verdadeira se de facto a mesa for vermelha.

Já a afirmação: “A mesa é azul” é falsa, se de facto a mesa for vermelha. Contudo, apesar de ser falsa, ela continua a ser uma frase com sentido: é totalmente possível a mesa ter sido pintada de azul, em vez de vermelho. O predicado “azul” satisfaz as condições linguísticas do seu uso. Não há qualquer problema lógico em juntar “mesa” e “azul”, mesmo que a “investigação empírica” nos mostre que a mesa é, afinal, vermelha.

Contudo, se afirmarmos “a mesa é cega”, a afirmação nem sequer pode ser considerada falsa: é uma afirmação que não faz sentido: não é verdadeira nem falsa, mas simplesmente *nonsense*. E não existe qualquer investigação empírica que possa comprovar ou refutar que “a mesa é cega”.

Todavia, não estará este pressuposto, de que as palavras não mudam de significado, simplesmente errado? Curiosamente, achamos que os autores estão a esquecer-se de uma importante lição: o próprio Wittgenstein mostrou que, por diversas razões, as palavras podem mudar o seu sentido. Por exemplo: uma palavra pode perder o seu significado prático, como acontece quando o objeto que era nomeado pelo conceito deixa de existir; ou o significado inicial pode ser alargado para referir outros objetos, etc.

Seguindo Klagge (1989), podemos defender que o facto de existir mudanças conceptuais de conceitos e palavras pode ser um obstáculo à pretensão que Hacker e Bennett oferecem em relação à rigidez dos conceitos, dos significados e das ligações entre conceitos (cf. Klagge, 1989: 320-324).

Pode dar-se o caso de, por exemplo, defender-se a ideia de que palavras ou conceitos podem ser aplicadas além das suas significações tradicionais de modo a tornar mais claro esses conceitos (cf. Crane, 2015: 254). Por exemplo, Stephen Mulhall, um filósofo também de Oxford, aponta nesse sentido em relação ao conceito de “imagem” no contexto da discussão sobre a projetabilidade da linguagem:

A palavra “imagem” denota, entre outras coisas, pinturas abstratas, pinturas representativas e filmes, isto é, imagens em movimento. Embora pinturas e filmes abstratos tenham algo em comum com pinturas representativas, parece não haver características comuns relevantes a uma pintura *dripping* de Jackson Pollock e uma imagem projetada de Humphrey Bogart, e ainda assim não temos inclinação – certo? – de dizer que a palavra tem um significado particular numa conversa sobre *Casablanca* e outro quando a conversa se foca em *Lavender Mist*. [...] Se alguém construísse tal padrão de uso a partir do zero, poderíamos achar tal bastante enigmático. Mas se o observarmos como um processo de desenvolvimento histórico, o enigma dissolve-se. No caso de “imagem”, o foco original na pintura representacional naturalmente permite uma extensão do uso do termo para fotografias e, por consequência, para filmes; e os desenvolvimentos na pintura também tornaram natural uma extensão diferente do termo para incluir telas de um estilo não-representacional (Mulhall, 2014: n/p).

Mulhall pretende mostrar que há conceitos que podem evoluir com o tempo, cujo uso pode ser estendido além do seu contexto original, não colocando em causa qualquer transgressão de sentido.

Note-se o seguinte exemplo: imagine o leitor que está, neste momento, na Antiga Grécia. Estando a conversar com alguém, essa pessoa afirma: (a) “amanhã vou voar para o Egito!”. Obviamente, dado que nos encontramos na Antiga Grécia, acharíamos que a mesma já teria bebido demasiado vinho

naquela altura, pois encontramos-nos numa época em que é impossível para a humanidade voar – apenas aves poderiam voar. Há, portanto, um impedimento lógico-gramatical que não permite que os dois conceitos sejam usados na mesma frase de forma a que faça sentido.

Contudo, é perfeitamente compreensível por todos nós se alguém, atualmente, afirmar: (b) “amanhã vou voar para Ottawa”. Neste caso, o verbo “voar” mudou¹² o seu contexto de uso e permitiu um novo enquadramento lógico-gramatical, fazendo com que a frase faça sentido e não haja, afinal, qualquer transgressão dos limites do sentido. Para ser claro: o significado de (a) no conceito “voar” está ligado ao que os pássaros e aves faziam. Já o significado em (b) altera-se, mantendo o sentido original, sendo estendido também a pessoas (cf. Crane, 2015: 256).

Portanto, é plausível descartar a crítica feroz à Neurociência da abordagem isolacionista como consequência da sua recusa em compreender a flexibilidade da linguagem natural. (cf. Pöyhönen, 2014: 99)

Esta ideia poderá estar, de resto, presente no parágrafo 67 das *Investigações Filosóficas* (1953) de Wittgenstein:

Não consigo pensar em nenhuma expressão melhor para a plausibilidade dessas semelhanças do que “semelhanças de família”; dado que as várias semelhanças entre os membros de uma família: a constituição, as feições, a cor dos olhos, a marcha, o temperamento e assim por diante – sobrepõem-se e cruzam-se da mesma maneira – E direi: “jogos” formam uma família (Wittgenstein, 1953: § 67).

Podemos, assim, seguindo Wittgenstein, defender que os usos distintos da palavra “voar” no nosso exemplo constituem uma semelhança de família que poderá elucidar a relação conceptual entre o cérebro e a atribuições de fenómenos psicológicos a essa parte do corpo humano em particular, e não ao ser humano inteiro, como afirmado em (FM).

Podemos, inclusive, afirmar que poderá ser uma nova descoberta conceptual que partes específicas do cérebro ou do corpo pensam ou interpretam ou percecionam. Esta afirmação é diferente da ideia de que tal teria o sentido de declarar que não conseguimos percecionarmos algo sem ter um lobo occipital saudável, que não conseguimos falar corretamente sem uma área de Broca funcional, entre

¹²Mais claro ainda: “voar” estava antes associado àquilo que as aves faziam; a mudança do significado passa por incluir o facto de, atualmente, existirem aviões que nos permitem deslocar rapidamente de uma cidade do mundo para outra do outro lado do mundo.

outros, algo que Bennett e Hacker aceitam facilmente. Mas poderá ser realmente uma inovação conceptual¹³ da Neurociência a tese que defende que partes do cérebro pensam ou percebem ou decidam. Se esse for o caso, então, talvez não haja qualquer falácia mereológica da Neurociência.

Patricia Churchland, defensora da abordagem reducionista analisada no próximo capítulo (cf. Capítulo 3), defende algo semelhante, afirmando

Visto que no decurso da evolução científica normal os significados das palavras sofrem mudanças em resultado da mudança de teorias, então quando uma hipótese é avançada que resultaria em mudanças de significação, tal não é por si mesmo uma objeção à hipótese [...] Dai que aquilo que a ouvidos antigos soava como algo de estranho acaba por parecer completamente habitual e correto. Tal acontece porque os ouvidos antigos são informados pela teoria antiga [...] A pretensão de que a adequação de uma nova teoria está comprometida se os seus termos não conseguem preservar a sinonímia com os termos da velha teoria é perniciosamente conservadora, porque implica que a preservação do *status quo* deve ultrapassar as considerações da adequação empírica (Churchland, 1986: 274).

Fica, assim, claro que a hipótese da invariância conceptual poderá perder a sua força se aceitarmos uma metafilosofia diferente da isolacionista, como veremos de seguida.

II.2.2. A Objeção de Dennett

Na parte final do livro *Fundamentos Filosóficos da Neurociência*, Hacker e Bennett analisam em dois apêndices o trabalho de dois dos mais importantes filósofos da mente da atualidade: o primeiro sobre Daniel Dennett, e o segundo sobre John Searle. Os apêndices são críticas ao trabalho de ambos, procurando mostrar que estão equivocados em vários pontos e que as suas ideias se encontram sujeitas a vários erros. Essas reflexões levantarão considerações metafilosóficas que vale a pena analisar.

Daniel Dennett, filósofo da mente americano, afirma que aceita a conceção da natureza do trabalho filosófico dos seus dois mestres diretos, Quine e Ryle (cf. Dennett, 1994: 237). Ora, Bennett e Hacker fazem notar que, logo à partida, esta afirmação é estranha, porque Quine e Ryle viam a Filosofia

¹³ Uma análise crítica desta posição pode ser encontrada em Slaney & Maraun (2005).

de forma oposta: enquanto o segundo acreditava que a Filosofia teria um carácter único e excepcional, o primeiro acreditava, como vimos no Capítulo 1, que a Filosofia está em continuidade com a Ciência (cf. Hacker & Bennett, 2003: 413).

Dennett, seguindo Quine, afirma que “a filosofia está unida às ciências físicas e é efetivamente a sua continuação” (Dennett, 1987: 5). Além disso, o filósofo americano defende que os problemas filosóficos são resolvidos com a combinação da investigação científica com a melhoria do nosso esquema conceptual, sempre considerado à luz das evidências empíricas (cf. Symon, 2002: 450).

Como podemos já antecipar, as ideias de Dennett são vistas por Hacker como totalmente equivocadas: como observámos em (9) e (10), a investigação empírica não pode, supostamente, resolver qualquer problema filosófico, assim como não pode resolver qualquer problema matemático.

Esta visão contraria todo o trabalho de Dennett, que é visto como um pensador que consegue agregar conhecimentos teóricos e práticos nas suas posições filosóficas em diversos problemas que investiga. Mesmo pensadores que discordam do autor de *Consciousness Explained* concordam com a relevância “do papel da investigação empírica na solução dos problemas filosóficos tradicionais” (Symon, 2002: 8).

Embora Hacker e Bennett levantem uma consideração nova, afirmando que “reconhecemos que por vezes os conceitos precisam de ser revistos ou argumentados para finalidades científicas”, os autores discordam da metafilosofia de Dennett, baseado no naturalismo de Quine, mais especificamente, (i) no naturalismo metafilosófico fraco (cf. Capítulo 1) que defende o uso da ciência empírica como auxílio para resolver os problemas filosóficos.

Dennett irá responder a todas as críticas de Hacker e Bennett num capítulo que nasce de uma sessão de “Authors and its Critics” na *American Philosophical Association* (cf. Robinson, 2007). Começa por considerar as ideias em que ambos estão de acordo:

(a) As lições da Filosofia da Linguagem Comum (*Ordinary Language Philosophy*) que Dennett aplicou nas suas ideias de Psicologia Popular, são importantes e devem ser consideradas nas investigações da Filosofia da Mente e da própria Neurociência. Os conceitos mentais que são usados por esta última têm pressupostos em relação ao seu uso que, se usados de forma errónea, podem colocar em causa as investigações, resultados e modelos empíricos, sendo que o filósofo, como um “especialista nas nuances do sentido”, pode ser exatamente a pessoa indicada para realizar esta “higiene conceptual”;

(b) As sobras do cartesianismo que podem ser encontradas em muitos neurocientistas – o teatro cartesiano, o homúnculo dentro da nossa mente que observa o espetáculo da nossa subjetividade, etc., são um erro e devem ser totalmente eliminados. Dennett é conhecido exatamente por ser um dos primeiros pensadores que refuta explicitamente o que apelidou de “cartesianismo materialista” para designar precisamente a ideia equivocada de que existe um lugar específico no cérebro – que Dennett chama de “Teatro Cartesiano” onde o filme da subjetividade acontece. Além disso, o filósofo americano concorda com a leitura de Hacker da investigação de Libet sobre ação voluntária;

(c) A ideia de que o termo *qualia*, usado por Brian Farrell (1950) e depois explicitamente utilizado por Thomas Nagel (1974), é um termo inútil e incoerente. Para Dennett, a introspeção não é uma forma de percepção interior e subjetiva, nem um “olho da mente” (*mind's eye*). Como veremos, Hacker acha incoerente afirmar-se “eu tenho uma experiência visual de vermelho”, porque “ver não é uma coisa que se tenha ou se possua” (veremos mais sobre este assunto no Capítulo 6 sobre *qualia*) (cf. Dennett, 2007: 74-75).

Contudo, as compatibilidades acabam aqui. O que acontece de seguida é um ataque severo de Dennett às considerações de Bennett e Hacker que será útil para retirar implicações metafisológicas para o problema metodológico.

Em primeiro lugar, Dennett explica que a principal ideia de Hacker, presente em (FM) é semelhante à que o próprio tinha proposto na sua distinção entre o nível pessoal e subpessoal de explicação. Note-se a seguinte citação que ilustra a distinção:

Uma análise da nossa maneira comum de falar sobre dores mostra que não foi possível descobrir no cérebro eventos ou processos que exibissem as características dos supostos “fenómenos mentais” da dor, porque falar sobre dores é essencialmente não mecânico, e os eventos e processos do cérebro são essencialmente mecânicos (Dennett, 1969: 91).

Por exemplo, numa reflexão sobre o argumento do Quarto Chinês de John Searle, Dennett responde através da “resposta do sistema” que o mesmo confunde os dois níveis de explicação no seu argumento: “Eu” compreendo chinês; o meu cérebro, não. Recordemos a experiência mental de Searle: alguém, que não fala chinês, está fechado dentro de um quarto onde há símbolos em chinês dentro de caixas; essa

pessoa possui um livro de instruções em inglês, que explica como combinar os símbolos chineses e como enviar sequências de símbolos chineses para fora do quarto, quando são introduzidos no quarto outros símbolos chineses, através de uma pequena abertura, de modo a que a pessoa de fora não consiga perceber o que se passa no quarto; a pessoa que está dentro do quarto não sabe que os símbolos introduzidos são ‘perguntas’ e os que saem são ‘respostas’.

Podemos concluir que o sistema, no seu todo, fala, portanto, chinês, na perspectiva das pessoas que estão do lado de fora. Logo, o sistema passa o famoso Teste de Turing, embora a pessoa dentro do quarto saiba que no fundo não percebe chinês. Searle concluiu que a experiência mental do Quarto Chinês torna clara a possibilidade de um sistema que tem “intencionalidade atribuída”, mas sem “intencionalidade intrínseca” (semântica genuína), ou seja, que a sintaxe não é suficiente, mas é também necessária a compreensão dos símbolos envolvidos, deitando assim por terra todo o projeto da Inteligência Artificial Forte (cf. Searle, 1980).

Ora, Dennett avança com a seguinte crítica onde a distinção entre nível pessoal e subpessoal é apresentada:

A resposta dos sistemas sugere corretamente, na minha opinião, que Searle confundiu diferentes níveis de explicação (e atribuição). Eu entendo inglês, meu cérebro não – nem, mais particularmente, a parte apropriada (se tal pode ser possível) que opera para “processar” frases recebidas e executar as minhas intenções de atos de fala (Dennett, 1980: 429).

Tornando assim implícito que, para Dennett, (FM) é uma versão similar da distinção que apresenta, vai concluir que Hacker está a querer “ensinar a missa ao Papa”, dado que foi ele o primeiro a avançar com o pressuposto da falácia mereológica (Dennett, 2007: 77)

Ademais, o problema para Dennett é que os autores fazem uma leitura extrema da passagem de Wittgenstein (8) (ver acima, secção II.1.1, pág. 33). O filósofo americano assume que o seu trabalho foi profundamente influenciado por Wittgenstein, tendo procurado estender esta posição a outras entidades como robots e partes do cérebro dado que, de certa forma, “eles se assemelham a um ser humano vivo” – algo que Hacker critica claramente – e essa semelhança é suficiente para usar, legitimamente, sem transgredir qualquer limite de sentido, um determinado atributo psicológico para caracterizar um comportamento específico (cf. Dennett, 2007: 78).

Dennett aborda, inclusivamente, este tema específico no seu *Content and Consciousness*, remetendo para o tipo de erro que uma interpretação radical de Wittgenstein pode levantar:

[...] a distinção entre os níveis pessoal e subpessoal de explicação não é nova. A filosofia da mente iniciada por Ryle e Wittgenstein é, em grande medida, uma análise dos conceitos que usamos no nível pessoal [...] [sc. e que] os níveis pessoal e subpessoal não devem ser confundidos. [sc. Mas] a lição foi ocasionalmente mal interpretada [...] como a lição de que o nível pessoal de explicação é o único nível de explicação quando o assunto é mentes e ações humanas. Num sentido importante, mas restrito, isso é verdade, pois, como vemos no caso da dor, abandonar o nível pessoal é parar de falar sobre dor. Noutra sentido importante, é falso, e é isso que muitas vezes é esquecido. O reconhecimento de que existem dois níveis de explicação dá origem ao ónus de relacioná-los entre si, e esta é uma tarefa que não está fora da província do filósofo (Dennett, 1969: 95-96).

Assim, para Dennett, Bennett e Hacker cometem exatamente o erro que aponta mais de três décadas antes: os autores afirmam que o nível pessoal de explicação é “o único nível de explicação quando o assunto em causa são as mentes e as ações humanas”, e tal poderá ser um equívoco.

Como vimos, a posição de Hacker é assumir que podemos investigar os limites do sentido através do exame cuidadoso do uso das palavras:

O contrassenso é frequentemente gerado quando uma expressão é usada contrariamente às regras do seu uso. [A] A expressão em questão pode ser uma expressão comum, não técnica, caso em que as regras para o seu uso podem ser extraídas do seu emprego comum, podendo receber explicações sobre o seu significado. Ou pode ser [B] um termo técnico, caso em que as regras para o seu uso devem ser extraídas da introdução do termo pelo teórico e das explicações que este oferece das regras estipuladas (Bennett & Hacker, 2003: 6).

O problema deste pressuposto – que também foi notoriamente usado por Gilbert Ryle na identificação dos “erros categóricos” – é que ninguém encontrou ainda estas regras do sentido. Dennett afirma até que Ryle estava a fazer bluff! – o filósofo não tinha uma lógica das expressões comuns ou ordinárias para apoiar a sua argumentação (cf. Dennett, 2007: 83).

A legitimidade de (FM) é justamente o facto de que a sua aplicação instanciada por vários neurocientistas ir contra as “regras de uso” sendo, assim, uma “transgressão dos limites do sentido”, utilizando as duas expressões mais vezes empregues por Bennett e Hacker no seu livro.

O problema é que, sem sabermos quais são essas “regras” e, por coerência, quais são os “limites”, afirmar constantemente as duas expressões parece ser um exercício filosófico pobre e, no fundo, irrelevante. Dennett chega a afirmar que, embora tal tenha sido um projeto filosófico notório do passado, nunca nenhum filósofo articulou com detalhe quais eram essas “regras de uso” das expressões comuns (*ordinary*).

Para Dennett, o trabalho da Linguística pode auxiliar a compreender este ponto. Quando um linguista analisa uma certa linguagem ou expressão que parece desviar-se do uso normal das palavras, tal não implica que ele esteja a apontar para uma regra explícita da sintaxe ou semântica de uma linguagem particular, mas poderá estar apenas a dar conta de uma característica da mesma.

Por exemplo, na língua portuguesa, a figura de estilo do oxímoro levanta várias dúvidas em relação ao sentido, se tomado literalmente. Vejam-se os seguintes exemplos:

- “O PIB de 2020 terá um *crescimento **negativo”;
- “Mohammad Gandhi era a favor da *guerra **pacífica”;
- “Com a pandemia da COVID-19, as ruas da cidade estavam com um *silêncio **ensurdecedor”;
- “Essa é a *sábua **ignorância a que Sócrates se referia”.

Nestes quatro exemplos, os termos indicados com * têm um significado que, normalmente, não faz sentido ao ser acompanhado de imediato por termos que significam o oposto, assinalados a **.

Ora, esta análise, que Hacker considera ser uma investigação *a priori*, vai ser vista por Dennett como uma investigação *a posteriori*. Para o autor de *Consciousness Explained*, quando Hacker questiona “Como podemos investigar os limites do sentido?” (Bennett & Hacker, 2003: 6), o filósofo britânico procura dar a entender que essa investigação é uma tarefa que requer apenas considerações apriorísticas sobre os conceitos, as suas ligações e implicações com outros conceitos, respondendo imediatamente à sua própria questão, cometendo uma clara falácia da petição de princípio.

Todavia, para Dennett, essa investigação poderá ser considerada, de facto, uma investigação empírica:¹⁴ “[...] ao contrário do que qualquer filósofo possa afirmar, examinar o uso de palavras é uma investigação empírica, que geralmente produz verdades e falsidades quotidianas e comuns e está sujeita à correção por observações e objeções” (Dennett, 2007: 81).

Assim, a simples descrição de que “o trabalho do filósofo é analisar o uso das palavras e as transgressões dos limites do sentido” é apenas uma afirmação declarativa inútil de carácter prescritivo que, como iremos ver, não ajuda a pensar a normatividade da natureza do trabalho filosófico.

Como Justin Sytsma (2007) afirma, a ideia de criticar a aplicação de certos conceitos como problemáticos ou de indicar qual o uso correto desses conceitos tem dois pressupostos problemáticos: (i) assumir que existe, de facto, um significado único e bem-definido do conceito e (ii) que a pessoa que analisa a correta ou incorreta aplicação desse conceito conhece (i). Ora, ambas estas ideias são controversas e assumem teses sobre a natureza da linguagem que podem estar equivocadas (cf. Pöyhönen, 2014: 100).

O projeto filosófico da abordagem isolacionista passa por analisar a forma como as pessoas empregam variados termos comumente, para depois aplicar essa mesma análise a uma segunda comunidade, a comunidade da Neurociência, afirmando que, como a primeira tem umas supostas regras de uso, então a segunda deve seguir exatamente as mesmas supostas regras. Mas talvez Hacker esteja a falhar a análise de um conceito importante de Wittgenstein, como já indicámos.

O autor das *Investigações Filosóficas* apresentou o conceito de “jogos de linguagem” com o objetivo de demonstrar que as palavras não podem ser compreendidas fora do contexto em que são usadas. Pode muito bem dar-se o caso, como apontámos acima seguindo Crane (2015), que, de facto, pode fazer sentido atribuir predicados psicológicos ao cérebro se considerarmos que estamos perante um jogo de linguagem diferente do jogo da linguagem comum, ou que, na argumentação de Dennett, podemos estender ao nível subpessoal a aplicação pragmática de termos e conceitos usualmente feita ao nível pessoal (cf. Crane, 2015: 256).

Quando neurocientistas como Crick argumentam que “o cérebro acredita em x ”, o que Crick está de facto a considerar é o nível subpessoal de explicação (cf. Dennett, 2007: 90).

Segundo Hacker, há dois tipos de expressões ou conceitos:

¹⁴Seguindo a mesma crítica, Patricia Churchland vai afirmar que “a análise conceptual (...) é um tipo de empreendimento antropológico, e deve ser possível ser completado em alguns anos” (Churchland, 1996a: 290). Uma segunda crítica pode ser encontrada em Machamer & Sytsma (2009).

[A] Os conceitos comuns, não-técnicos, cujas regras para o seu uso podem ser tornadas explícitas pelo uso comum e as explicações padrão para o seu significado;

[B] Os conceitos técnicos ou especializados, cujas regras de uso são introduzidas pelo teórico que cria o conceito e oferece explicações para o seu uso (cf. Bennett & Hacker, 2003: 6).

Para Dennett, esta divisão é profundamente problemática, principalmente o pressuposto presente em [B]. A existência de conceitos técnicos como definidos por Hacker é, para Dennett, falsa, a não ser que o filósofo de Oxford esteja a definir o termo “técnico” de forma tão específica que restrinja a quantidade de conceitos que são de facto “técnicos” segundo essa definição, dado que nenhum dos conceitos analisados na obra *Fundamentos Filosóficos da Neurociência* são, para o filósofo da *Tufts University*, conceitos [B] – dado que não foram “batizados” por um especialista. Se esse é o caso, então os conceitos do livro pertencem à categoria [A] (cf. Dennett, 2007: 84). Para Dennett, “os conceitos, como são usados na ciência cognitiva, são realmente ‘linguagem comum’ – não termos técnicos que foram explicitamente estipulados por alguma teoria” (Dennett, 2007: 92).

Hacker tinha já antecipado, de certa forma, a grande novidade filosófica do seu livro em co-autoria com Bennett em 1990, quando afirmou no seu livro *Wittgenstein: Meaning and Mind*:

Se neurofisiologistas, psicólogos, cientistas da inteligência artificial ou filósofos desejam mudar a gramática existente, para introduzir novas formas de falar, eles podem fazê-lo; mas as suas novas estipulações devem ser explicadas e as condições de aplicação devem ser estabelecidas. O que não pode ser feito é argumentar que, uma vez que sabemos o que “pensar”, “ver” ou “inferir” significa e sabemos o que “cérebro” significa, devemos saber o que “o cérebro pensa, vê, e infere” significa. Pois sabemos o que esses verbos significam apenas na medida em que dominamos o seu uso existente, o que não permite aplicá-los ao corpo ou às suas partes, salvo derivativamente. Nem se pode cruzar o novo uso “técnico” com o antigo, como, por exemplo, o neurocientista tipicamente faz nas suas teorizações. Pois tal produz um conflito de regras e, portanto, incoerência no uso desses termos pelos neurocientistas (Hacker, 1990: 148-149).

O problema, contudo, é que Hacker se encontra equivocado, como vimos, na ideia da existência de “regras” lógico-gramaticais, dado que o próprio nem sequer oferece um vislumbre do que essas regras possam ser.

Para Dennett, o argumento central de Hacker comete, como já indicámos, uma falácia *petitio principii*: se nos tivesse apresentado essas regras sempre nomeadas, mas nunca apresentadas, e nos tivesse demonstrado com exatidão de que forma é que certas expressões de neurocientistas entram em conflito com essas regras, poderíamos estar numa posição em que poderíamos concordar ou discordar das mesmas. Mas como essas regras são uma quimera filosófica, Hacker poderá estar a cometer o próprio erro que parece apontar aos neurocientistas, sobrepondo o seu jogo de linguagem ao jogo de linguagem da Ciência Cognitiva (cf. Dennett, 2007: 83).

A abordagem isolacionista, ao mostrar desprezo por qualquer contacto com o empírico, comete o erro metodológico de não se aperceber que poderá ter entrado no “jogo de linguagem” da Neurociência, tentando aplicar um “exame do uso das palavras” através das lentes oculares do próprio “jogo de linguagem” da linguagem ordinária ou comum para um território cujas regras são distintas, assim como os sentidos a serem aplicados a determinados conceitos neurocientíficos.

Este ponto é tornado óbvio quando Hacker considera que expressões como “o cérebro pensa, decide, infere, cria, percebe, etc.” não possam ser, de facto, uma inovação conceptual útil para a investigação. O filósofo de Oxford prefere assumir que se trata de uma confusão conceptual (cf. Bennett & Hacker, 2003: 70), sem apresentar qualquer argumento para escolher essa terceira opção (i.e. uma confusão conceptual) em detrimento das duas opções anteriores (i.e. uma descoberta que o cérebro participa dessas funções humanas; ou uma inovação linguística teoricamente útil).

Todavia, para Dennett, pode ser realmente um facto empírico que partes do cérebro estejam envolvidos nos processos psicológicos, e que tal envolvimento justifique a aplicação de uma postura intencional (*intentional stance*), que captura, para o autor, a parte central das “regras” de uso das expressões descritas.

Uma consequência da crítica de Dennett é negar que haja (FM): de facto, os cientistas não atribuem, *literalmente*, crenças e desejos e memórias a partes do cérebro – tal seria realmente uma falácia. Mas a aplicação do que Dennett chama de “hemi-demi-semi-proto-quase-pseudo-intencionalidade” a partes do cérebro, por exemplo, é um passo teórico útil para compreendermos melhor como é que as pessoas funcionam e como é que o cérebro consciente opera (cf. Dennett, 2007: 88-89).

O próprio Hacker tinha avançado com esta possível hipótese no seu livro de 1990, antecipando o erro que viria a criticar ferozmente em 2003:

Se toda essa [sc. ciência cognitiva] deve ser considerada seriamente, parece mostrar, primeiramente, que a observação gramatical de que esses predicados, no seu uso literal, são restritos aos seres humanos e o que se comporta como seres humanos é um *simpliciter* errado ou demonstra “inércia semântica” que foi superada pelo avanço da ciência, dado que as máquinas realmente se comportam como seres humanos. Em segundo lugar, se faz sentido literal atribuir predicados epistêmicos e até perceptivos a máquinas construídas para simular certas operações humanas e executar determinadas tarefas humanas, parece plausível supor que o cérebro humano deve ter uma estrutura funcional abstrata semelhante à do *design* da máquina. Nesse caso, certamente fará sentido, afinal, atribuir a variedade de predicados psicológicos ao cérebro humano (Hacker, 1990: 160-161).

É surpreendente que Hacker, treze anos antes, considere a possibilidade, por razões plausíveis, de estender predicados psicológicos do todo – dos seres humanos e do que se comporta como eles – a partes do ser humano. O único argumento positivo que encontramos em Hacker é a defesa da abordagem isolacionista. Focaremos as considerações metafisológicas que esta levanta de seguida. Mas antes, iremos introduzir a análise crítica oferecida por John Searle a algumas das ideias apresentadas na obra *Fundamentos Filosóficos da Neurociência*.

II.2.3. A Objeção de Searle

Tal como Dennett, Searle considera não haver uma distinção tão nítida entre a Ciência e a Filosofia. Neste aspecto, o filósofo de Berkeley segue a tradição influenciada por W.V. Quine.

Segundo Searle, o trabalho do filósofo é formalizar as grandes questões conceptuais de modo a que a Ciência possa fornecer-lhes uma resposta. Assim, Hacker concorda que o trabalho da Filosofia é clarificar e elucidar as questões filosóficas, mas rejeita completamente, como vimos em (9) e (10), que a Ciência possa responder ou fazer alguma coisa para tal. Os problemas filosóficos não podem resolvidos, solucionados ou dissolvidos pela experiência científica, pois uns são sobre o sentido e *nonsense*, e a segunda sobre a verdade ou falsidade. Searle vai criticar em vários aspetos esta posição metafisológica.

Para Searle, o cerne do debate está na interpretação que Hacker e Bennett fazem de Wittgenstein em (8) (ver acima, secção II.1.1, pág. 33): a interpretação de Hacker baseia-se na ideia de que o comportamento funciona como um critério lógico para aplicação de determinados predicados psicológicos ou mentais. Há uma conexão necessária entre o sentido dos conceitos mentais e o

comportamento observável. Por isso mesmo, e como vimos anteriormente, nenhum cérebro exibe o comportamento necessário para ser atribuído o predicado de ser consciente dado que um cérebro não se pode comportar como uma pessoa no seu todo.

Contudo, para o autor de *Mente, Cérebro e Ciência*, as consequências das pressuposições baseadas em Wittgenstein dos autores não são a que estes relatam. Assumindo que os mesmos estão certos em relação ao facto de que devemos distinguir entre a dor e o comportamento da dor, isto é, entre os critérios de aplicação dos predicados mentais e o comportamento observável de certos fenómenos, não podemos confundir as distinções.

Para Searle, Hacker, como os antigos ultrabehavioristas, confunde “a evidência comportamental dos estados mentais com a existência dos próprios estados mentais” e, na versão de Hacker de behaviorismo, “confunde os critérios que servem de base para a aplicação dos conceitos mentais com os próprios estados mentais” (Searle, 2007: 102-103).

Assim, para Searle, os isolacionistas cometem o erro de confundir a atribuição de predicados psicológicos aos factos atribuídos por esses predicados psicológicos. O ponto de Searle é focar-se num aspeto central da afirmação de Wittgenstein: o foco deste é aquilo que podemos dizer (*say*), como sublinhado em (8).

Searle faz o seguinte exercício: troquemos o verbo “dizer” (*to say*) para “ser” (*to be*), de forma que (8) fique (8*): “somente seres humanos e o que se comporta como um ser humano vivo podem *ser* conscientes”.

Ora, se assumirmos (8*) como uma regra lógica, esta é obviamente falsa e não somente sem sentido, dado que existem muitos outros seres, como crustáceos, moluscos, diversos mamíferos (entre outros) que, apesar de não se comportarem como seres humanos, poderão ser conscientes se tiverem os necessários mecanismos neurobiológicos e sistemas nervosos.

O ponto de Searle é o de que o comportamento é, neste caso, irrelevante para sabermos se um determinado ser vivo é ou não consciente: esta questão não pode ser determinada por uma mera análise linguística, como parecem Hacker e Bennett supor (cf. Searle, 2007: 104).

Assim, Searle não tem dúvidas de acusar Hacker de confundir falaciosamente as regras de uso linguístico dos conceitos ou palavras com a própria ontologia do mundo e de fenómenos, sendo este um erro que era também cometido pelos antigos behavioristas, que confundiam a evidência de estados mentais pela ontologia desses estados mentais: “É uma falácia dizer que as condições para a operação bem-sucedida do jogo de linguagem são as condições para a existência do fenómeno em questão” (Searle, 2007: 105).

Para o filósofo de Berkeley, há pelo menos duas formas de atribuir predicados psicológicos ao nível subpessoal do cérebro: (i) ver o cérebro como um sujeito e agente (e.g. o cérebro pensa); (ii) ver o cérebro como o local onde esses processos psicológicos ocorrem (e.g. o pensamento acontece no cérebro).

Ora, o argumento central de Hacker é afirmar que, se seguirmos (8), não faz sentido que (i) seja verdadeiro. Contudo, a partir deste passo, Hacker também conclui que (ii) não faz sentido e é um erro conceptual. Mas para Searle, a questão em (i) é substancialmente diferente de (ii) e ambos os casos requerem argumentos independentes. O facto de (8) poder aplicar-se a (i) não permite, necessariamente, que se aplique a (ii).

Se seguirmos com rigor o argumento de Wittgenstein, mesmo assumindo que ele não é problemático, é duvidoso que coloque em causa (ii): o argumento apenas indica que o agente dos fenómenos mentais em questão deve ser um sistema capaz de exibir o comportamento apropriado desses fenómenos mentais. Embora não possamos, seguindo (8), afirmar que o cérebro vê, mas apenas a pessoa, tal não nos informa minimamente sobre se podemos identificar a experiência visual no cérebro. O argumento de Wittgenstein apenas nos obriga a considerar o cérebro como parte do mecanismo causal de um sistema (e.g. uma pessoa) que na sua totalidade exibe o comportamento de ver (cf. Searle, 2007: 108).

Deste modo, Searle conclui que o máximo que o argumento de Wittgenstein pressupõe é que o cérebro não pode funcionar como um sujeito dos processos mentais, mas tal em nada informa sobre se o cérebro é o local onde esses processos mentais acontecem: “O seu argumento contra o cérebro como sujeito não pode ser aplicado ao cérebro como *locus*” (Searle, 2007: 111).

Ademais, mesmo que se aceite (FM), a sua aplicação para este problema em particular pode ser colocada em causa dado um pormenor importante: Hacker e Bennett nunca chegam a indicar o que é uma pessoa (cf. Searle, 2007: 119). Parecem, contudo, cometer o erro de assumir que o cérebro faz parte da pessoa – daí poder haver uma falácia mereológica, dado que a parte (i.e. o cérebro) é confundida com o todo (e.g. a pessoa). Mas esta pressuposição é falsa: um cérebro não faz parte de uma pessoa tal como uma gota de água não faz parte do oceano. O cérebro fará, no máximo, parte do corpo da pessoa, e não da pessoa em si. Este erro cometido por Hacker é irónico, dado que é puramente a continuação do cartesianismo e da concepção de identidade pessoal tradicional, sendo um erro categorial que não pode ser permitido (cf. Harré, 2012: 352).

Além de Searle criticar a interpretação behaviorista radical de Hacker e Bennett de Wittgenstein, os autores colocaram também em causa as pressuposições metafilosóficas de Searle. Segundo eles, Searle afirma que a Filosofia e a Ciência são semelhantes em dois aspetos:

- (i) Universais no seu conteúdo;
- (ii) Visam a verdade.

Mas também são dissemelhantes em três aspetos:

- (iii) A Filosofia, ao contrário da Ciência, debate questões cujas respostas não podem ser alcançadas através de um método sólido e sistemático;
- (iv) As questões da Filosofia são questões mais gerais, de “enquadramento”, enquanto as questões da Ciência são questões mais particulares e individuais;
- (v) As questões da Filosofia são essencialmente problemas conceptuais (cf. Searle, 1999: 2069).

Para Searle, a relação entre a Filosofia e a Ciência é próxima porque, no fundo, possuem o mesmo tipo de empreendimento: ambos são uma atividade cognitiva que procuram melhorar a nossa posição epistémica no mundo. Mas, embora semelhantes nesse aspeto, têm diferentes objetivos. O objetivo da Filosofia é colocar a pergunta certa esclarecendo confusões teóricas e conceptuais. Logo que esse esclarecimento é alcançado, a Ciência entra em ação resolvendo o problema em questão. O problema que, anteriormente era visto como filosófico passa, neste processo, a ser um problema científico que pode ser resolvido através dos métodos empíricos da ciência.

Esta visão já tinha sido defendida por Bertrand Russell no seu famoso *The Problems of Philosophy*, onde afirmou que “[...] logo que um conhecimento definido relacionado com qualquer assunto se torna possível, este assunto deixa de se chamar filosofia, e torna-se uma ciência independente” (Russell, 1967: 90). No mesmo sentido, J. L. Austin via o trabalho filosófico de forma semelhante – recorde-se que Austin foi professor de Searle – afirmando que:

Na história da investigação humana, a filosofia ocupa um lugar do sol centralizado primordial, seminal e tumultuoso: de tempos a tempos atira para o exterior uma parte de si mesmo para se tornar numa ciência, num planeta, fresco e equilibrado, progredindo de modo estável em direção a um estado final muito distante. Isto aconteceu há muito tempo com o nascimento

das matemáticas, e mais uma vez com o nascimento da física: só no século passado testemunhámos outra vez o mesmo processo [...] com o nascimento da lógica matemática, através do trabalho conjunto de filósofos e de matemáticos. Será possível que o próximo século assistirá ao nascimento [...] de uma verdadeira e abrangente ciência da linguagem? Nessa altura, ter-nos-emos visto livres de mais uma porção da filosofia (ainda sobrará muita), da única maneira de nos vermos livres da filosofia, dando-lhe um chuto para cima (Austin, 1961: 180).

Searle vai seguir esta linha metafilosófica. Esta forma de investigação racional sobre o mundo parece explicar porque é que a Filosofia parece não ter qualquer sucesso epistémico nem fazer qualquer avanço ao passo que a Ciência parece ser bem-sucedida: assim que a Filosofia consegue aprimorar o problema de tal forma que pode ser estudado sistematicamente, o mesmo é transferido do escopo dos filósofos para o escopo dos cientistas que, fazendo uso dos instrumentos científicos, resolvem a questão com mais ou menos dificuldade (cf. Searle, 1999: 2070).

Ora, embora Hacker e Bennett aceitem a conceção da Filosofia como elucidação conceptual, não a concebem como sendo uma proximidade entre a Ciência e a Filosofia, defendendo, antes, a abordagem isolacionista, argumentando que existe efetivamente uma divisória radical profunda entre a prática filosófica e a prática das ciências naturais em geral (cf. Bennett & Hacker, 2003: 437-438).

O cunho naturalista da metafilosofia de Searle é assim criticado pelos autores isolacionistas, porque não faz sentido que a Filosofia tenha como objetos de investigação os fenómenos naturais que Searle investiga, como seja, por exemplo, a consciência. Nem é o trabalho da Filosofia criar teorias e hipóteses sobre esses fenómenos naturais, algo que Searle tenta fazer em variados temas e problemas.

Ora, para Hacker e Bennett, Searle está enganado em quase todas as suas pressuposições metafilosóficas. É difícil perceber como é que a Ciência é universal, como afirmado em (i), dado que se trata somente de um empreendimento para investigar os fenómenos naturais, ficando assim de fora as Artes, a História ou a Matemática. Mas, ao contrário da Ciência, a Filosofia pode ter um carácter universal se entendida como sendo uma investigação conceptual sobre qualquer domínio de pensamento em que conceitos e palavras são usadas, seja na História, na Ciência, nas Artes ou na Matemática.

A questão do grau presente em (iv) também parece ser duvidosa, dado que há questões filosóficas efetivamente gerais (e.g. o que é uma boa vida) mas também especificamente particulares (e.g. um movimento voluntário é um movimento antecedido por um ato de vontade?) (cf. Hacker e

Bennett, 2003: 440). As respostas de Searle não são muito convincentes: reafirma, baseando-se no seu próprio trabalho, algumas ideias metafilosóficas que procuram funcionar como prova para a sua posição.

Não será de estranhar esta fraca produção metafilosófica dado que Searle começa por afirmar que “(...) na minha experiência, disputas sobre a natureza da filosofia tendem a ser infrutíferas e usualmente expressam apenas preferências por diferentes projetos de investigação” (Searle, 2007: 122).

Podemos, contudo, ainda retirar algumas ilações relevantes. Para Searle, é falso que a Filosofia não possa oferecer conceções teóricas gerais sobre fenómenos naturais, dado que foi exatamente isso que Searle fez durante a sua carreira, tendo apresentado uma teoria geral dos atos de fala e significado, uma teoria geral da consciência ou uma teoria geral da intencionalidade, entre outras.

Assim, o ponto central da discussão e onde os filósofos discordam não é tanto sobre o que significa fazer Filosofia, dado que, supostamente, todos concordam que se trata de uma investigação conceptual, mas, antes, do que é que significa exatamente fazer uma investigação conceptual, isto é, qual é a natureza dessa análise conceptual (cf. Searle, 2007: 122).

No fundo, a discordância fundamental entre Searle, Dennett e os isolacionistas parece ser a natureza dos problemas filosóficos e científicos. Para os segundos, a Filosofia é fundamentalmente uma investigação *a priori*. Para os primeiros, abraçando a intuição de Quine, embora algum do trabalho filosófico possa realmente ser análise conceptual, este também pode envolver interação direta ou indireta com investigação efetivamente empírica.

II.2.4. A Objeção da “Prática e Anti-excepcionalidade Filosófica”

Estamos perante o problema metodológico e sobre que metodologia devemos adotar nas investigações sobre os problemas da Filosofia da Mente. Como é que os profissionais da Filosofia, isto é, os investigadores em Filosofia da Mente, pensam sobre este problema?

No artigo “O que os filósofos da mente realmente fazem? Alguns dados quantitativos”, Joshua Knobe (2014) apresenta algumas considerações sociológicas para a nossa discussão. Comparando uma amostra dos artigos mais citados entre 2009 e 2013 com uma amostra dos artigos mais citados de um período do século XX (entre 1960-1999) em Filosofia da Mente, os seguintes resultados¹⁵ foram encontrados:

¹⁵ O autor divide os métodos aplicados nos artigos de cada período por a) *A priori*: artigos que não fazem uso sistemáticos de estudos empíricos; b) *Experimental*: artigos que reportam resultados experimentais originais; e c) *Baseia-se no Empírico*: artigos

- Na amostra do século XX, a distribuição dos métodos utilizados pelos filósofos era a seguinte:

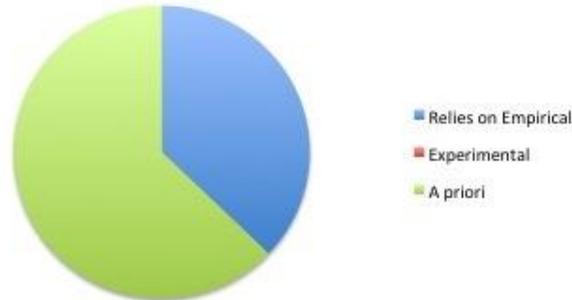


Figura 2.1.: Métodos utilizados no Século XX (cf. Knobe, 2014: n.p).¹⁶

Sucintamente, a maior parte dos artigos eram *a priori*, com uma minoria a incorporar algum trabalho empírico e nenhum a referenciar resultados originais de estudos experimentais.

- A distribuição da amostra dos artigos entre 2009 e 2013 é a seguinte:

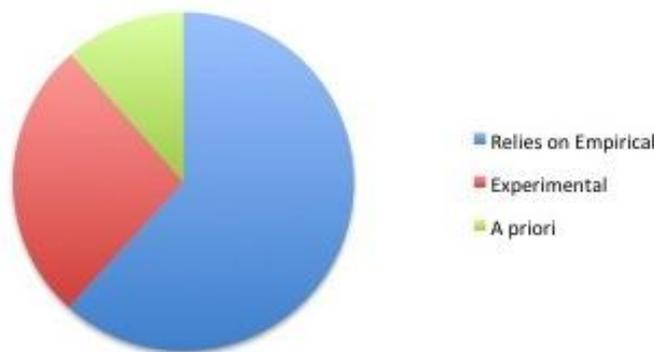


Figura 2.2.: Métodos utilizados entre 2009-2013 (cf. Knobe, 2014: n.p)¹⁷

que não se reportam a resultados experimentais originais, mas que fazem uso de resultados de estudos empíricos presentes na literatura científica.

¹⁶ Legenda da Figura 2.1.: cerca de 5/8 da investigação é efetuada por métodos *a priori*; 3/8 é efetuada por métodos baseados indiretamente no empírico; nenhuma investigação é efetuada por métodos empíricos diretos.

¹⁷ Legenda da Figura 2.2.: cerca de 5/8 da investigação é efetuada por métodos baseados indiretamente no empírico; 2/8 da investigação é efetuada por métodos empíricos diretos; e somente 1/8 da investigação é efetuada por métodos *a priori*.

Um resultado flagrante deste segundo gráfico é mostrar que há uma mudança radical na forma de fazer Filosofia da Mente na contemporaneidade. Apenas uma minoria de artigos se fundamenta em investigações puramente *a priori*. A maior parte dos artigos refere de alguma forma descobertas empíricas, e uma larga percentagem faz referência a resultados originais de experimentos (cf. Knobe, 2014).

Assim, poderá surgir a convicção que a Filosofia da Mente deve atualizar-se e deve esquecer a imagem tradicional do trabalho filosófico, ilustrada em (9) e (10). Antes, as investigações *a posteriori* devem ser consideradas na tarefa filosófica,¹⁸ se o objetivo for alcançar algum progresso nas difíceis questões que a consciência ou a subjetividade têm criado à ciência contemporânea.

Será que esta investigação de carácter sociológico em relação às crenças dos investigadores da Filosofia da Mente sobre a natureza do seu trabalho é um argumento contra a abordagem isolacionista? Embora estes dados sejam interessantes, estas considerações descritivas pouco parecem fazer pelas considerações normativas que devemos retirar para a metafilosofia.

O facto de a maior parte dos investigadores da Filosofia da Mente na atualidade fazer uso de metodologias tendencialmente empíricas e menos apriorísticas não serve de prova a favor de determinada metafilosofia, dado que tal afirmação constituiria uma versão da falácia *ad populum*, o que deve ser evitado. Apesar disso, alguns investigadores discordam efetivamente de uma defesa apriorística da Filosofia e argumentam que a abordagem isolacionista já não faz sentido e é obsoleta (veremos detalhadamente porquê no seguinte capítulo).

Em primeiro lugar, Williamson argumenta que a ideia, presente na abordagem isolacionista, que a Filosofia tem um papel excecional e único, diferente de todas as outras ciências, é problemática:

[...] a pressuposição comum do excecionalíssimo filosófico é falsa. Mesmo a distinção entre o *a priori* e o *a posteriori* acaba por obscurecer semelhanças subjacentes. Embora existam diferenças metodológicas reais entre a filosofia e as outras ciências, como realmente praticadas, elas são menos profundas do que se costuma supor (Williamson, 2007: 3).

Para Williamson, a Filosofia é uma Ciência como outras ciências e, mais importante, o conhecimento filosófico é conhecimento sobre a realidade e não somente elucidacões de sentido (Williamson, 2007).

¹⁸ Tal poderá implicar que a formação do investigador de Filosofia da Mente mude radicalmente: deve ser incorporado na sua formação ensinamentos básicos de Neurociência, Psicologia Evolutiva e Ciência Cognitiva em geral, de modo a que o investigador tenha ferramentas suficientes para compreender a linguagem usada pelo mundo científico.

Assim, esta posição critica o radicalismo da abordagem isolacionista no aspeto central em relação ao grau de isolamento entre a Filosofia em relação à Ciência. Podemos distinguir metodologicamente que os problemas conceptuais da Filosofia são diferentes dos problemas factuais da Ciência, sem, contudo, ignorar que ambos se tocam e interagem em variados aspetos, seja em questões ou problemas, argumentos, teorias, conceitos, entre outros.

Williamson¹⁹ argumenta contra esta posição devido à obscuridade epistémica da distinção entre o *a priori* e o *a posteriori*. Num artigo recente, o filósofo defende a visão oposta da abordagem isolacionista e da tese de excecionalidade da Filosofia, argumentando a favor de uma visão anti-excecionalista da Filosofia:

[sc. A filosofia], como a matemática, é uma ciência, mas não uma ciência natural. Os seus métodos não são primariamente experimentais, embora possam basear-se nos resultados da ciência natural. Como a matemática fundacional, os seus métodos são abduativos e dedutivos. Como nas ciências naturais, o progresso da filosofia consiste na construção de melhores modelos e não na descoberta de novas leis. Não nos devemos preocupar se a filosofia é *a priori* ou *a posteriori*, porque a distinção é epistemologicamente superficial (Williamson, 2019: 19).

Como veremos no capítulo seguinte, uma crítica que irá ser levantada pela abordagem reducionista a este programa de investigação isolacionista será exatamente a forma como a Filosofia e a Ciência são definidas pela última. Ao contrário do assumido, dado que a Filosofia e a Ciência não são essencialmente diferentes, a primeira poderá ser reduzida à segunda com o objetivo de alcançar progresso epistémico a respeito do problema da mente consciente, por exemplo.

Assim, o grande erro da abordagem isolacionista poderá ser a sua posição metafilosófica que, como vimos, defende que a Filosofia, além de não fazer parte da Ciência ou de não se relacionar diretamente com esta, não faz qualquer contribuição epistémica para o conhecimento humano, dado que se defende uma visão anticognitivistista da Filosofia. Pelo contrário, a Filosofia, como disciplina de segunda ordem, procura esclarecer ou elucidar “aquilo que já sabemos”: a sua tarefa é a compreensão pela clareza.

¹⁹ Embora Williamson admita, logo a seguir, que irá oferecer uma defesa de uma determinada forma de fazer filosofia de *armchair*. Há, assim, uma tensão entre as duas posições que poderá ser problemática, como defendido por Kornblith (2009).

Finalmente, e como veremos no capítulo quinto, Georg Northoff também critica esta abordagem isolacionista, argumentando que não existem conceitos puramente filosóficos e conceitos puramente científicos, como é dado a entender em (9) e (10). Antes, o filósofo e neurocientista alemão considera as seguintes ideias:

(11) Não existe uma Neurociência livre de conceito: os conceitos são relevantes para a Neurociência no teste da plausibilidade conceptual das descobertas neurocientíficas;

(12) Não existe uma Filosofia relevante e sem factos do mundo natural; os factos são relevantes para uma Filosofia naturalmente relevante ao testar a plausibilidade empírica dos conceitos filosóficos (cf. Capítulo 5).

De (11) e (12), conclui, apelando à fórmula kantiana, “filosofia sem ciência é vazia, ciência sem filosofia é cega” (Northoff, 2014a).

Esta conclusão vai fundamentalmente contra os pressupostos apresentados ao longo deste capítulo, enquadrando-se num quadro (novamente) com influência de Quine – que apresentaremos ao detalhe no próximo capítulo – de uma epistemologia naturalizada que entende a Filosofia não como oposta à Ciência, mas, antes, uma continuação, abstratamente formalizada, da mesma (cf. Capítulo 5).

III.3. Conclusão:

Defendendo uma visão muito particular do trabalho filosófico e concluindo que o método filosófico é irrelevante para a Neurociência e vice-versa, dado que ambas as disciplinas estão preocupadas com diferentes assuntos – a Filosofia com o que faz ou não sentido, e a Ciência com o que é verdadeiro ou falso – o problema metodológico é solucionado pelo paralelismo entre as duas disciplinas e pela consequente conclusão da impossibilidade de haver uma interdisciplinaridade que resolva os problemas levantados pelos fenómenos mentais e subjetivos. Como tentámos argumentar, contudo, a visão metafilosófica do isolacionismo confunde as suas considerações descritivas com as normativas: é partindo de, por exemplo, (9) e (10), que o trabalho filosófico de Dennett e Searle é considerado confuso e, em muitos casos, como “contrassenso”. Ora, devido a uma definição essencialista da Filosofia, dado que a Filosofia é mais do que apenas uma elucidação de sentido, concluímos com a ideia de que a abordagem isolacionista é demasiado exclusiva, ignorando outras considerações metafilosóficas sobre a Filosofia e a Ciência que podem ter real valor e que passaremos a analisar nos próximos capítulos.

III. NEUROFILOSOFIA REDUTIVA

Como vimos no capítulo anterior, a primeira resposta ao problema metodológico tem como base a tese de que a Filosofia e a Ciência são fundamentalmente diferentes e, por isso, não podem trabalhar em conjunto. Neste capítulo, iremos investigar a abordagem que pode ser considerada como oposta às ideias principais da abordagem isolacionista, defendendo que, pelo contrário, a Neurociência tem muito a contribuir para a Filosofia, desde que esta se atualize para um método que incorpore o conhecimento empírico.

A Abordagem do Reduccionismo (AR) procura, antes de mais, ser uma “aplicação de descobertas neurocientíficas a questões tradicionais da filosofia” (Bickle, 2009: 4). O objetivo é reformular o trabalho filosófico que é tido, até então, como um trabalho de análise conceptual baseado numa metodologia *a priori*, por uma metodologia que incorpore os dados da Ciência e possa usá-los para resolver os problemas tradicionais da Filosofia da Mente. Nas palavras de Churchland:

Os factos [*sc.* da ciência] são os factos; a análise linguística [*sc.* *a priori*] poderá informar-nos sobre o que uma pessoa comum acredita, mas não poderá ser um guia fidedigno sobre os factos do mundo, sobre a real natureza das coisas (Churchland, 1996a: 290).

Contudo, (AR) é bastante mais radical que a primeira citação: na sua base está a tese fundamental de que os conceitos filosóficos devem, tendencialmente, ser substituídos por conceitos neurocientíficos. Vejamos quais os princípios-base pelos quais se rege (AR) para depois analisarmos os pressupostos metafilosóficos desta abordagem. Para tal, iremos analisar dois proponentes, em diferentes versões, de (AR): Patricia Churchland e John Bickle.

III.1. Pressupostos da Abordagem Reducionista

III.1.1. Negação do Dualismo

O primeiro passo de um proponente reducionista é negar o dualismo (quer de substâncias quer de prioridades) por achar que o uso da Lei de Leibniz (LL) – especificamente, do princípio da identidade dos indiscerníveis – para provar a dualidade da mente/corpo é inválido. A primeira formulação explícita aparece na secção 9 do *Discurso de Metafísica* de Wilhelm G. Leibniz (cf. Loemker 1969: 308).

A (LL) afirma que se x e y são idênticos, então x e y devem partilhar necessariamente todas as mesmas propriedades. Em notação lógica, (LL) é expressa da seguinte forma:

$$(LL) \quad \forall F (Fx \leftrightarrow Fy) \longrightarrow x = y.$$

O objetivo é, então, encontrar uma propriedade que seja verdadeira sobre estados cerebrais, mas que não o seja sobre estados mentais (ou vice-versa). Veja-se (1) o argumento a favor do dualismo:

(1.1) Os meus estados mentais [EM] são conhecidos introspectivamente como estados do meu eu consciente;

(1.2) Os meus estados cerebrais [EC] não são conhecidos introspectivamente como estados do meu eu consciente;

(1.3) [Aplicando (LL)], EM *diferentes* EC.

O argumento (1) parece, assim, provar logicamente que os estados mentais são diferentes dos estados cerebrais. Se a aplicação de (LL) for legítima quanto às premissas (1.1) e (1.2), então estaremos perante um argumento sólido a favor do dualismo.

Qual é então o erro do argumento para o reducionista? A premissa (1.1) não parece ser problemática, dado que os *qualia* são definidos exatamente como algo que é conhecido introspectivamente. Em contraste, a premissa (1.2) aparenta ser mais problemática. Para o reducionista, o argumento patenteia cometer uma petição de princípio dado que só se pode concluir a veracidade da premissa se se aceitar desde logo a conclusão do argumento. Ou seja, se de facto estados mentais são idênticos a estados cerebrais, então quando alguém introspecciona um estado mental, introspecciona de

facto um estado do seu cérebro: pode não descrevê-los como idênticos – a identidade pode ser um facto do próprio mundo que é independente do conhecimento que tem dessa identidade.

Tal aponta para um erro de raciocínio que o argumento comete, nomeadamente uma falácia intensional. O defensor do reducionismo não aceita assim a aplicação de (LL) nestes argumentos, porque tal passo é ilegítimo. Analisemos o seguinte (2) argumento:

- (2.1) A aspirina é reconhecida por João como aliviando a dor;
- (2.2) O ácido acetilsalicílico [AA] não é reconhecido por João como aliviando a dor;
- (2.3) [Aplicando (LL)], a aspirina não é idêntica a AA (cf. Churchland, 1981: 329).

Apesar da verdade das premissas, a conclusão é claramente falsa: a identidade dos termos é obviamente verdadeira. Deste modo, o argumento é claramente inválido. O problema é que a “propriedade” descrita em (2.1) e em (2.2) consiste apenas no conhecimento do sujeito. Mas tal propriedade não é uma propriedade genuína do próprio objeto. Além disso, mesmo em versões mais interessantes do argumento, os problemas continuam, porque assumem logo de início o que o argumento pretende supostamente concluir, cometendo uma petição de princípio (cf. Churchland, 1981: 325-333).

Assim sendo, o (AR) avança com o conceito de reducionismo que passaremos a detalhar de seguida.

III.1.2. O Conceito de Reduccionismo

Embora o conceito de reducionismo possa ter várias definições dependendo da sua aplicação, usaremos este termo seguindo o seu uso em Filosofia da Ciência. Neste sentido, “redução” é aplicada entre duas teorias: a teoria reduzida (T_R) encontra-se numa relação particular com uma teoria mais básica (T_B). Assim, qualquer fenómeno de (T_R) que se diga ser reduzido a outro fenómeno de (T_B) é, antes de tudo, uma consequência da redução das teorias a que pertencem, e não do fenómeno em si (cf. Churchland, 1989: 278).

Assim, não são os fenómenos que são diretamente reduzidos. O que é reduzido é uma Teoria, que compreende um conjunto de especificidades que poderão ou não ser conectadas com o conjunto de especificidades da teoria reduzida. O tipo de ambição reducionista tem o nome de “redução interteórica” cujo funcionamento exemplificaremos através de exemplos da história da Ciência.

Veja-se o exemplo da luz: quando afirmamos que a luz foi reduzida à radiação eletromagnética, esta afirmação tem dois sentidos relevantes. O primeiro é afirmar que a teoria corpuscular da luz foi reduzida à teoria da radiação eletromagnética. O segundo afirma que a primeira teoria foi reduzida de tal forma que podemos reduzir o conceito de luz ao conceito de radiação eletromagnética.

O importante a retirar desta análise é que o segundo passo é dependente do primeiro, e não vice-versa: apenas havendo a redução de uma teoria à outra é que podemos alcançar a redução dos fenómenos da teoria reduzida aos fenómenos da nova teoria.

Por analogia e focando no problema mente-corpo, a redução dos estados mentais a estados cerebrais apenas poderá ser alcançada se houver, antes, a redução de uma teoria sobre estados mentais a uma teoria neurobiológica. Se (T_n) for de tal forma reduzida a (T_b) , então a redução de estados mentais a estados cerebrais poderá ser legítima.

A redução da teoria, se for bem-sucedida, tem duas implicações positivas. Em primeiro lugar, com a redução de uma teoria à outra, ganhamos unificação explicativa: se conseguirmos de facto reduzir uma teoria a outra, então a compreensão que temos dos fenómenos dessa teoria é aumentada.

Uma segunda implicação relevante é que com a redução interteorética alcançamos uma simplificação ontológica: com o sucesso da redução, dois fenómenos que pareciam ser ontologicamente distintos poderão ser, afinal, o mesmo; ou um dos fenómenos que se acreditava ter propriedades ontológicas deixa de ter existência. Com este processo, há uma simplificação ontológica que é garantida pelo sucesso da redução de uma teoria a outra.

Voltando ao exemplo da luz, antes do desenvolvimento da teoria eletromagnética, acreditava-se que a luz era um fenómeno específico dissemelhante dos efeitos eletromagnéticos que eram tidos como fenómenos distintos da luz. Ora, com a redução de uma teoria à outra, compreendemos que os dois fenómenos que antes eram tidos como dissímeis são, afinal, o mesmo. Este processo garante uma simplificação ontológica relevante.

Como veremos mais à frente, pode dar-se o caso de, em vez de haver redução de um fenómeno a outro, poderá haver a eliminação do fenómeno reduzido (cf. Churchland, 1989: 281). Ora, o reducionismo, desde a sua primeira aparição, tem sido entendido pelo menos de três formas, que passaremos a detalhar de seguida (cf. Ney, s/d).

III.1.2.1. Redução como Tradução (RcT)

A primeira forma de redução que podemos analisar aparece com Rudolf Carnap. A grande ambição de Carnap era alcançar a unificação das ciências numa só linguagem comum através da redução de diversas teorias a uma teoria comum. Como vimos no Capítulo 1, Quine viria a criticar exatamente este aspeto da filosofia neopositivista. Podemos formular a primeira versão de redução da seguinte forma:

(RcT): (T_2) é reduzida a (T_1) quando todas as verdades sobre (T_2) tenham sido traduzidas para a linguagem de (T_1) ;

Nas palavras de Carnap, (RcT) é descrita da seguinte forma: “Diz-se que um objeto (ou um conceito) é redutível a um ou mais objetos se todas as afirmações sobre ele puderem ser transformadas [sc. traduzidas] em afirmações sobre esses outros objetos” (Carnap, 1967: 1).

Carnap irá, assim, defender que a linguagem ou teoria da física poderia servir para reduzir todas as restantes linguagens – inclusive os conceitos filosóficos – a uma linguagem universal. Esta redução permitiria a unificação da ciência e do seu método, algo fundamental para os positivistas lógicos de Viena que procuravam defender que todas as afirmações empíricas poderiam – e deveriam – ser reduzidas a um único idioma e a um método uniformizado de investigação de todo o tipo de entidades por todas as ciências (cf. Carnap, 1934: 32).

O ponto relevante para a nossa investigação é a motivação dos positivistas lógicos para alcançar tal redução: dado os métodos pouco rigorosos da filosofia e das ciências sociais em geral, baseados nas intuições e análise da linguagem puramente *a priori*, seria necessário procurar uma metodologia mais rigorosa. Portanto, o objetivo de Carnap era substituir esses métodos epistemicamente pobres por métodos usados nas ciências empíricas: a intuição deixava de ter um papel relevante na investigação e era substituída pelos factos (*data*) e por verificação intersubjetiva.

O objetivo, então, de reduzir todas as afirmações de todas as ciências a uma linguagem comum serviria para alcançar uma objetividade epistémica que garantisse que os fenómenos estudados pudessem ser compreendidos na sua totalidade. Mais importante, a finalidade do método de redução como tradução serviria para tornar mais robusta a capacidade de previsão das teorias científicas – algo que era impossível com os métodos da Filosofia.

Assim, alcançada a redução por tradução, os cientistas poderiam ter uma visão mais rigorosa do mundo e do seu conhecimento. Otto Neurath, na mesma linha de Carnap, viria a afirmar que as

variadas ciências empregam diversos conceitos e teorias e que o trabalho do projeto da unificação da ciência é exatamente perceber se esses diferentes conceitos se estão a referir, no fundo, aos mesmos objetos de estudo dado que os diferentes conceitos poderão ter a mesma função. Para conseguir compreender esse fenómeno, é necessário fazer uma avaliação periódica regular do discurso científico de modo a que os conceitos não sejam multiplicados “além da necessidade” (cf. Neurath, 1983: 172-173).

Comprendemos, assim, dois pontos fundamentais:

- (i) A preocupação de aplicar uma forma de navalha de Ockham linguística de modo a minimizar os conceitos e termos de uma teoria para alcançar simplicidade teórica, dado que uma teoria deve conseguir explicar e prever fenómenos fazendo uso somente dos conceitos estritamente necessários; a tradução dos termos mentalistas da Filosofia, por exemplo, em conceitos neurocientíficos seria um desses exemplos de redução que seria epistemicamente útil;
- (ii) A unificação da Ciência dá-se pela compreensão de que, talvez, diferentes ciências usam diferentes conceitos para referir o mesmo fenómeno: apenas através da redução poderíamos alcançar essa realização; por exemplo: pode dar-se o caso de que os conceitos da Biologia “estímulo” e “resposta” podem ser dois termos usados para referir o mesmo fenómeno que a Física descreve por “causa” e “efeito”.

Uma consequência filosófica relevante desta abordagem reducionista seria adotar uma teoria coerentista da verdade, em oposição à teoria por correspondência. A teoria da verdade por correspondência defendia que uma proposição p é verdadeira se p corresponder a um determinado estado do mundo ou realidade, algo que é fundamentalmente extralinguístico. Se p corresponder a um estado do mundo, será verdadeira; se p não corresponder a um estado do mundo, será falsa.

Pelo contrário, a teoria coerentista da verdade (tCV) pode ser formalizada da seguinte forma:

(tCV): p é verdadeira se e só se é coerente com um conjunto particular de ρ s.

Assim, nesta versão da teoria da verdade, a justificação de afirmações dá-se por coerência com outras afirmações, e não com uma correspondência direta com o mundo externo. Neurath apresentou, para defender esta teoria, a famosa metáfora do marinheiro que tem de reconstruir o navio encontrando-se em alto mar. A metáfora procura defender que é impossível sustentar uma base imutável e

absolutamente objetiva da Ciência, pelo que a base das afirmações científicas poderá estar no consenso – ou na falta dele – dos cientistas, não havendo, assim, um padrão fixo de verdades (cf. Neurath, 1983: 181).

Assim, dado que não existe um mundo externo que possamos comparar às nossas afirmações para alcançar uma justificação sólida, resta alcançar a coerência com outras afirmações sobre p . Percebemos, assim, porque é que (RcT) é um fundamento importante do reducionismo para alcançar a unificação das ciências: quanto mais coerente for a teoria na relação com outras teorias, mais justificada ela estará, e, assim, mais verdadeira será. Alcançando a tradução das linguagens das diferentes ciências numa linguagem comum, essa conexão é mais fácil de ser alcançada entre diversas disciplinas.

É importante frisar que o foco na tese de que as diferentes ciências deveriam reduzir-se à linguagem da Física é somente algo historicamente contingente. Esta versão reducionista não defende que somente a Ciência Física poderia servir de linguagem comum às restantes disciplinas. O que interessava, antes, era que essa linguagem comum fosse universal com o objetivo de alcançar a unificação da Ciência. O facto de ser a linguagem da Física preterida por Carnap e Neurath seria um facto contingente e histórico do seu tempo, dado que esta disciplina parecia corresponder exatamente aos critérios de universalidade e confirmação por intersubjetividade que a Filosofia, por exemplo, não conseguiria alcançar dado o seu uso persistente de conceitos subjetivos. O reducionismo, nesta versão, é um reducionismo para uma linguagem comum, e não necessariamente para linguagem da Física.

Esta forma particular de reducionismo levanta uma incongruência difícil de ser ultrapassada devido à sua ligação com (tCV): esta teoria da verdade leva a paradoxos inconsistentes se se assumir a relação de coerência como consistência entre proposições. Veja-se o seguinte exemplo:

- O sujeito S_1 acredita na proposição p em consistência com um conjunto de proposições \mathcal{P}_S ;
- O sujeito S_2 acredita na proposição $\neg p$ em consistência com um conjunto de proposições $\neg \mathcal{P}_S$;

Ora, nesta situação, tanto S_1 como S_2 têm as suas crenças justificadas e por isso verdadeiras segundo (tCV). Ora, tal é inconsistente, porque apenas S_1 ou S_2 podem estar certos sobre o valor de verdade de uma proposição, e não ambos ao mesmo tempo (cf. Walker, 1989). Vejamos, assim, uma outra forma de ver o reducionismo.

III.1.2.2. Redução como Derivação (RcD)

Uma segunda forma de entender a redução aparece com Ernest Nagel e a sua influente obra *A Estrutura da Ciência*. Para Nagel, o relevante da redução não é tanto a tradução de termos noutros termos como defende (RcT) mas, antes, a ideia de que as leis da ciência que queremos reduzir sejam demonstradas como sendo uma consequência lógica dos pressupostos teóricos da ciência-base (cf. Nagel, 1961: 345-358).

Quando alcançamos essa demonstração, observamos que existe um número limitado de princípios-base que são necessários para explicar as verdades de ambas as ciências. O objetivo da redução por derivação é aproximar as ciências de uma teoria comum que possa integrar os vários domínios das ciências naturais que têm em comum um conjunto limitado de princípios (cf. Nagel, 1961: 336).

Para Nagel, há dois tipos de redução: as reduções homogêneas e as reduções heterogêneas. A diferença está no papel de tradução dos conceitos: na primeira versão, poderá de facto existir uma relação de tradução entre os conceitos da teoria reduzida e a teoria-base. Contudo, tal não é obrigatório como assume (RcT): nas reduções heterogêneas, a ciência a ser reduzida poderá incluir conceitos que não estão presentes na ciência-base (cf. Nagel, 1961: 342). Não é, assim, necessário traduzir sempre os conceitos de uma teoria para a outra.

Veja-se o seguinte exemplo: caso estejamos interessados em reduzir os conceitos mentalistas da Filosofia da Mente como “crença”, “mente” ou “dor” a uma teoria física, esta não terá qualquer conceito que traduza diretamente os conceitos da primeira teoria. Quando se dá este caso, é necessário adicionar pressupostos (*assumptions*) às leis da ciência-base (e.g. a teoria física) que indiquem as relações entre os termos mentalistas e os termos da ciência-base. A estes pressupostos chamaremos de “leis-ponte”, por influência de Carl Hempel, que irão permitir a derivação – e não a tradução – das leis da ciência a ser reduzida (e.g. os conceitos mentalistas) à ciência-base (e.g. conceitos da física) (cf. Nagel, 1961: 351-354).

Carl Hempel formalizou esta estrutura de pensamento explicitamente da seguinte forma. Assumindo que (B_1) e (B_2) são conceitos que constituem a linguagem da ciência-base e (T_1) e (T_2) são conceitos que constituem a linguagem da ciência que queremos reduzir:

(3.1) A ocorrência de (B_1) causa a ocorrência de (B_2) (uma lei na ciência-base);

(3.2) Se algo é (B_1) , então é (T_1) (lei-ponte);

(3.3) Se algo é (B_2), então é (T_2) (lei-ponte);

(3.4) Logo, a ocorrência de (T_1) causa a ocorrência de (T_2) (uma lei na ciência a reduzir) (cf. Hempel, 1966: 105).

Assim, obtemos uma lei da ciência que queremos reduzir às leis da ciência-base, alcançando uma derivação dos termos através das leis-ponte em vez de uma tradução entre os conceitos.

O tipo de leis-ponte que é necessário para alcançar uma derivação pode variar e é objeto de discussão intensa entre os filósofos da ciência. Nagel pensou inicialmente que bastaria que as leis-ponte tivessem uma forma condicional para que a derivação fosse bem-sucedida. Mas com a discussão da redução de Nagel, as formas bicondicionais ou de identidade têm vindo a ganhar mais relevo na literatura científica (e.g. Sklar, 1967) dado que as relações de identidade representam maior progresso no objetivo da unificação das ciências: apenas as leis-ponte como identidade seriam consideradas reduções genuínas.

Neste sentido aponta Jaegwon Kim ao afirmar que

É indiscutivelmente analítico que a redução deve simplificar; afinal de contas, as reduções devem reduzir [...] Nesse ponto, as leis-ponte [Algo é um T_1 se e só se for um B_1] estão em falta de variadas formas. Como [Algo é um T_1 se e só se for um B_1] deve ser uma lei contingente, os conceitos [T_1] e [B_1] permanecem distintos; portanto, as leis-ponte não produzem uma simplificação conceptual. Ademais, como temos apenas um “se e só se” contingente de propriedades de conexão entre [T_1] e [B_1], [T_1] e [B_1] permanecem propriedades distintas e não há simplificação ontológica [...] Se queremos simplificação ontológica das nossas reduções, é necessário encontrar uma maneira de aprimorar as leis-ponte em relações de identidades (Kim, 1998: 96-97).

Assim, segundo esta visão, apenas quando há uma relação de identidade é que alcançamos genuínas reduções. Mas fazemos notar que esta é uma visão radical da visão tradicional de Nagel que põe em primeiro lugar a simplificação ontológica, sendo a unificação da Ciência somente uma consequência secundária deste processo. Mas a perspectiva tradicional de Nagel era diferente: a unificação da Ciência deve ser o objetivo relevante da redução por derivação, bastando que as leis-ponte tenham forma de condicionais, não exigindo, assim, relações de identidade (cf. Ladyman, Ross & Spurrett, 2007: 49).

Com a influência do trabalho publicado por Thomas Kuhn nos anos 60, os filósofos da ciência deram cada vez mais importância à ideia de que os seus modelos de redução deveriam poder incorporar

exemplos da história da Ciência. Por exemplo, seguindo um exemplo do próprio Nagel, seria difícil perceber como a redução da termodinâmica à mecânica estatística poderia ocorrer aplicando o modelo de Nagel original.

Nesse sentido, Kenneth Schaffner (1967) levantou duas críticas ao modelo reducionista de Nagel:

(i) O modelo de Nagel de redução pressupõe que a dedutibilidade da teoria a ser reduzida à teoria-base em conjunção com o auxílio das leis-ponte ocorria diretamente. O problema é que, para que a derivação fosse realmente bem-sucedida, seria necessário corrigir ou modificar de variadas formas a primeira, dado que alguns pormenores da teoria a ser reduzida poderiam ser falsos. Seguindo Schaffner e o exemplo da teoria de Maxwell, podíamos compreender que um dos fundamentos centrais da teoria física ótica, a lei da intensidade de Fresnel, estava incorreta. Além de ser necessário modificar ou eliminar aspetos da primeira teoria, poderia dar-se o caso de se ter de incorporar ou adicionar factos sobre alguns dos fenómenos que a teoria procurava explicar e descrever (e.g. seguindo o exemplo da redução da física ótica ao eletromagnetismo, é necessário acrescentar a ideia de que a primeira apenas se aplica a uma frequência determinada, e não a todas). Assim, se queremos reduzir a física ótica à teoria de Maxwell, a derivação ocorre não da teoria física ótica tradicional, mas de uma versão corrigida e alterada da mesma (cf. Schaffner, 1967: 137-147);

(ii) O modelo de Nagel de redução assumia um fenómeno particular de nome “invariância conceptual” entre as duas teorias. Ora, de facto, aquando de uma redução, poderá acontecer que um conceito poderá alterar o seu significado ou ser entendido de uma nova perspetiva (cf. Capítulo 2) ou poderá, até, ser totalmente rejeitado pela teoria-base. Schaffner oferece o exemplo da evolução do conceito de “gene” da teoria genética de Mendel como sendo o resultado da redução a uma teoria bioquímica (Schaffner, 1967: 243). Assim, a teoria que alcançamos com a derivação poderá não ser constituída por *todos* os fenómenos ou leis da teoria original, sendo que a invariância conceptual poderá ser falsa;

Assim, nesta atualização do modelo nageliano, os vários exemplos da história da Ciência podem já ser integrados com maior facilidade teórica e passa a ter a seguinte formalização:

- (4.1) (T^*) é uma versão reconceptualizada e corrigida de T ;
- (4.2) A redução de uma teoria T a uma teoria B envolve uma versão T corrigida (T^*) ;
- (4.3) As leis-ponte não ligam os conceitos e fenômenos de T a B , mas de (T^*) a B ;
- (4.4) Assim, (T^*) é derivado de B em conjunção com as leis-ponte;
- (4.5) O mesmo acontece com B : além de (T^*) ser uma modificação de T , dado que (T^*) é derivado de B , B é modificado para (B^*) ;
- (4.6) Logo, (B^*) deriva (T^*) em conjunção com as leis-ponte.

O modelo de Schaffner, que se trata, no fundo, de uma atualização do modelo de Nagel, tem em comum com este a tese da (RcD) de que a redução interteórica ocorre da derivação de T para B , com o auxílio das leis-ponte que relacionam os conceitos e fenômenos (ou modificações ou eliminações) das duas teorias, e pode ser formulada da seguinte forma:

(RcD): A teoria (T_2) é reduzida à teoria (T_1) quando todas as leis de (T_1) forem derivadas das leis de (T_2) ;

Esta é uma tese que será relevante para a análise da tentativa de reduzir os termos mentalistas da Filosofia a conceitos de uma co-teoria da Neurociência e Psicologia. Mas antes de analisarmos este debate, vejamos uma terceira forma de entender o reducionismo.

III.1.2.3. Redução como Explicação (RcE)

Uma outra forma de considerar a redução é o modelo da redução como explicação que aparece no artigo “Da Redução”, em 1956, de Paul Oppenheim e John Kemeny que formalizaram a redução da seguinte forma. Uma teoria (T_1) é reduzida a uma teoria (T_2) em relação a um conjunto de dados observacionais (O) se:

- (5) Red (T_1, T_2) se e só se:
 - (5.1) Voc (T_2) contém termos que não estão Voc (T_1) ;
 - (5.2) Todos (O) de (T_2) estão em (T_1) ;
 - (5.3) (T_1) é tão sistematizada como (T_2) (cf. Oppenheim & Kemeny, 1956: 15);

Assim, a redução de uma teoria (T_1) a uma teoria (T_2) ocorre se o vocabulário (Voc) de (T_2) possuir termos na sua composição que não estão presentes no vocabulário de (T_1). Os dados observacionais (O) têm de ser explicados tanto por (T_2) como (T_1). Finalmente, a sistematização presente em (5.3) refere-se à capacidade de uma teoria de conseguir adaptar a sua complexidade à capacidade de prever e explicar mais dados observacionais (O).

Tal ponto é importante devido à motivação dos autores para alcançar reduções em Ciência. Nas palavras dos mesmos: “(...) o papel de uma teoria não é oferecer-nos mais factos, mas organizá-los num sistema que se possa gerir facilmente” (Oppenheim & Kemeny, 1956: 11). Nesse sentido, estes autores encontram-se na linha da unificação teórica da Ciência através da redução ou eliminação de teorias cujas observações (O) podem ser explicadas por uma teoria mais abrangente que explica igualmente essas observações.

É importante notar a diferença entre o modelo (RcE) e os modelos anteriores (RcT) e (RcD). Segundo os autores, o problema dos modelos anteriores é que procuram estabelecer uma ligação direta entre duas teorias, algo que é logicamente demasiado exigente. Na versão (RcE), essa ligação é, pelo contrário, indireta e sem tomar como fundamento central a tradução dos conceitos da teoria. Assim, o objetivo é explicar todos os factos sem que seja necessário um método de tradução direta – como acontece com (RcT) – ou de segunda ordem – como acontece com (RcD).

Neste novo modelo, os termos da teoria a reduzir não têm de estar conectados com os termos da teoria-base como defendiam os anteriores modelos, mas com os dados observacionais que, naturalmente, se conectam indiretamente aos termos da segunda teoria. Assim, a ligação entre os termos é logicamente menos exigente – e por isso mais simples – do que uma tradução total e completa (cf. Oppenheim & Kemeny, 1956: 16).

Podemos formalizar esta tese de (RcE) da seguinte forma:

(RcE): A teoria (T_2) é reduzida à teoria (T_1) quando todas as observações explicadas por (T_2) são igualmente explicadas por (T_1);

Deste modo, este modelo pretende incluir os anteriores modelos na sua composição, assumindo que os casos de tradução direta são somente casos especiais e raros e que a maior parte das reduções interteóricas acontecerão de forma explicitada em (4). Portanto, a redução terá como objetivo auxiliar

a eliminar teorias e conceitos que são supérfluos, procurando que a Ciência alcance níveis mais elevados de parcimónia teórica, objetivo que, se alcançado, terá como consequência a unificação da Ciência.

III.1.3. Redução em Patricia Churchland

Patricia Churchland irá usar o reducionismo para resolver o problema mente-corpo. Para a filósofa americana, a identificação das propriedades e conceitos de (T1) com (T2) não é necessária e é muitas vezes inútil. As reduções na história da Ciência não implicam somente reduções “suaves” (*smooth*) que envolvem cruzamentos de identificação entre teorias, mas, pelo contrário, reduções “bruscas” (*bumpy*) onde esses cruzamentos não são de todo necessários dado que são problemáticos por envolverem uma revisão da teoria inicial – como vimos em Nagel – e dos seus conceitos que podem ser eliminados²⁰ e não identificados para depois serem reduzidos. Segundo Churchland, é este o caso que acontece nas seguintes teorias pré-científicas, como por exemplo:

- A teoria do flogisto: desenvolvida por Georg Stahl, no século XVII, acreditava que todos os corpos combustíveis teriam uma matéria de nome “flogisto” que seria libertada para o ar durante os processos de combustão ou calcinação que seria depois absorvida pelas plantas;
- A teoria calórica: desenvolvida no século XVIII, defendia-se que o calor era um fluido invisível e sem odor que todos os corpos orgânicos teriam, cuja quantidade iria influenciar a quantidade de calor que esse corpo teria; quanto maior fosse a quantidade de fluido, maior seria a temperatura;
- A teoria dos demónios: tendo várias versões em diferentes períodos da história do pensamento, esta teoria acreditava que os problemas de doença mental se deviam à possessão demoníaca em pessoas saudáveis; tal teoria servia para explicar a esquizofrenia, epilepsia, síndrome de Tourette, distúrbio de dupla personalidade, entre outros problemas mentais.

Em todos estes casos, a teoria a reduzir eliminou conceitos centrais – a existência do flogisto, do fluido calórico, ou dos demónios – e nenhuma identificação interteorética foi alcançada ou necessária (cf. Churchland, 1981: 284). Este ponto é importante para Churchland, porque o reducionismo no debate

²⁰O primeiro autor a referir o termo “eliminativismo” é James Cornman no seu artigo de 1968, “On the Elimination of ‘Sensations’ and Sensations” (Cornman, 1968), embora a ideia, na sua base, já tivesse sido defendida por Broad no seu clássico *The Mind and its Place in Nature* (Broad, 1925: 607-611).

da mente e do cérebro dá-se, erroneamente, pela lupa das reduções suaves que requerem uma identificação particular de, por exemplo, estados mentais com estados cerebrais.

Ora, a filósofa quer evitar os constrangimentos filosóficos das teorias identidade-tipo ou identidade-espécime. Estas teorias de origem fisicalista defendem que tudo o que existe é algo físico, ou seja, tudo é extensão. Os fenómenos mentais, como pensamentos ou emoções, são idênticos a certos fenómenos físicos – a nossa neurofisiologia associada a x estado mental: ter um estado mental consiste em estar num estado físico particular.

Apliquemos esta distinção ao fenómeno mental: a teoria identidade-tipo (*type*) afirma que para cada tipo de fenómeno mental – sentir-se triste ou feliz, querer ou odiar algo – há um estado físico correspondente. Assim, por exemplo, as dores são idênticas à ativação de fibras C , a minha dor é idêntica à minha ativação das fibras C , e a dor de outra pessoa é idêntica à ativação das suas fibras C ; a mesma ativação mantém-se igual ontem, hoje e amanhã, o mesmo processo mental corresponde ao mesmo processo físico.

Por outro lado, os teóricos da identidade-espécime (*token*) defendem que a minha dor ontem foi definitivamente idêntica a um estado físico. E a minha dor hoje é definitivamente idêntica a um estado físico. Mas esses estados físicos poderiam ser todos diferentes: no primeiro caso pode ser um conjunto de neurónios x a disparar, enquanto que no segundo poderá ser outro conjunto y de neurónios a disparar.

A resposta identitária oferece um programa de investigação mais forte, dado que afirma que tipos específicos de estados físicos, por exemplo, uma libertação de endorfinas, é idêntica aos tipos de estados mentais associados a ela, por exemplo, sentir-se feliz, e que este é o caso para todos os seres humanos e animais não-humanos com uma neurobiologia complexa.

Ora, Churchland quer evitar que se pense o reducionismo desta forma tão restrita dado que esta assume que é realmente necessário que haja cruzamento de identificação entre conceitos e fenómenos das teorias em jogo – reduções do tipo suave. Mas, para a filósofa, tal pode ser perigoso para o próprio desenvolvimento das Neurociências:

No entanto, fazer essas identificações é um requisito que prejudica gravemente o desenvolvimento teórico da neurociência, pois implica que apenas um tipo de redução na extremidade retentiva do espectro [*sc.* reduções suaves] é válida e nenhuma dos estilos revisionários das relações redutivas [*sc.* reduções bruscas] será aceitável nem a eliminação será considerada sequer possível (Churchland, 1986: 284).

Ora, o objetivo da Abordagem do Reduccionismo será reduzir os conceitos mentalistas da Filosofia a conceitos de uma coteoria científica que una o conhecimento da Neurociência com a Psicologia Cognitiva que irá corrigir e modificar os vários aspetos e conceitos da teoria-base até poder absorver ou derivar a teoria a reduzir.

A própria história da Ciência parece estar preenchida de reduções suaves mas, para Churchland, tal é apenas uma impressão geral que não presta atenção aos pormenores dos aspetos formais. Veja-se o exemplo da “redução” da mecânica clássica de Newton (MC) à teoria da relatividade restrita de Einstein (TRr) oferecida pela filósofa.

Parece estarmos perante uma redução do lado retentivo do espectro das reduções dado que as duas teorias têm fenómenos e conceitos em comum, como a massa, o movimento, a aceleração, entre outros. Contudo, esta impressão é ilusória dado que as duas teorias são substancialmente diferentes: dá-se origem a um caso de eliminação, e não de redução propriamente dita, entre (TRr) e (MC).

Tomemos atenção ao conceito de massa em (MC) e em (TRr). $MC_{(massa)}$ é vista como uma propriedade intrínseca da natureza, não variando de observador para observador tendo um valor absoluto. Já $TRr_{(massa)}$ é totalmente diferente: o seu valor depende do observador, das condições em que a observação é efetuada e é, fundamentalmente, um valor relativo. Ora, para (TRr), $MC_{(massa)}$ não existe nem faz qualquer sentido. Já para (MC), $TRr_{(massa)}$ também não faz qualquer sentido. Assim, $MC_{(massa)}$ não pode ser identificado com $TRr_{(massa)}$. Antes, o segundo conceito substitui por completo o primeiro conceito na nova teoria, não através de uma redução suave, mas de uma eliminação.

O que foi reduzido, tecnicamente, foi um análogo de (MC), (MC*), em conjunção com pressupostos falsos e leis análogas – nomeadamente a tese de (MC) funciona somente em determinadas condições específicas e limitadas – que permite ser logicamente derivada de (TRr), de forma que conseguimos, com este passo, explicar o sucesso da teoria reduzida (dentro dos seus limites) e que a teoria-base (TRr) tem maior poder explicativo e preditivo como teoria científica do que (MC).

Neste exemplo, estamos perante um caso de redução interteorética bem-sucedida embora não tenha sido necessário alcançar relações de identidade entre (MC) e (TRr) dado que as duas teorias são, na sua essência, radicalmente diferentes (cf. Churchland, 1986: 286-288).

Um segundo exemplo que pode ser oferecido para demonstrar uma redução que leva à eliminação é a teoria física anterior à de Newton. Aristóteles tinha defendido que um objeto se mantém em movimento apenas se uma força for continuamente aplicada a esse objeto. Esta é uma descrição ordinária (*folk*) do movimento físico que as crianças podem compreender fazendo observações

rudimentares: por exemplo, empurrando um carrinho de brincar, esta irá notar que, se deixar de empurrar o carro, o seu movimento irá também parar.

Claro, esta prototeoria física estava profundamente equivocada tendo como contraexemplo disponível no tempo de Aristóteles o movimento de projétil: como se poderia explicar que um projétil continuaria em movimento depois de receber um impulso inicial? É esta visão errónea, mas de observação comum, que Newton e a sua primeira lei procurava destronar: um objeto em movimento retilíneo uniforme irá continuar uniformemente o seu movimento a não ser que uma força contrária interrompa o mesmo – é apenas necessária uma aceleração inicial para fazer mover um objeto, mas não uma força contínua. Claro, a nova teoria (MC) consegue explicar por que é que o carro empurrado pela criança irá parar: a força da gravidade e o facto de o atrito das rodas ser uma força contrária ao movimento inicial.

Ora, a teoria aristotélica não foi reduzida suavemente com ligações identitárias efetuadas entre esta teoria e (MC). Pelo contrário, a maior parte dos conceitos e leis da prototeoria física foram falsificados e substituídos por novos conceitos e leis de (MC) que, por sua vez, foi falsificada por (TRr) (cf. Churchland, 1986: 289-290).

Para Churchland, este segundo exemplo é útil porque mostra como é que as crenças comuns ou ordinárias (*folks*) sobre a natureza do movimento podem estar equivocadas mesmo que tal pareça ser claro ou óbvio para muitas pessoas intuitivamente,²¹ uma lição que a filósofa irá procurar aplicar ao problema entre os conceitos mentalistas e os conceitos neurocientíficos.

A autora de *Neurofilosofia* retira quatro lições relevantes para a discussão da abordagem redutiva:

- (i) Reduções de uma teoria podem levar à falsificação e conseqüente substituição de outra teoria;
- (ii) Por vezes, o que é substituído é uma teoria ordinária ou popular (*folk*) que pode ser baseada em observações (erradas);
- (iii) Apesar de uma teoria ordinária ou popular ser muito intuitiva e evidente, esta pode estar equivocada;
- (iv) Uma nova teoria, depois de se tornar familiar, pode ser aplicada causalmente como a teoria ordinária ou popular (cf. Churchland, 1986: 291).

²¹McCloskey (1983) investigou empiricamente que muitas pessoas com menos conhecimento ou informação são mais aristotélicas do que newtonianas em relação ao movimento, acreditam numa teoria física da intuição.

Igualmente importante é a ideia metafilosófica de que, com a redução de teorias por outras, as próprias questões centrais de investigação e conceitos poderão perder o seu sentido inicial: há, assim, uma reconfiguração do que a teoria precisa de explicar e descrever.

Para ilustrar este ponto relevante, a filósofa americana pede-nos para considerar o geocentrismo que defendia que o planeta Terra era o centro do universo. O Sol e a Lua giravam à volta da Terra uma vez por dia, e as estrelas estavam fixas numa esfera celestial que rodava pelo movimento do eixo da Terra. Neste sistema, os teóricos do geocentrismo procuravam responder à pergunta “O que faz a esfera celestial rodar?”, sendo que haveria diferentes hipóteses e respostas.

Com a refutação do modelo geocêntrico com Copérnico e depois Galileu, e com o aparecimento da mecânica clássica de Newton, a pergunta inicial deixava de fazer sentido dado que a nova teoria postulava que não existia qualquer esfera celestial. Assim, o que precisava de uma explicação inicialmente, com a mudança de teoria, deixa de precisar de explicação.

Ademais, deixamos de poder contar com observações que pareciam ser relevantes e que contavam como evidência para suportar o modelo ptolemaico, e passamos a poder explicar o aparente movimento das estrelas em referência à teoria newtoniana pelo movimento da Terra e não das estrelas. Finalmente, a própria evidência que era central no modelo anterior de que a Terra era o centro do universo e um corpo estático e estável é falsificada (cf. Churchland, 1986: 293).

Os “factos” da teoria reduzida deixam de precisar de uma explicação sendo transformados em pseudofactos com a teoria-base, que passa a poder explicar porque é que esses pseudofactos eram tidos como factos na teoria reduzida. O objetivo de Churchland ao analisar este exemplo da história da Ciência é tentar apontar para a tese de que, em relação ao problema mente-cérebro, o mesmo pode acontecer: o que é tomado como um facto evidente da Psicologia *Folk* pode deixar de precisar de ser explicado numa teoria futura da Neurociência que poderá, por sua vez, explicar porque é que esses fenómenos eram tidos como factos na teoria que foi alvo de redução.

Ora, que teoria deve ser alvo de redução para a abordagem reducionista? Vejamos a teoria que será tendencialmente substituída por uma teoria neurobiológica futura segundo os reducionistas.

III.1.4. Teoria da Psicologia Popular (*Folk*)

Depois de termos esclarecido os diferentes aspetos da redução, o próximo passo da Abordagem do Reduccionismo é perceber que teorias são candidatas para uma redução interteórica relevante – relembramos que uma redução não envolve primariamente a redução de um fenómeno (e.g. estado mental de tristeza) a outro fenómeno (e.g. estado cerebral associado à tristeza), mas, antes, a redução de uma teoria a outra. Que teoria é, assim sendo, candidata a ser reduzida pela teoria neurobiológica?

Wilfrid Sellars identificou uma estrutura padrão de crenças, desejos, dores, sonhos, etc., que formam uma teoria que usamos para descrever algo. Esta teoria, de nome “Psicologia Popular (*Folk*)” (PF), postula várias entidades – estados mentais intencionais – que se relacionam entre si e com o comportamento humano. Fazemos uso desta teoria para prever o comportamento dos outros e para descrever e classificar os nossos próprios estados internos. Nas palavras da autora:

A psicologia popular é a psicologia do senso comum – a tradição psicológica em virtude da qual explicamos o comportamento como o resultado de crenças, desejos, percepções, expectativas, objetivos, sensações e assim por diante. É uma teoria cujas generalizações conectam estados mentais a outros estados mentais [...] Essas generalizações caseiras são o que fornecem a caracterização dos estados e processos mentais (Churchland, 1986: 299)

Esta teoria parece não ser reconhecida por nós como tal: os seus conceitos são tão comuns que não compreendemos que as nossas observações já estão dependentes ou sujeitas a uma teoria específica. A (PF) é uma teoria empírica e as suas entidades são teoréticas. Como todas as teorias empíricas, (PF) é, em princípio, falsificável.

Assim, a (PF) é constituída por três ideias fundamentais:

- (6.1) A capacidade de atribuir estados mentais aos seres humanos (e animais não-humanos);
- (6.2) A capacidade de explicar o comportamento dos seres humanos (e animais não-humanos) em referência aos seus estados mentais;
- (6.3) A capacidade de prever o comportamento humano (e animais não-humanos) em diferentes circunstâncias (cf. Stich & Ravenscroft, 1992; Stich & Nichols, 2003).

É importante perceber por que razão os defensores de (AR) insistem na ideia de que, apesar de os conceitos mentalistas serem tão imediatos e óbvios, tal não significa que não façam parte de um esquema conceptual ou de uma teoria constituída por esses fenómenos mentais. Como vimos, o reducionista precisa de (i) uma teoria (ii) que possa ser corrigida e revisionada (iii) para depois ser reduzida a uma teoria-base ou, no caso extremo, eliminada. A (PF) terá esse papel: será esta teoria que será reduzida a uma teoria neurobiológica (futura).

Como fizemos notar anteriormente, poderá dar-se o caso de a nova teoria ser radicalmente diferente da (PF), cujos conceitos e fenómenos podem ser totalmente eliminados de forma análoga ao conceito de massa na mecânica clássica. Como também vimos, questões que parecem fundamentais (e.g. o problema mente-cérebro) podem ser totalmente dissolvidas com a nova teoria que terá outras questões a necessitar de explicação.

Sellars procura explicar alegoricamente como é que a teoria da (PF) chegou até aos nossos dias e por que razão é empregada por todas as pessoas falantes. Segundo o filósofo, haveria um primeiro estágio onde os seres humanos apenas usariam uma linguagem de Ryle baseada somente em aspetos descritivos das propriedades públicas localizadas no espaço e no tempo.

Neste primeiro estágio, não há conceitos, como crenças e desejos, e as atitudes proposicionais não existem (cf. Sellars, 1956: 309). No entanto, no segundo estágio, um indivíduo [Jones] cria uma teoria para tentar explicar o comportamento dos seus semelhantes. Nesta fase, a teoria permite perceber que o comportamento verbal público é o culminar de um processo que começa com um “discurso interno” ou privado (cf. Sellars, 1956: 317-318).

Neste estágio, a teoria apenas se aplica aos seus semelhantes: apenas estes têm crenças, desejos, etc. Contudo, na terceira fase do mito, os indivíduos aprendem a aplicar a si mesmos a teoria que anteriormente aplicavam aos outros. Eis como Sellars encerra o mito:

(...) Uma vez que Jones desenvolveu a teoria de que o comportamento verbal manifesto é a expressão de pensamentos, e tendo ensinado aos seus compatriotas a fazer uso da teoria na interpretação do comportamento do outro, foi necessário somente um pequeno passo para usar essa linguagem na autodescrição. Assim, quando Tom, observando Dick, possui evidências comportamentais que justificam o uso da proposição (na linguagem da teoria) “Dick está a pensar em ‘*p*’”, Dick está a utilizar a mesma evidência comportamental, podemos afirmar na linguagem da teoria, “estou a pensar em ‘*p*’”. E agora acontece que – precisa mesmo de ser? – Dick pode ser treinado para fornecer autodescrições razoavelmente fidedignas fazendo uso da linguagem da teoria, sem ter de observar o seu comportamento.

Jones faz isso, aproximadamente, aplaudindo as declarações de Dick de “estou a pensar em ‘*p*’” quando a evidência comportamental apoia fortemente a afirmação teórica “Dick está a pensar em ‘*p*’”; e franzindo a testa com declarações de “eu estou a pensar em ‘*p*’”, quando a evidência não sustenta essa afirmação teórica. Os nossos ancestrais começam a falar do acesso privilegiado que cada um de nós tem aos seus próprios pensamentos. O que começou como uma linguagem de uso puramente teórico ganhou o papel de reportar (Sellars, 1956: 320).

Assim, o importante desta história alegórica é que, no mito de Sellars, há uma mudança fundamental que ocorre entre os segundo e terceiro estágios: a forma de expressão “eu estou a pensar em ‘*p*’” – que era um postulado teórico de uma teoria – passa a ter um papel de reportar o mental sem necessitar de fazer inferências diretas do comportamento do outro. Além desta diferença, os indivíduos deixam de perceber que algo que era visto como um postulado teórico passa a parecer imediato e infalível para o sujeito.

Portanto, fazemos uso da (PF) e dos seus conceitos como “desejar”, “acreditar”, “pensar”, entre outros, para nos descrever a nós mesmos e aos outros. E grande parte das vezes aceitamos essas descrições como verdadeiras e infalíveis. Além disso, além de utilizarmos esta teoria tacitamente para nos descrever a nós mesmos e aos outros, fazemos uso da mesma para explicar o comportamento em geral das pessoas. Finalmente, usamos esta teoria para prever e antecipar o comportamento dos outros (cf. Churchland, 1994: 308).

São estas componentes da teoria que fazem com que o reducionista acredite que a (PF) é a teoria candidata a ser reduzida. Eis uma possível formulação do (7) argumento anti-Psicologia Popular (*Folk*):

- (7.1) As atitudes proposicionais (crenças, desejos, pensamentos, etc.) são postulados teóricos da (PF);
- (7.2) A (PF) é falsa;
- (7.3) Os postulados teóricos da (PF) não existem;
- (7.4) Logo, as atitudes proposicionais não existem.²²

²² Alguns autores põem em causa a validade deste argumento, dado que defendem que podemos defender a existência das atitudes proposicionais sem ter que defender a Psicologia Popular (*folk*) (cf. Kitcher, 1984; Von Eckardt, 1994).

Mas por que é que devemos acreditar na (AR) e aceitar que (PF) é falsa e deve ser substituída por outra teoria de carácter neurobiológico? Patrícia Churchland, defensora de (7), oferece várias razões a favor da premissa (7.2) do nosso argumento:

(a) (PF) tem graves deficiências explanatórias: o seu escopo é demasiado limitado e parece ser apenas aplicável a linguagens comuns (*folk*) humanas; tal aponta para o carácter superficial e incompleto desta teoria;

(b) (PF) tem um pobre desenvolvimento, marcado pela estagnação e a falta de progresso, dado que a mesma teoria se mantém inalterável desde a sua origem;²³

(c) (PF) é incompatível com a visão do mundo científico emergente: a ciência contemporânea tem divulgado inúmeras descobertas integrando-as numa poderosa imagem do nosso desenvolvimento. A (PF) parece não fazer parte desta imagem, dado que a Neurociência consegue explicar muito da cognição humana sem apelar a elementos da (PF);²⁴

(d) Indução histórica: Teorias do tipo “*folk*” têm uma fraca existência histórica – a Cosmologia *folk*, a Física *folk*, a teoria da possessão demoníaca: todas elas estão atualmente refutadas e eliminadas; o mesmo deverá acontecer à (PF) (cf. Churchland, 1981).

Assim, Churchland e os reducionistas em geral consideram que há razões suficientes para acreditar que a (PF) será eliminada muito brevemente, assim que a Neurociência atingir a maturidade.

Outra tese fundamental é, como vimos, a tese negativa contra o valor da intuição. Para os reducionistas, devemos ser céticos acerca das evidências que podemos alcançar através da introspeção dado que, se de facto a Psicologia *Folk* é real, este mecanismo estará profundamente enviesado pelos seus conceitos. Assim, para Churchland, alcançar determinadas crenças através da introspeção será como observar demónios, flogísticos ou esferas celestiais – embora nos pareça clara e realista a sua existência, devemos olhar com ceticismo para este tipo de evidências.

²³ Podemos, claro, responder que a estagnação é sinal que está correta e é eficaz e que, precisamente por esse motivo, não precisa de ser modificada. Ou, como Hacker, que não há progresso porque apenas teorias (reais) podem ter progresso e a (PF) não é uma teoria.

²⁴ Tal pode ser visto como falso: muita investigação contemporânea – em Psicologia, História, Teoria Moral, Criminologia, etc. – usa a (PF) na sua visão do mundo e tem sido útil na própria investigação científica, como argumentado por Horgan e Woodward (1985) ou Greenwood (1991).

Esta tese parece ter algum suporte empírico baseado em várias investigações que levantam a suspeita sobre a razoabilidade do conceito de introspeção. Por exemplo, Nisbett e Wilson demonstraram que a introspeção falha, algumas vezes, em três ocasiões fundamentais:

- (i) Os sujeitos desconhecem a existência de um estímulo que influenciou a reação;
- (ii) Os sujeitos desconhecem a existência de uma reação;
- (iii) Os sujeitos desconhecem que um estímulo afetou efetivamente uma reação (cf. Nisbett & Wilson, 1997).

É baseando-se nestas (e outras) evidências que um reducionista assume uma posição anti-introspeção em relação às atitudes proposicionais e ao seu valor epistémico e cognitivo.

Vejamos, de seguida, a posição de John Bickle em relação ao reducionismo.

III.1.5. Redução em John Bickle

John Bickle é outro defensor da Abordagem do Reduccionismo (AR), embora a sua posição possua algumas diferenças em comparação com a forma de reducionismo que analisámos na anterior secção. Para Bickle, o papel da (PF) é importante dado que esta é a teoria candidata a ser reduzida por uma teoria neurobiológica do futuro. Contudo, o filósofo do Mississippi difere de Churchland num aspeto interessante. Para o primeiro, não importa tanto se a (PF) é de facto uma teoria – uma crítica que analisaremos de seguida – mas, antes, que possamos construir uma conceção popular (*folk*) do mental:

Não importa se [a psicologia *folk*] está realmente “armazenada na cabeça” como uma teoria [...] ou se constitui mais do que apenas uma teoria (ou seja, desempenha outros papéis além de gerar previsões e explicações sobre o comportamento). Tudo o que importa [...] é que a psicologia popular *possa ser tratada como* a teoria à qual pertence a nossa ontologia do senso comum (Bickle, 1998: 43).

Assim, Bickle procura defender uma posição minimalista que não requer uma reivindicação epistémica tão exigente quanto à constituição da (PF) como uma abordagem teórica, algo que será motivo de discussão intensa, como veremos.

Da mesma forma que Churchland, Bickle recusa os argumentos da linguagem ordinária ou comum dado que a terminologia e as explicações neurobiológicas parecem ser claramente incompatíveis com o discurso do dia-a-dia sobre o mental e os estados mentais (cf. Bickle, 1998: 44). Nas palavras de Jerome Shaffer:

O facto de não fazer sentido falar de eventos mentais que ocorrem em algum momento no corpo resulta numa teoria da identidade que não pode ser verdadeira. Tal ocorre porque os eventos físicos correspondentes ocorrem em algum momento do corpo e, se eventos físicos são idênticos aos eventos mentais, então esses eventos mentais devem acontecer no mesmo ponto do corpo. Mas esses eventos mentais não acontecem em nenhuma parte do corpo, porque qualquer afirmação de que aconteçam aqui ou ali não teria sentido. Portanto, os eventos mentais não podem atender às condições de coexistência no espaço e, portanto, não podem ser idênticos aos eventos físicos (Shaffer: 1968: 48-49).

Contudo, para Bickle, esta posição comete uma óbvia petição de princípio, dado que o que está em causa é exatamente a natureza dessas entidades teóricas da Psicologia Popular (*Folk*), isto é, se de facto esta teoria – ou algo que se comporta como uma teoria – for reduzida a uma teoria neurobiológica do futuro, então já não haverá nenhuma lacuna como Shaffer sugere.

Para Bickle, o problema tradicional mente-corpo pode ser substituído por outro problema mais útil: a reformulação interteórica do problema mente-corpo. Esta reformulação é útil dado que pode integrar todas as soluções tradicionais ao problema se consideradas como análogos preditivos do sucesso explicativo ou da impossibilidade de redução de todas ou algumas componentes da Psicologia Popular (*Folk*) (cf. Bickle, 1998: 46).

Veja-se, por exemplo, a reformulação interteórica do dualismo de propriedades. Embora a teoria eletromagnética não seja reduzível à mecânica quântica, esta mantém-se como um ramo da Física: o facto de esta ser irreduzível não serve para demonstrar a tese de que existem objetos que não são físicos. Antes, pelo contrário, somente demonstra que a mecânica quântica não esgota as propriedades físicas do universo, i.e., além da massa, velocidade ou gravidade, existe uma propriedade adicional das partículas, nomeadamente a carga elétrica, que deve ser tomada em consideração.

Ora, o mesmo poderá acontecer em relação aos postulados da Psicologia Popular (*Folk*): o facto de alguns dos elementos desta teoria não serem redutíveis às leis da Física (tese anti-reducionista) não demonstra por si só que esses elementos não são, na sua substância, não físicos, mas, antes, que possuem propriedades não físicas de objetos físicos (cf. Bickle, 1998: 46).

Um segundo exemplo que pode ser oferecido é a variação funcionalista da perspectiva da reformulação interteórica: pode ser defendido que a Psicologia Popular (*Folk*) poderá ser reduzida suavemente não à ciência da Física, mas à Psicologia Computacional. Nas palavras de Jerry Fodor:

Fomos conduzidos ao funcionalismo (e, portanto, à autonomia da Psicologia) pela suspeita de que existem generalizações empíricas sobre os estados mentais que não podem ser formuladas no vocabulário de teorias neurológicas ou físicas: supomos que a neurologia e a física não fornecem predicados-tipo projetáveis que subtendam os domínios dessas generalizações (Fodor, 1981: 25).

Assim, a ideia de Fodor é considerar a tese anti-reducionista através de uma defesa da redução interteórica da Psicologia Popular (*Folk*) a uma teoria da Psicologia Computacional, dado que, em comum, existem estruturas básicas e a continuação da postulação das atitudes proposicionais (cf. Bickle, 1998: 48).

Pode-se, também, defender uma redução relativamente brusca da Psicologia Popular (*Folk*) a uma teoria futura da Psicologia Computacional. Por exemplo, o trabalho de Stephen Stich (1983) pode ser interpretado neste sentido: optando por uma correção exaustiva da teoria a ser reduzida (PF), defendendo uma perspectiva causal estrita que permite, por um lado, capturar as generalizações comportamentais que cada indivíduo possui mas, por outro lado, não necessitando de aplicar a ontologia específica da Psicologia Popular (*Folk*), poderá alcançar-se uma redução mais do lado direito (cf. Figura 3.1) do espectro das reduções interteóricas (cf. Bickle, 1998: 49)

Finalmente, pela reformulação interteórica do problema mente-corpo, podemos também interpretar o eliminativismo materialista como a sugestão de que, por exemplo, o conceito de “dor” não terá qualquer sinónimo ou identificador co-extensivo de um conceito da linguagem popular ou *folk*. Nas palavras de Feyerabend, um dos primeiros defensores do eliminativismo:

Toda a discussão interessante, isto é, toda a discussão que leva a um avanço do conhecimento, termina numa situação em que uma mudança decisiva de significado ocorra [...] uma nova teoria das dores não muda a(s) [natureza das] dores [...] ela muda o significado de ‘Estou em dor’ (Feyerabend, 1962: 58).

Assim, na reformulação em análise, a redução da Psicologia Popular (*Folk*) cairá na versão radical e extrema de redução brusca, levando não à correção de conceitos centrais da teoria reduzida, mas à sua eliminação.

Uma diferença que Bickle considera fundamental entre a sua abordagem reducionista e a abordagem reducionista de Churchland é o facto de a segunda não defender a relevância de um critério formal para avaliar o espectro das reduções interteóricas, isto é, um critério que ajude a determinar a diferença exata e específica entre reduções suaves ou bruscas, entre ligações identitárias ou eliminações.

Todavia, para o filósofo, este critério é muito relevante e não pode ser ignorado em nome de critérios sociológicos e pragmáticos da ciência. É isto que parece sugerir Churchland ao afirmar que a procura por um critério formal é irrelevante: dado que diferenciamos as reduções através de “uma variedade de considerações sociais e pragmáticas” que podem incluir “o capricho dos investigadores principais da área, o desejo de preservar ou romper com hábitos do passado, as oportunidades para divulgar a teoria, obter financiamento ou discípulos” (Churchland, 1986: 283).

Assim, para Churchland, estes fatores meramente descritivos influenciam demasiado a procura por um critério formal de distinção das reduções, pelo que este não deve ser relevante para um investigador que procura reduzir uma teoria a outra.

Porém, para Bickle, esta perspectiva descritiva ou pragmática não pode justificar a normatividade do problema da redução por meros aspetos sociológicos da ciência, em contraposição com as consequências ontológicas do problema em questão (cf. Bickle, 1998: 54). Deste modo, irá tentar apresentar critérios formais para distinguirmos o tipo de reduções, as suas consequências ontológicas e metafisológicas, como veremos de seguida.

III.1.6. A Teoria da Redução Interteórica

Bickle irá avançar com uma nova proposta de redução que, como veremos, poderá ser interpretada como uma Redução como Explicação (cf. II.1.2.3.). Vejamos com mais detalhe os pressupostos filosóficos da visão que Bickle detém da redução.

O primeiro passo para conceber o reducionismo é utilizar a perspectiva do estruturalismo. Esta perspectiva não deve ser confundida com o estruturalismo francês da Filosofia Continental: trata-se, antes, de uma abordagem da Filosofia da Ciência analítica que será útil para fundamentar a natureza da

redução. Especificamente, Bickle fará uso da abordagem semântica da Filosofia da Ciência, que defende dois pressupostos fundamentais (cf. Mormann, 1991):

- (i) As teorias científicas não devem ser concebidas primariamente como entidades linguísticas (e.g. um conjunto de axiomas, proposições ou provas), mas como modelos;
- (ii) A ferramenta apropriada para uma explicação formal da estrutura de (i) não é a lógica de primeira ordem (como concebido pelos positivistas lógicos) nem a metamatemática, mas, antes, a matemática (cf. Bickle, 1998: 59).

Há, contudo, uma intensa discussão em relação a que tipo de (ii) deve ser adotada: alguns autores defendem a teoria dos conjuntos (cf. Suppes, 1956; Sneed, 1975); outros defendem a abordagem dos espaços de estados (*state-space approach*) (cf. van Fraassen, 1972; Suppe, 1974), ou ainda abordagens computacionais da Ciência Cognitiva (cf. Giere, 1988).

Seja como for, o primeiro autor a defender explicitamente a abordagem semântica é Suppes, que suspeitava da possibilidade de axiomatização das teorias científicas com instrumentos de lógica de primeira ordem por não ser “nem natural, nem simples formalizá-las em lógica de primeira ordem” (Suppes, 1956: 248), além de ser “totalmente impraticável” (Suppes, 1956: 250).

Mas o que significa a tese apresentada em (i)? Por que devem ser as teorias científicas vistas como modelos? Para Bickle, os modelos não são representações das entidades que a teoria procura clarificar, mas, antes, são essas mesmas entidades. Nas palavras dos autores que melhor formalizaram esta ideia:

Em vez de se afirmar que determinadas equações são um modelo de fenômenos subatômicos ou económicos, propomos defender que os fenômenos subatômicos ou económicos são os modelos da teoria representada por essas equações (Balzer, Moulines & Sneed, 1987: 2).

Mais relevante, os modelos não devem ser vistos como as unidades fundamentais da análise estruturalista de uma teoria particular, mas, antes, como um subconjunto de “modelos potenciais” da teoria em questão (cf. Bickle, 1998: 63).

O próximo passo de Bickle é conjugar esta perspectiva estruturalista com a ideia relevante de correção da teoria reduzida avançada por Clifford Hooker, que tinha avançado com um modelo de

correção baseado na intuição de Schaffner de que a forma como o modelo de Nagel é concebido é errónea. Contudo, o filósofo americano identifica vários problemas na perspectiva de Hooker sobre a redução que tentará resolver:

- (a) Hooker nunca chega a aplicar quantitativamente e ao detalhe a sua redução a casos concretos da história da Ciência;
- (b) Hooker não oferece uma formalização precisa do conceito de “relação análoga” entre a teoria reduzida e a teoria-base;
- (c) Hooker não oferece nenhum critério formal especializado para distinguir reduções suaves de reduções bruscas (cf. Hooker, 1981: 223).

Ora, Bickle quer resolver os problemas (a) – (c). Embora os detalhes técnicos da abordagem sejam complexos e fora do âmbito deste trabalho, o relevante da sua perspectiva pode ser caracterizado pela seguinte intuição pressuposta: em vez de se olhar para o reducionismo em termos de derivações sintáticas (como fazia a tradição anterior), este novo reducionismo interpreta a relação entre teorias como uma construção, isto é, um modelo potencial da estrutura teórica dos conjuntos entre a teoria reduzida e a teoria-base.

Aplicado diretamente ao problema central da Filosofia da Mente, o problema mente-corpo, o reducionismo de Bickle pressupõe uma capacidade de previsão – como a abordagem de P. Churchland – de que, à medida que a Neurociência se desenvolve e atinge teorias mais amadurecidas, a teoria-base ganhará cada vez mais capacidade de reduzir os elementos – quer pela redução suave, ou brusca – da Psicologia Popular (*Folk*).

Bickle não tem dúvidas que tal será o caso utilizando um argumento indutivo por analogia: já existem evidências de conceitos que foram reduzidos de forma bem-sucedida na Neurociência contemporânea. O filósofo oferece o exemplo da aprendizagem associativa (*associative learning*). Segundo o reducionista, este fenómeno psicológico prova que já existem reduções a teorias neurocientíficas de teorias psicológicas-cognitivas.

Mesmo a forma mais simplificada de aprendizagem associativa (i.e. o condicionamento de Pavlov) faz uso de conceitos como representação e computação – em vez de se referirem ao conteúdo – para explicar o fenómeno em questão (cf. Bickle, 1998: 166-175). Bickle quer demonstrar dois aspetos importantes: que as teorias de aprendizagem associativa fazem uso de recursos cognitivos (citando os

trabalhos de Rescorla e Wagner, 1972) para depois mostrar de que forma é que essas teorias cognitivistas são reduzidas a teorias neurobiológicas.

Para alcançar este segundo passo, Bickle utiliza a investigação científica de Eric Kandel (1984) que afirmava: “O nosso objetivo é sugerir como a psicologia cognitiva pode começar a convergir com a neurobiologia para gerar uma nova perspectiva no estudo da aprendizagem” (Kandel, 1984: 376), afirmando no final do estudo que “tentámos fornecer versões neuronais dos modelos de condicionamento de Rescorla e Wagner” (Kandel, 1984: 388).

Para Bickle, este é um caso de redução bem-sucedida:

Em termos da minha teoria da redução, os conjuntos empíricos de base da teoria de Rescorla e Wagner [RW] relacionam-se diretamente, embora de forma heterogénea, com os conjuntos e as relações da teoria de Hawkins e Kandel [HK]. As ontologias das duas teorias são sobre os mesmos fenómenos – os sistemas do mundo real e as suas aplicações empíricas sobrepõem-se – embora eles descrevam esses fenómenos em diferentes níveis teóricos. O que temos na redução da teoria de Rescorla e Wagner à teoria de Hawkins e Kandel é uma redução interteórica suave (...): uma redução que justifica a ontologia da teoria reduzida em virtude da estruturação de seus elementos básicos amorfos e especificados funcionalmente nos componentes subjacentes e na dinâmica desses tipos de nível inferior (Bickle, 1998: 179).

Mas por que razão Bickle considera que este primeiro caso serve como evidência de uma redução do tipo suave? O filósofo considera os critérios que apontámos anteriormente: a [HK] não faz qualquer tipo de correção a [RW]; pelo contrário, [HK] aceita a formalização matemática de [RW], sendo que essa formalização encaixa perfeitamente os dados e evidências experimentais obtidos por [HK]. Por isso mesmo, estamos perante um caso de redução suave que conecta os aspetos cognitivos aos aspetos neurofisiológicos.

Com esta nova abordagem da redução interteórica, Bickle pensa ter alcançado algo que Churchland – tal como, anteriormente, Hooker – pensava ser demasiado complexo: uma metodologia objetiva²⁵ que consegue dar uma orientação precisa sobre a localização específica de uma redução no espectro das reduções interteóricas (Figura 3.1):

²⁵ Alguns autores rejeitam que Bickle tenha sido bem-sucedido. Por exemplo, João Fonseca afirma que “[Bickle não consegue] fornecer um procedimento efetivo que localize no espectro de uma forma absoluta uma certa redução arbitrária. Essa localização só pode ser estabelecida de um modo paroquial e interno e não de um modo absoluto e externo” (Fonseca, 2006: 129-130).

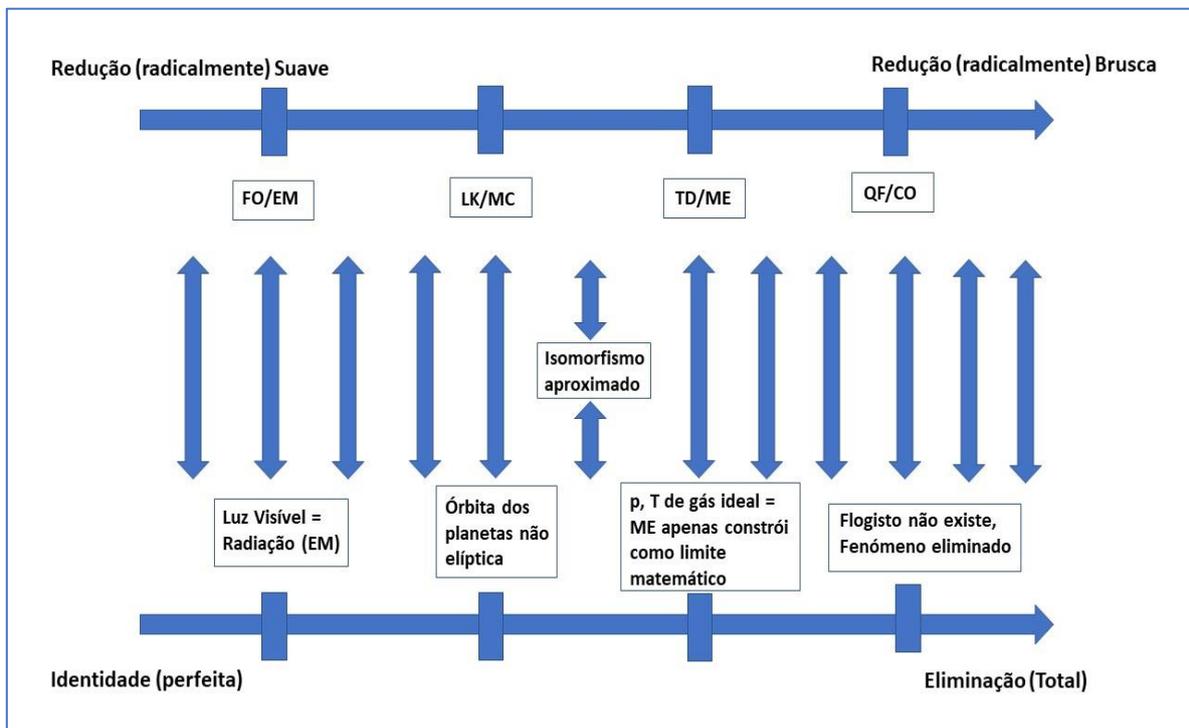


Figura 3.1. Seta de cima: o espectro das reduções interteóricas. Alguns casos históricos de acordo com a quantidade de correções da teoria reduzida. Seta de baixo: o espectro ontológico. Alguns casos históricos de acordo com as consequências ontológicas, desde identidades suaves a eliminações. FO: Física Ótica (teoria corpuscular da luz); EM: Eletromagnetismo de Maxwell; LK: Leis de Kepler; MC: Mecânica Clássica; TD: Termodinâmica Clássica; ME: Mecânica Estatística; TF: Teoria do Flogisto; TCO: Teoria da Combustão (Oxigênio) (traduzido de Bickle, 1996).

Desta forma, Bickle pensa ter conseguido resolver o desafio dos críticos do reducionismo a que chama de “desafio de ‘demonstra ou silencia-te’” (*put up or shut up challenge*): este desafio passaria pela ideia de que podemos aceitar de forma mais ou menos concreta que o reducionismo pode ser alcançado nas ciências da Física, tal como demonstrado amplamente por Churchland e Bickle em variadas ocasiões.

Contudo, o argumento continua, tal não demonstra que o mesmo possa ser alcançado nas ciências do cérebro: a redução da Psicologia Popular (*Folk*) a uma teoria neurofisiológica parece ser mais complexa do que a redução da mecânica clássica à teoria de Einstein, por exemplo.

Bickle procura contradizer este argumento anti-reducionista enfrentando o desafio de provar (*put up*) reduções alcançadas na Neurociência, apresentando para tal o fenómeno da aprendizagem

associativa (cf. Bickle, 1998: 166-174). Ademais, Bickle irá continuar ao longo dos anos a apresentar variadas tentativas de redução de fenômenos mentais a teorias neurocientíficas, como seja o caso da consolidação da memória (cf. Bickle, 2003: 46-51), da consciência (cf. Bickle, 2003: 163-212) ou ainda da aprendizagem espacial (cf. Bickle *et al.*, 130-137).

Portanto, Bickle procura demonstrar que a redução interteórica (especificamente da Psicologia à Neurociência) envolve uma metodologia que pode ser simples, embora os detalhes sejam deveras complexos: “intervir neuronalmente e rastrear os efeitos no comportamento” (Bickle, 2006: 425). Pragmaticamente, alcançamos uma redução quando um cientista consegue intervir no nível neuronal ou fisiológico de certo fenômeno mental conseguindo detectar objetivamente alterações neurofisiológicas, num primeiro nível, e alterações comportamentais, num segundo nível:

Quando essa estratégia é bem-sucedida, os eventos celulares ou moleculares em neurónios específicos nos quais as experiências intervêm [...] explicam diretamente os dados comportamentais. Essas explicações deixaram de lado as explicações intervenientes, incluindo o psicológico, o processamento cognitivo de informações, e até o neurocientífico-cognitivo [...] esses sucessos equivalem a reduções da mente a componentes moleculares nos neurónios [...] (Bickle, 2006: 426).

Podemos, assim, concluir que a redução interteórica proposta por este autor é um tipo de Redução como Explicação (RcE) (cf. III.1.2.3.): alcançamos uma redução quando conseguimos uma teoria neurofisiológica que consegue explicar os dados observacionais (O) da teoria reduzida – a Psicologia Popular (*Folk*).

Qual é, então, a diferença mais substancial entre a abordagem reducionista de Bickle e a de Churchland, dado que ambos defendem o reducionismo como instrumento para resolver o problema mente-corpo? A concepção de Bickle, tal como a de Churchland, aceita que terá de haver uma substancial mudança conceptual entre as teorias, colocando de parte relações de identidade isomórficas entre todos os postulados da Psicologia Popular (*Folk*) e os postulados de neurofisiologia – neste ponto, ambos os reducionistas estão de acordo.

Porém, Bickle rejeita a versão eliminativista radical de Churchland que considera que o destino de *todos* – o quantificador universal é o ponto central desta tese – os conceitos da teoria mentalista será o mesmo, em termos ontológicos, que o flogisto ou o fluido calórico: antes, Bickle prefere assumir uma posição mais rigorosa e menos radical, substituindo os conceitos representacionais da Psicologia

Popular (*Folk*) (e.g. as atitudes proposicionais) pelos conceitos representacionais de uma teoria cognitiva-neurofisiológica (em vez de negá-los totalmente, como faz Churchland) (cf. Bickle, 1998: 211) para depois alcançar reduções tanto do tipo suaves, como bruscas.

Numa analogia interessante, Bickle refere que a teoria da relatividade restrita de Einstein continua a usar os conceitos de massa, velocidade, gravidade, etc., embora esses conceitos sejam diferentes da sua versão clássica apresentada na mecânica clássica de Newton. Deste modo, encerra a sua obra afirmando: “Identidades estritas entre entidades e propriedades, não. Eliminação total das diferentes versões do fluido calórico ou flogisto, novamente não” (Bickle, 1998: 211). Deste modo, fica claro a diferença substancial entre as duas abordagens reducionistas dos dois autores analisados.

III.2. Reflexão Crítica:

Depois de termos refletido sobre as várias dimensões da Abordagem do Reduccionismo (AR), focaremos agora a nossa atenção sobre as críticas e problemas principais desta abordagem metodológica.

III.2.1. A Objeção da “Não-Teoria”:

Uma das formas de rebater esta proposta reducionista é defender, como sugerimos anteriormente, que a Psicologia Popular (*Folk*) não é, efetivamente, uma teoria (cf. Ratcliffe, 2006). Relembremos que esta tese é fundamental para que ocorra realmente uma redução interteórica que leve à eliminação.

Ora, esta tese pode de facto estar equivocada: pode ser argumentado que o uso de atitudes proposicionais como crenças ou desejos, aplicadas a nós e aos outros, não se deve a um domínio de uma teoria – a Psicologia Popular (*Folk*) – sobre essas atitudes proposicionais, mas, antes, à simulação desses estados mentais (cf. Gordon, 1986; 1992; Goldman, 1992).

Veja-se o seguinte exemplo: quando observamos um colega de trabalho a dirigir-se para a fila do bar da universidade, não consultamos uma teoria para chegar à conclusão que este tem fome e por isso deseja comer, mas, diferentemente, colocamo-nos nos “sapatos” desse colega simulando o que estaríamos a desejar quando nos dirigíssemos para a fila do bar, descobrindo, assim, o estado mental do colega – tem fome.

Assim, esta crítica propõe a tese fundamental de que o ser humano não explica ou prevê o comportamento dos outros indivíduos apelando a uma teoria particular, mas, numa estratégia alternativa, fá-lo simulando como é que cada um agiria numa situação contextual semelhante.

A crítica segue afirmando que a forma como prevemos ou explicamos o comportamento alheio faz uso dos mesmos mecanismos cognitivos que usamos para explicar e prever o nosso próprio comportamento, que é possuído por cada um de nós. Contraria-se, portanto, a ideia de que apelamos a uma teoria – que aprendemos previamente –, pelo que a existência de uma teoria que possui como postulados centrais as atitudes proposicionais deixa de ser necessária.

Por que razão a teoria da simulação refuta a abordagem reducionista (AR)? O relevante desta objeção passa pela ideia de que esta crítica nega um dos pressupostos fundamentais do reduccionismo (seja a versão de Churchland ou a de Bickle): a necessidade de postular uma teoria – neste caso, a

Psicologia Popular (*Folk*) – que será reduzida à teoria-base – neste caso, uma teoria neurofisiológica futura. Veja-se a formulação deste argumento (8) contra (AR):

- (8.1) (AR) postula a teoria da (PF) para ser alvo de redução;
- (8.2) (PF) não é uma teoria (argumento da simulação);
- (8.3) Portanto, (AR) não possui uma teoria que seja alvo de redução;
- (8.4) Se não existe uma teoria alvo de redução, então (AR) é falsa;
- (8.5) Portanto, (AR) é falsa.

Fica, assim, claro de que modo a refutação da (PF) como teoria põe em causa o projeto reducionista da (AR).

Embora esta objeção seja interessante, não parece crucial. Suponha-se, por exemplo, que ninguém tenha subscrito uma teoria que postula a existência de anjos. A demonstração de tal facto não nos parece ser útil como prova contra a existência de anjos. Mesmo que não haja uma teoria dos anjos, nada impossibilita que a existência dos anjos seja real. O mesmo pressuposto é válido, por exemplo, com a existência dos prótons e electrões: muitos séculos antes, não haveria uma teoria que postulasse a existência destas partes fundamentais dos átomos.

Apesar disso, os electrões e prótons existem há milhões e milhões de anos e o facto de que a visão pré-científica ou popular (*Folk*) do mundo não concebesse a existência destas partes, tal não invalida que estas existiam e continuem a existir. Podemos, assim, generalizar a seguinte tese:

(9) Um sujeito *S* não necessita de consultar uma teoria sobre *x* para que *x* exista, dado que a falha em possuir uma teoria sobre *x* é consistente com:

- (i) A não existência de *x* (e.g. anjos);
- (ii) A existência de *x* (e.g. electrões e neutrões).

Assim, com (9), demonstramos que o facto de alguém rejeitar que um sujeito tacitamente faça uso dos postulados teóricos da teoria da (PF) não serve para refutar o pressuposto desta como teoria que será depois alvo de redução.

Além disso, alguns defensores da ideia de que (PF) é uma teoria (e.g. Gordon, 1992) utilizam dados da Psicologia do desenvolvimento para reforçar a defesa da sua tese (refutando assim (8.2)): os psicólogos Henry Wellman e Alison Gopnik sugerem que as crianças passam por diferentes fases

semelhantes às que um indivíduo passaria se estivesse a aprender uma teoria e a fazer uso dos seus postulados teóricos.

Os autores também sugerem que as crianças parecem atribuir crenças a elas mesmas de forma semelhante à atribuição de crenças aos outros indivíduos, provando que a atitude proposicional da crença é utilizada como postulado pela teoria da (PF), em contraposição com a explicação da simulação (cf. Gopnik e Wellman, 1992).

Finalmente, como fizemos notar, não é necessário aceitar a tese de que a Psicologia Popular (*Folk*) é realmente uma teoria: como indicado, Bickle sugere que basta aceitar que os seus postulados intencionais funcionam *como se* fossem parte de uma teoria, sendo irrelevante do ponto de vista da redução se, de facto, o conjunto de postulados formam, efetivamente, uma teoria (cf. Bickle, 1998: 43).

III.2.2. A Objeção da “Realização Múltipla”

Uma segunda forma de colocar em causa a (AR) é avançada por dois filósofos americanos: Hilary Putnam (1975) e Jerry Fodor (1974). Putnam fez notar que um determinado estado mental x pode ser realizado por diversos e distintos processos físicos que não partilham, entre si, nada de relevante. Especificamente, ficou famoso o exemplo de Putnam em relação ao estado mental da dor. O mesmo estado de dor parece poder ser imputado:

- (i) A diferentes padrões neuronais no sistema nervoso da mesma pessoa;
- (ii) A diferentes padrões neuronais de seres vivos diferentes (humanos, polvos, ratos, moluscos, etc.);
- (iii) A diferentes padrões neuronais de seres que não possuem sequer um sistema nervoso baseado na biologia orgânica terrestre (como eventuais extraterrestres ou computadores digitais);

Assim, o objetivo de Putnam é mostrar que, por causa de (i), (ii) e (iii), qualquer identidade entre o fenómeno mental “dor” e um processo neurofisiológico específico é, assim, falsa, dado que (i) na mesma pessoa, diferentes circuitos neuronais podem ser ativados para o mesmo estado mental de dor, (ii) o estado mental de dor faz parte de diversas espécies biológicas como ratos, polvos e moluscos e (iii) podemos supor mundos possíveis ou imaginários onde um extraterrestre ou um computador digital

poderão sentir dor embora tenham sistemas nervosos radicalmente diferentes do sistema nervoso do ser humano ou dos animais não-humanos.

Podemos formalizar esta crítica da seguinte forma. O reducionismo pode ser concebido defendendo a tese:

(Red): Para todos os sujeitos S (x tem dor = x instancia a propriedade física P);

A crítica de Putnam a (Red) é exatamente fornecer contraexemplos que demonstram que a equivalência bilateral e unitária entre “dor” e uma propriedade física específica P é falsa, dado que, para “ x tem dor”, poderão existir diversas propriedades físicas P s que podem ser instanciadas, e não somente uma, como parece sugerir o reducionista. Nas palavras de Putnam:

Considere aquilo que o [sc. reducionista] tem de alcançar para cumprir as suas reivindicações. Ele necessita de especificar um estado físico-químico tal que qualquer organismo (e não apenas um mamífero) sofra [sc. dor] se e somente se (a) possuir um cérebro com uma estrutura físico-química adequada; e (b) o seu cérebro está nesse estado físico-químico. Tal significa que o estado físico-químico em questão deve ser um estado possível em cérebro de mamíferos, no cérebro reptiliano, no cérebro de um molusco [...] Mesmo que esse estado possa ser encontrado, deve ser nomologicamente certo de que também será um estado do cérebro de qualquer vida extraterrestre que seja capaz de sentir dor [...] Portanto é pelo menos possível que a evolução paralela, em todo o universo, possa sempre levar a um e o mesmo “correlato” físico da dor. Mas essa é certamente uma hipótese ambiciosa (Putnam, 1975: 436).

A consequência ontológica desta crítica é defender que “dor” é uma propriedade funcional que tem – no mundo atual ou possível – diferentes instanciações neurofisiológicas (*realizers*). Deste modo, não existe uma propriedade física P com a qual a propriedade “dor” possa ser identificada unilateralmente.

Jerry Fodor argumentou também a favor desta tese, aplicando-a de forma mais geral à tese do reducionismo: Fodor irá focar-se no modelo de redução que tinha disponível na literatura científica do seu tempo, o modelo de Nagel. Ora, o importante a reter da objeção da realização múltipla é que parece apenas afeitar – se afetar de todo – o que podemos chamar de “reducionismo clássico”, sendo que o reducionismo de Bickle poderá escapar-se a essa refutação específica. Vejamos, assim, quais os problemas deste argumento.

Uma primeira forma de refutar a ideia de Putnam é afirmar que o argumento pressupõe diversos factos sobre a vida mental e a neurofisiologia associada de diversos seres vivos como mamíferos ou polvos ou moluscos. Mas será que, por exemplo, moluscos sentem realmente dor? Podemos conceber que, apesar de um molusco reagir a estímulos, esse comportamento possa ser automático com total ausência de dor. Ou, se queremos conceber que os moluscos possuem de facto dores, pode ser defendido que estas são qualitativamente distintas das dores dos seres humanos, que são seres muito mais complexos com uma neurofisiologia associada muito mais intrincada.

Assim, esta crítica sugere que o argumento de Putnam está refém de evidências empíricas sobre o sistema nervoso e a vida mental de outros seres vivos que poderão não estar ainda disponíveis ou que, se existem, são inconclusivas. Ora, esta forma de refutar a objeção em causa não nos parece devastadora, dado que apenas refere que a tese precisa de ser complementada empiricamente não tendo, assim, um problema lógico na sua estrutura que refute por completo a tese defendida.

Aos ombros de David Lewis (1969), um contra-argumento importante a esta objeção foi avançado. O filósofo norte-americano começa por admitir que o argumento da realização múltipla pode ser um obstáculo ao reducionismo, mas apenas se este for concebido na sua versão forte: apenas se defendermos uma teoria da identidade-espécime em que estado mental x é identificado unicamente com um espécime de uma propriedade física P .

Contudo, grande parte dos reducionistas contemporâneos – como Churchland ou Bickle – não aceitam esta tese. Antes, concebem uma versão menos exigente de reducionismo, defendendo que, em alguns casos, pode ocorrer efetivamente que um estado mental x se identifique não com um espécime (*token*) de uma propriedade física P , mas, antes, com um tipo (*type*) de propriedade física P que poderá, essa sim, ser multiplicadamente realizável no nível reduzido.

Veja-se o exemplo avançado por Churchland sobre o conceito de “temperatura” na teoria termodinâmica clássica. A temperatura num gás é idêntica à energia cinética molecular média. A temperatura num sólido, no entanto, é idêntica à energia cinética molecular máxima, uma vez que as moléculas de um sólido estão ligadas em estruturas de rede e, portanto, possuem maior constrangimento de movimentos vibratórios entre as moléculas do que num gás ou num líquido. Já a temperatura no plasma é algo totalmente dissemelhante, dado que os constituintes moleculares de um plasma foram destruídos (cf. Churchland, 1986: 356; Churchland, 1988).

Podemos, a partir deste exemplo, afirmar que o conceito de “temperatura” da termodinâmica clássica é realizável microfisicamente em diversos estados físicos distintos embora o conceito em análise

tenha sido reduzido e levado a simplificações ontológicas na teoria-base. Outros exemplos da história da Ciência são usados por diferentes autores para mostrar este ponto (cf. Hooker, 1981; Enç, 1983).

Pode-se, contudo, colocar em causa este exemplo da história da Ciência afirmando que, por muito interessante que seja a redução do conceito de “temperatura” na teoria termodinâmica, tal é irrelevante, filosoficamente, para analisar reduções interteóricas na Filosofia da Mente (cf. Zangwill, 1992: 215).

Mas por que razão Bickle considera que, mesmo aceitando que a objeção da realização múltipla possa ser válida, com efeito, para o reducionismo clássico, esta não afeta a sua abordagem reducionista particular? O segredo está, como vimos, no conceito de “modelo potencial”: para Bickle, (T_2) e (T_1) não se encontram numa relação de uma função 1 para 1 – reducionismo clássico – mas, antes, a teoria-base possui vários modelos potenciais que serão relacionados com o modelo potencial único da teoria reduzida:

Portanto, a perspectiva [*sc.* de Bickle] da redução interteórica permite que os casos que envolvem teorias reduzidas multiplicadamente realizáveis na teoria-base. A relação de redução p pode relacionar-se com muitos modelos de potenciais espécimes [*sc.* *tokens*] da teoria de redução a um único modelo potencial da teoria reduzida. Esta [...] é uma refutação interessante contra a aplicabilidade dos argumentos da realização múltipla contra o reducionismo [*sc.* de Bickle]” (Bickle, 1998: 115).

Deste modo, Bickle pensa livrar-se da força inicial da objeção da realização múltipla. Churchland considera que esta refutação do reducionismo é, no fundo, uma falácia do espantalho: o exemplo extraído da termodinâmica mostra que as reduções podem ser relativas a um domínio de fenómenos e, ao contrário do que o crítico reducionista quer fazer parecer, é uma parte normal da restante evolução da Ciência.

Assim, o facto de não haver reduções conectadas “um a um” (*one-to-one*) entre as teorias, isto é, o facto de as reduções serem relativas ao domínio não significa que sejam reduções falsas ou impostoras. Veja-se um segundo exemplo a favor deste contra-argumento:

Dois volumes de um gás podem ter a mesma temperatura, mas as distribuições de velocidades das suas moléculas constituintes serão bastante diferentes mesmo que o seu valor médio seja o mesmo. Para ser consistente, os funcionalistas devem negar novamente

o sucesso redutivo à mecânica estatística, pois, como diriam, a temperatura de um gás é compreendida de maneira diferente nos dois casos (Churchland, 1986: 357).

Concluimos, assim, que a objeção da realização múltipla não parece eficaz para rebater a abordagem reducionista (AR) de Churchland e Bickle – embora possa sê-lo contra o reducionismo clássico. Assim, analisemos uma terceira objeção comumente avançada pelos críticos reducionistas.

III.2.3. A Objeção da “Redução Ontológica”

Uma objeção que pode ser levantada a (AR) é o facto de os reducionistas estarem, supostamente, mais preocupados em reduzir os estados mentais a estados cerebrais do que nas reduções propriamente ditas. Relembramos que a tese reducionista é, no fundo, uma tese sobre a natureza ontológica dos postulados da Psicologia Popular (*Folk*).

Ora, se tal é o caso, então a própria teoria a ser reduzida perde qualquer utilidade e autonomizá-la como teoria fundamental da Neurociência parece ser um exercício fútil dado que os postulados da primeira irão desaparecer (cf. Bickle, 1998: 145). Assim, se a mente é o cérebro, então, segue a crítica, o foco do trabalho reducionista é querer reduzir a ontologia da mente à ontologia do cérebro, sendo apenas necessário a segunda para alcançar uma explicação do fenómeno mental.

Será que esta objeção aponta para um problema fundamental da (AR)? A resposta de Bickle é interessante e parece promissora, baseando-se na seguinte ideia de Paul Churchland: “São as reduções interteóricas suaves [*smooth*] que motivam e sustentam afirmações de identidade entre teorias, e não o contrário” (Churchland, 1985: 50). O que tal significa?

A ideia é afirmar que as consequências ontológicas – sejam elas quais forem – não são o objetivo principal da (AR), sendo que estas são somente secundárias, dependendo primariamente do tipo de reduções interteóricas alcançadas entre as teorias em questão. Nas palavras de Bickle: “[esta crítica] vê a ontologia do reducionismo como guiando a sua metodologia [...] mas [...] as conclusões ontológicas dependem do resultado da teoria da redução [*sc.* e não vice-versa]” (Bickle: 1998: 146).

Patricia Churchland já tinha também afirmado que muita da confusão filosófica sobre o reducionismo poderia ficar facilmente resolvida quando os investigadores da área perceberem que a questão de saber se os estados mentais se reduzem a estados cerebrais deve ser colocada de outra

forma: existe alguma teoria sobre o mental que possa ser reduzida a uma teoria cerebral (cf. Churchland, 1986: 279)?

Assim, a forma de rebater este argumento é indicar que a objeção confunde exatamente a motivação de um reducionista: a ontologia é logicamente posterior, isto é, uma consequência secundária, à redução propriamente dita – o que se quer alcançar não é a redução do conceito mental de “dor” a um estado (ou uma variedade) cerebral, mas, antes, a redução de uma *teoria* do mental a uma *teoria* cerebral.

Igualmente interessante é o diagnóstico que Bickle faz desta objeção: trata-se de uma forma tradicional e comum, entre os filósofos, de ver os problemas ontológicos como prioritários em relação aos problemas metodológicos ou de outro tipo sendo exactamente esse o erro que esta objeção comete – vemos, aqui, um pressuposto metafilosófico interessante que recusa a forma tradicional de ver o trabalho filosófico (cf. Bickle, 1998: 147).

Além disso, segundo o filósofo americano, é falso que mesmo havendo redução, todo o escopo da teoria reduzida – os seus conceitos, estruturas, práticas, fundamentos, explicações – seja totalmente eliminado. Tal dependerá, novamente, da posição no espectro de reduções (cf. Figura 3.1) em que a redução irá incidir. Uma redução suave pode, além de preservar relações de identidade entre as teorias, reter práticas científicas específicas que façam ainda sentido mesmo que a redução tenha sido alcançada: “os cientistas ainda usam física ótica e a termodinâmica de equilíbrio com mais frequência do que a química do flogisto” (Bickle, 1998: 144).

A razão para tal acontecer é que as teorias reduzidas são, normalmente, mais gerais no seu escopo de explicação, pelo que a prática científica e tecnológica tende a estar associada às teorias reduzidas – considerando somente, novamente, as reduções suaves. Daí que, para Bickle, não é necessário para um reducionista do seu tipo recomendar a total eliminação da Psicologia Popular (*Folk*): esta pode ainda manter a sua utilidade, tanto conceptual como pragmaticamente (cf. Bickle, 1998: 144). Vejamos, assim, uma outra forma de criticar o projeto reducionista.

III.2.4. A Objeção do “Isolacionismo Metodológico”:

Como observámos no segundo capítulo, a abordagem reducionista será problemática para alguém que defende uma concepção isolacionista da Filosofia e da Neurociência, dado que (AR), no seu âmago, defende que a primeira deve ser tendencialmente reduzida à segunda.

Peter Hacker, ao analisar o reducionismo, declara diretamente o seu isolacionismo, afirmando que:

Precisamente porque as diversas formas de reducionismo ontológico e explicativo são teses metafísicas relacionadas com a lógica das atribuições existenciais e as estruturas lógicas da explicação, não estão abertas à validação ou invalidação científica. Se tiverem de ser confirmadas ou refutadas, então sê-lo-ão por argumentos analíticos (Hacker & Bennett, 2003: 356-357).

Para um isolacionista, a (AR) não passa de uma confusão conceptual que deve ser posta de parte por cometer dois erros centrais:

- (i) Não parece haver qualquer razão para afirmar que duas pessoas que estejam no mesmo local da cidade e que se apercebam de que está a começar a chover formem a crença “está a começar a chover” tenham de ter ativados exactamente os mesmos circuitos neuronais correlacionados à crença “está a começar a chover”. Ou seja: os critérios de identificação de estados mentais são diferentes dos critérios de identificação dos estados neuronais. Veja-se a seguinte contradição: se o estado mental x é idêntico ao estado neuronal y , então, sempre que o sujeito S está no estado mental x , está, também, no estado neuronal y ; mas uma crença não pode ser idêntica a um estado neuronal y dado que o sujeito da crença é S e o estado neuronal y não é um adequado sujeito, sendo que apenas o cérebro de S pode ser o sujeito adequado para um estado cerebral y , finalmente, o sujeito S não é idêntico ou redutível ao seu cérebro – nem a nenhuma parte particular do seu corpo, sendo antes o seu conjunto;
- (ii) Se (AR) está correta e um estado mental y é idêntico a um estado neuronal y , então poderíamos identificar exactamente em que zona do nosso cérebro a crença “está a começar a chover” se localiza e, além disso, essa zona teria de ser a mesma para todos os seres humanos

que formassem a crença “está a começar a chover”. Mas tal não faz qualquer sentido (cf. Hacker & Bennett, 2003: 360-361).

Ora, como já fizemos notar, estas duas críticas podem, de facto, ser eficazes contra um certo tipo de reducionismo que advoga relações identitárias “um a um” entre estados mentais e estados neuronais. Mas, como também fizemos notar, a (AR) de Churchland e Bickle não se identificam primariamente com essa abordagem do reducionismo clássico, pelo que esta refutação não parece afetar diretamente o projeto reducionista em geral, mas apenas uma perspetiva específica sobre o reducionismo.

O importante a reter da sua crítica neste ponto é o seu pressuposto metafilosófico (como debateremos já de seguida): Hacker não aceita a conceção de Filosofia ou Ciência usada pela (AR). Como veremos, os reducionistas são tipicamente filósofos que montam a sua teoria nos ombros de filósofos como Quine, defendendo um naturalismo²⁶ que vê a Ciência e a Filosofia em continuidade, e não, como notámos no segundo capítulo, duas disciplinas fundamentalmente separadas.

Ademais, como veremos, os reducionistas podem ser considerados os filósofos que radicalizaram ao máximo as pretensões metafilosóficas de Quine, tendo extremado a sua posição.

Restará apontar que, além da discórdia metafilosófica, Hacker também põe em causa a tese de que a Psicologia Popular (*Folk*) é de facto uma teoria de variadas formas:

- (i) O facto de conceitos se relacionarem entre si não implica que estes formem uma teoria;
- (ii) O facto de a linguagem (neste caso, sobre o mental) implicar a interligação de vários conceitos não implica que haja uma teoria; apenas aponta para o carácter holístico e geral de uma linguagem, sendo composta de conceitos e de regras particulares para o seu uso;
- (iii) O facto de adquirir-se um conceito não significa que se aprenda uma teoria;
- (iv) Os conceitos mentais ou psicológicos não são entidades ou postulados teóricos como os genes ou os protões; diferentemente, são abstrações (cf. Hacker e Bennett, 2003: 368-370).

Ora, como vimos na secção II.2.1., negar que a Psicologia Popular (*Folk*) seja uma teoria não refuta a abordagem reducionista diretamente, mas fá-lo indiretamente: dado que (AR) necessita de uma teoria candidata a ser reduzida e dado que necessita que essa teoria seja composta de postulados teóricos, se esses postulados teóricos não existem, e se, afinal, não há qualquer teoria, então a (AR) é refutada

²⁶Especificamente, a Abordagem Reducionista fará uso de um “naturalismo de substituição”, como veremos no quinto capítulo.

indiretamente por não conseguir oferecer uma teoria candidata à redução a uma teoria-base da Neurociência. A não ser que o reducionista ofereça outra teoria, esta objeção poderá ter a sua força filosófica e deve ser considerada.

Claro, para Bickle, basta aceitar, de certa forma, o âmago desta objeção, defendido um sentido minimal de Psicologia Popular (*Folk*), como indicámos acima: bastará que (PF) se comporte *como se* fosse uma teoria, não importando realmente se o é, de facto.

III.2.5. A Objeção Metafilosófica

Como desvendámos acima, há uma pressuposição metafilosófica da Abordagem do Reduccionismo que poderá ser problemática. Veja-se a primeira citação que Bickle tem no seu livro de 2003:

E, no entanto, existem filósofos que se recusam a reconhecer a filosofia científica como filosofia, que desejam incorporar os seus resultados num capítulo introdutório da ciência e afirmam que existe uma filosofia independente que não se preocupa com a investigação científica e tem acesso direto à verdade. Penso que tais afirmações revelam uma falta de julgamento crítico. Quem não vê os erros da filosofia tradicional não quer renunciar aos seus métodos ou resultados e prefere seguir um caminho que a filosofia científica abandonou. Eles reservam o nome de filosofia para as suas tentativas falaciosas de um conhecimento supercientífico e recusam-se a aceitar como filosófico um método de análise projetado segundo os padrões da investigação científica. O que é necessário para uma filosofia científica é uma reorientação dos desejos filosóficos (Reichenbach, 1957: 305).

Ora, esta citação e forma de abrir um livro não é inocente filosoficamente: Bickle procura demonstrar ao leitor que, logo à partida, uma forma tradicional de fazer Filosofia não será encontrada naquele livro. Além disso, no seu prefácio, declara: “Estou a juntar as peças individuais que a ciência fornece para tornar explícito o ‘quadro geral’ que a maioria dos cientistas deixa implícito. Tal ‘filosofia’ é suficiente para mim” (Bickle, 2003: xv).

Assim, Bickle pretende enquadrar-se claramente na tradição naturalista de Quine que faz uso, direta ou indiretamente, isto é, criando experiências científicas diretamente, ou apelando a experiências científicas de outros investigadores, para filosofar sobre variados temas e problemas. Não será por isso espantoso que, por exemplo, na obra de 2003, Bickle não cite ou faça referência uma única vez a Peter

Hacker – aquele que seria o seu “adversário” filosófico direto – e refira Wittgenstein somente uma vez, declarando algo curioso: “[...] deve ficar claro da discussão que não estou a acusar a neurociência cognitiva de contrassenso” (Bickle, 2003: 160).

Fazemos referência a esta nota de rodapé no seguinte sentido: Bickle poderia de facto ignorar a filosofia de Hacker por simplesmente não conhecer nem a sua fonte primária (Wittgenstein), nem a sua tese filosófica (a existência do sentido e do seu oposto). Ora, fica claro que, curiosamente, Bickle conhece ambos, mas que na sua obra não direciona minimamente a sua atenção para discutir diretamente com Hacker (ou outro isolacionista) a relação entre Filosofia e Neurociência.

Esta consideração levanta assim uma objeção metafilosófica interessante, que já fizemos notar no primeiro capítulo: Bickle não discute qualquer princípio metafilosófico, porque assume, *a priori*, uma conceção de Filosofia como “científica”, como fica claro na citação de Reichenbach. Mas essa conceção é assumida sem discussão primária.

Portanto, Bickle parece confundir uma determinada conceção de fazer Filosofia como a única conceção filosoficamente relevante, cometendo uma falácia clara de confundir o nível descritivo metafilosófico com o nível normativo ou prescritivo.

Ora, como Hacker faz notar, a abordagem reducionista parece estar comprometida com três teses que poderão ser falsas:

- (i) Todo o conhecimento que seja conhecimento genuíno é científico;
- (ii) A Ciência natural é a única ciência que nos pode oferecer conhecimento;
- (iii) A única compreensão genuína do mundo é científica (cf. Hacker & Bennett, 2003: 372-373).

Mas não é óbvio ou claro porque temos de aceitar (i) – (iii) como verdades verossímeis. Lembremo-nos da conceção metafilosófica da Filosofia de Quine: “[a filosofia] é contínua com a ciência, e é mesmo parte da ciência [...] a filosofia está no fim abstrato e teórico da ciência” (Quine, 1969: 83).

Mas contra esta visão metafilosófica podemos contrapor outra, a do segundo Wittgenstein: “[quando estamos a fazer filosofia] não estamos a fazer ciência natural, nem sequer história natural”, e que, ao fazermos filosofia, “as nossas considerações não poderiam ser consideradas científicas” (Wittgenstein, 1953: § 109).

Se seguirmos Wittgenstein, (i) – (iii) deixam de fazer sentido:

- (i) Está equivocada porque existem muitos fenómenos que não são explicados cientificamente nem pretendem sê-lo: veja-se o caso do Direito, da Economia, ou da Arte; é claro que, dentro da Ciência, tudo o que podemos encontrar são explicações científicas; mas desta tese para uma generalização indutiva de todas as áreas do conhecimento, parece uma imprudência filosófica declarar que todo o conhecimento que é conhecimento genuíno é científico, dado que esta tese não pode sequer ser comprovada empiricamente, constituindo, seguindo o seu próprio critério, um exemplo de falso conhecimento;
- (ii) É, assim, uma tese igualmente implausível: uma explicação científica da *La Gioconda* deixa de fora qualquer interesse estético dessa obra de arte; embora possamos efetivamente analisar cientificamente este quadro, isto é, ir ao Museu do Louvre e analisar, por exemplo, o grau de pigmentação, os elementos químicos que constituem a pintura, ou mesmo a técnica de pintura utilizada – toda esta descrição científica não nos informará minimamente dos aspetos normativos ou estéticos da famosa pintura de Da Vinci;
- (iii) Está, também, errada: de facto, a Ciência não oferece qualquer compreensão do mundo, mas somente descrições; sendo que a elucidação da linguagem será o trabalho da Filosofia.

No entanto, a defesa do (ii) naturalismo metafilosófico forte procurará negar todas estas teses (cf. Capítulo 1). Hacker chega mesmo a atacar, diretamente, o pensamento de Quine numa nota de rodapé, afirmando que

Quine e os seus discípulos julgam que qualquer “conceptualização” se baseia na teoria. De modo que a teoria está sempre envolvida quando se usa a linguagem [...]. Eles sugerem que não há qualquer distinção entre verdades conceptuais [...] e verdades empíricas [...] – e, por implicação, que não há qualquer diferença entre normas de representação (que é o que a maior parte das verdades conceptuais são) e proposições de facto. Daí, também, pensarem que a filosofia é a continuidade da ciência, e parte do mesmo empreendimento [...]: a saber, a construção de teorias acerca do “mundo”. Do nosso ponto de vista, não há nada que possa estar mais longe da verdade (Hacker & Bennett, 2003: 380).

Fica claro, nesta citação, por que razão um isolacionista condena qualquer tentativa reducionista – seja a clássica ou a atual – de resolver os problemas da Filosofia da Mente.

Churchland chega a abordar de forma breve a crítica da falácia mereológica da Neurociência avançada pelos isolacionistas. Nas palavras da autora:

[...] a sugestão de que o cérebro se lembra ou tem conhecimento ou usa símbolos linguísticos é por vezes posta no pelourinho como um simples erro conceptual que consiste em pegar em categorias apropriadas para um domínio e aplicá-las a um domínio diferente e inapropriado. Mas o erro de categoria de uma pessoa é a teoria profunda de outra pessoa acerca da natureza do universo [...] O que há de importante para conseguir a verdade acerca dos cérebros não é saber se no uso comum ou popular (*folk*) as pessoas normais da rua dizem que as pessoas se lembram mas não dizem que os cérebros se lembram; em vez disso, é saber se devíamos dizer que os cérebros recordam – se, tendo em conta os factos empíricos, é uma hipótese razoável dizer que os cérebros recordam (Churchland, 1986: 273).

Ora, mas será que o isolacionista tem razão na objeção metafilosófica que levanta? O próprio Quine parece dar algumas indicações em sentido negativo. Talvez Hacker, por exemplo, esteja a utilizar um conceito ou definição de ciência substancialmente diferente da que Quine ou os naturalistas em geral – incluindo os reducionistas – usam como pressuposto metafilosófico.

Veja-se de que forma Quine define naturalismo como sendo “o reconhecimento de que é dentro da própria ciência, e não de uma filosofia *a priori*, que a realidade deve ser identificada e descrita” (Quine, 1981: 21). Ora, Hacker contrapõe a esta tese a ideia de que é absurdo que a Economia ou a História fiquem de fora do que pode ser considerado conhecimento genuíno apenas porque não são disciplinas científicas. Contudo, Quine nega esta ideia declaradamente: o filósofo naturalista incorpora dentro da sua definição de ciência variadas disciplinas como a Economia, a Sociologia ou a História (cf. Quine, 1995: 49).

Ora, podemos estar perante um problema de clareza conceptual. A conceção de “ciência” usada pelos isolacionistas é diferente da conceção de ciência dos naturalistas. De um lado, temos uma definição muito precisa de ciência: ciência equivale à ciência natural. Do outro, temos uma conceção de ciência mais alargada, que inclui diversas disciplinas, que podem ser mais ou menos empíricas.

Assim, o papel e a forma como a Filosofia será definida poderá estar dependente da forma como cada autor ou investigador define ciência: partindo de um pressuposto estreito de ciência, uma conceção de Filosofia anticientífica, como a apresentada pelos isolacionistas (cf. Capítulo 2), poderá fazer sentido. Todavia, partindo de um pressuposto ampliado de ciência, uma conceção naturalista da Filosofia poderá também fazer sentido.

Ademais, o próprio termo “naturalismo” pode ser interpretado com mais ou menos escopo teórico:

[...] naturalismo é um conceito aglomerado ou termo de semelhança de família; isto é, um conceito definido por uma mistura de qualidades, nenhuma das quais essencial. Naturalismo na filosofia pode ser distinguido por uma variedade de objetivos, crenças, princípios, premissas, procedimentos metodológicos, atitudes e valores (Overgaard, Gilbert & Burwood, 2013: 55).

Podemos, portanto, concluir que qualquer estratégia essencialista em relação à metafilosofia adotada poderá estar condenada dado que dificilmente poderemos acertar uma definição explícita e essencial do conceito “Filosofia” ou “Ciência” sem excluir, por um lado, formas praticadas na investigação filosófica corrente, assim como diversas disciplinas científicas.

No fundo, a predileção metafilosófica de cada investigador poderá estar mais ligada a preferências subjetivas – como a formação do investigador, os pares de que se rodeia, os problemas que investiga, etc., – do que por uma posição metafilosófica sólida e coerente.

A (AR), como uma tese naturalista forte, defende no seu escopo o que podemos apelar de “naturalismo ontológico”: a tese de que tudo o que existe deve ser descrito em termos das categorias naturais oferecidas pelas ciências. Mas, mais importante do que esta mera tese descritiva, é a sua tese prescritiva: a afirmação que tudo o que poderá vir a existir ou a ser esclarecido no futuro – dado que atualmente se encontra epistemicamente incerto por variadas razões contingentes como, por exemplo, a existência da vida mental subjetiva – terá, necessariamente, de ser descrito pelas mesmas categorias naturais. Ora, o reducionista poderá confundir as suas teses descritivas com as suas teses prescritivas e tal poderá ser o maior obstáculo desta abordagem filosófica.

Finalmente, e como veremos no capítulo 5, a Abordagem do Reduccionismo pode ser caracterizada por três características específicas: (i) monismo de domínio; (ii) monismo metodológico; e (iii) uma abordagem cérebro-redutora (*brain-reductive*) (cf. Northoff, 2014a: 99). Como exporemos nesse capítulo, podemos negar estas três características sem colocar de parte a importância da investigação neurocientífica do mental e sem cair, igualmente, numa forma tradicional de fazer Filosofia.

III.3. Conclusão:

Neste terceiro capítulo, avançamos com uma das teses mais relevantes para resolver o problema da relação entre a Filosofia e a Neurociência: ao contrário da anterior, que negava o problema e defendia a separação clara das duas disciplinas, a (AR) defende que a Filosofia – os seus conceitos, métodos e problemas – devem ser tendencialmente substituídos pelos conceitos de uma teoria evolutiva da Neurociência.

Analisámos, particularmente, as diversas formas de conceber a redução interteorética, a importância da defesa da Psicologia Popular (*Folk*) como uma teoria candidata à redução, assim como dois dos autores mais importantes desta perspetiva. De seguida, para concluir, analisámos as críticas mais relevantes a esta posição, fazendo notar de que forma as conceções metafisológicas dos autores influenciam diretamente as suas perspetivas metodológicas em questão.²⁷

²⁷ Este capítulo adapta e expande partes mínimas do capítulo “Como estudar a nossa mente?” (Gouveia, 2018) da nossa obra *Reflexões Filosóficas: Arte, Mente e Justiça* pela Editora Húmus.

IV. NEUROFENOMENOLOGIA

Como vimos anteriormente, o problema metodológico de relação entre a Filosofia e a Neurociência aparece na tentativa de compreender se estas duas disciplinas podem trabalhar em conjunto para lidarem com o problema da mente consciente. Analisámos, no capítulo 2, a importância do método filosófico da análise conceptual da linguagem. Neste capítulo, focaremos a nossa atenção sobre outro método usado em Filosofia: a fenomenologia.

A fenomenologia, como veremos, é o método filosófico de investigação da experiência quotidiana. O seu foco está na análise de diferentes fenómenos (do grego *phainomemon*, “aquilo que se mostra”) e do modo como estes aparecem à consciência, i.e., numa perspectiva da primeira pessoa. Assim, a fenomenologia é uma disciplina particularmente útil para compreender como é que os fenómenos aparecem à nossa mente, e como é que lhes fornecemos significados.

Já a Neurociência, como vimos, é o estudo científico do cérebro e do sistema nervoso, que lida com os aspetos cerebrais de uma perspectiva da terceira pessoa. Ora, muitos filósofos acreditam que nem uma perspectiva da primeira pessoa ou somente da terceira pessoa poderão ser suficientes para resolver ou dissolver problemas particulares da Filosofia da Mente. Antes, acreditam que, de algum modo, é necessário juntar estas duas ferramentas que parecem, por definição, inconciliáveis, numa metodologia de trabalho que nos ofereça vantagens explicativas em relação às suas concorrentes.

Historicamente, Edmund Husserl é visto como o filósofo que construiu a fenomenologia como uma disciplina filosófica coerente e como uma ferramenta concreta de estudo dos estados conscientes (e.g. *epoché*). Husserl, lembre-se, era um aluno de Brentano, filósofo que considerava que a propriedade mais importante da consciência era ser caracterizada por possuir intencionalidade (o “*aboutness*” da consciência, isto é: a consciência é sempre consciência de algo diferente dela). Uma atenção particular à fenomenologia do corpo foi desenvolvida por Maurice Merleau-Ponty na primeira metade do século XX.

Ora, como podemos já perceber, os dois métodos apresentam convergências quanto aos assuntos e temas investigados. Contudo, principalmente por discordâncias ontológicas e metodológicas entre a fenomenologia e uma Filosofia da Mente naturalizada, o diálogo entre as duas disciplinas é bastante penoso e controverso.

A Abordagem da Neurofenomenologia surge exactamente como tentativa de conciliar estas duas perspectivas. Analisemos, de seguida, os pressupostos desta abordagem.

IV.1. Pressupostos da Abordagem da Neurofenomenologia

IV.1.1. A Fenomenologia

A fenomenologia é um método da Filosofia que nasce no século XX com Edmund Husserl, pensador alemão que publicou o primeiro volume das suas *Investigações Lógicas* em 1900,²⁸ marcando o nascimento oficial desta abordagem ao estudo da mente consciente.

O objetivo de Husserl era claro: criar uma metodologia rigorosa que não pressuponha qualquer conhecimento pré-teórico sobre a mente ou qualquer objeto de estudo. Antes, o investigador deveria suspender qualquer juízo e focar-se no fenómeno “em si mesmo”.

Numa primeira fase, Husserl procura criticar a tese que dominava o estudo do mental na sua época: o psicologismo, a ideia de que todas as verdades sobre o mundo e sobre nós mesmos poderiam ser alcançadas a partir da dimensão psicológica do pensar. Ademais, seria um erro, para o filósofo alemão, reduzir a normatividade da lógica a uma mera análise descritiva da Psicologia.

Ora, Husserl²⁹ vai notar que tal tese é, fundamentalmente, falsa: o conceito de “pensamento”, aponta o filósofo, pode ser entendido em dois sentidos que é relevante distinguir. Um primeiro sentido é definir “pensamento” como uma atividade mental particular (*cogitatio*); diferentemente, “pensamento” pode também ser definido como o que pensamos em si mesmo, com existência ontológica própria e separada ou independente do pensar no primeiro sentido (*cogitatum*).

Veja-se o exemplo indicado no § 46 das *Investigações Lógicas*. Husserl indica que há uma diferença entre pensar a natureza do número e da sua relação com outros números, e o próprio acto mental de contar.

Assim, devemos distinguir claramente as essências das coisas que são totalmente independentes da *empíria*, da ligação destas com a experiência (*Erfahrung*) e vivência (*Erlebnis*) do ser humano. O foco da fenomenologia será então analisar a experiência considerada na sua manifestação no dia-a-dia: esta experiência não se reduz meramente à experiência sensorial que obtemos dos sentidos ou menos da vida mental em geral, mas, antes, é um conceito mais abrangente que inclui tudo o que “aparece à consciência” (cf. Overgaard, Gilbert & Burwood, 2013: 72).

²⁸ Antes, Husserl atribuía à Psicologia o papel de ciência fundamental do saber. Foi através da leitura de Frege e dos argumentos críticos ao psicologismo (recusando a tese de que a lógica se poderia reduzir à Psicologia) que Husserl moveu o seu foco para a fenomenologia.

²⁹ Husserl chega a afirmar que a sua teoria está em “guerra” direta com o psicologismo (cf. Husserl, 1969: 172)

É importante distinguir, desde já, duas concepções que o termo “fenomenologia” possui na literatura científica da área. Num sentido estrito, este conceito é usado pela Filosofia Analítica no sentido de descrever uma experiência subjetiva, aquilo que ficou famosamente conhecido na expressão de Thomas Nagel por “what is like to be *x*”, isto é, o conteúdo experiencial e subjetivo que é, por exemplo, uma percepção visual de uma paisagem (cf. Nagel, 1974: 435-450).

Mas o sentido em que o conceito de “fenomenologia” é empregado pelos fenomenólogos é um muito mais alargado que esta primeira significação: parte do seu sentido é realmente captado pela forma como os filósofos analíticos usam o termo, mas inclui-se igualmente o próprio objecto na relação de intencionalidade³⁰ – seguindo o exemplo, além da experiência visual da paisagem, também se inclui na sua análise o próprio objecto da percepção, neste caso, a paisagem (cf. Husserl, 2001: sec. 52).

Um exemplo interessante oferecido por Overgaard *et al.* baseado no pensamento de Husserl pode ser oferecido para compreender a diferença entre os dois sentidos:

Suponha que está sentado num jardim num lindo dia de primavera a admirar uma macieira em flor. A descrição fenomenológica estrita pode incluir descrições sobre as suas mudanças de atenção, a sua consciência tácita de mover os olhos, os sentimentos de alegria que acompanham a sua percepção e assim por diante. A descrição fenomenológica alargada [*sc.* da fenomenologia] adiciona algo sobre a cor das flores, a forma como a árvore de fundo se destaca no contexto de outros objetos (o relvado, a cerca, o alpendre do jardim e assim por diante) e a forma como os objetos são apresentados tendo lados e perfis que não podem ser percecionados [*sc.* diretamente] (a parte traseira do porta-malas, detalhes minuciosos das flores que só podem ser vistos de perto, etc.) e as qualidades estéticas que a árvore possui ao ser percecionada (é bela, a própria imagem da primavera e assim por diante) (Overgaard, Gilbert & Burwood, 2013: 73).

Assim, é importante compreender a intuição pressuposta neste exemplo: para a fenomenologia no sentido lato, o aspeto que mais lhe interessa analisar é o lado objetivo dessa relação entre sujeito e

³⁰ O termo “intencionalidade” não deve ser confundido com o termo “intensionalidade”: o primeiro irá ser detalhado de seguida e refere-se a algo diferente do próprio conceito que é referido (e.g. a consciência é sempre consciência de algo diferente de si mesma: por exemplo, do pôr-do-sol ou do sabor do café); já o segundo é um termo usado em lógica para referir a falta de preservação do valor de verdade depois de ocorrer a substituição dos termos coreferentes.

objeto, e não tanto o lado subjetivo do sujeito – embora esse polo seja também relevante na análise fenomenológica – ao contrário dos filósofos analíticos que se focam mais na experiência do sujeito (cf. McCulloch, 1995: 131).

Segundo Husserl, a relação entre um sujeito e um objeto ocorre sempre numa dualidade que é necessário distinguir com precisão: o conceito de “noesis” refere os modos como o objeto pode ser apreendido na sua intencionalidade;³¹ e o conceito de “noema” refere as potencialidades que o objeto tem de ser apreendido, afirmando que “não se deve confundir noema (correlato) e essência” (Husserl, 1980: 73).

Igualmente importante é a tese da *epoché*: segundo o filósofo alemão, vivemos com uma ideia implicitamente assumida, constantemente, de que o mundo externo é totalmente independente de nós e da nossa consciência – o que Husserl chama de “atitude natural” (cf. Husserl, 1983: sec. 30), indicando que devemos colocar “entre parênteses” todos os aspetos “em relação ao ser” (cf. Husserl, 1983: sec. 32).³²

Por causa disso, Husserl defende uma tese metafilosófica interessante que será alvo de crítica por alguns dos seus sucessores (como veremos de seguida): dado que as ciências empíricas se fundamentam na tese de que o mundo existe de forma independente (cf. “atitude natural”), as suas investigações e resultados devem ser ignorados na análise fenomenológica dos fenómenos dado que, na sua essência metodológica, a investigação científica já pressupõe a tese que é necessário colocar “entre parênteses”.

Todo este processo levará à “redução fenomenológica” permitindo que a investigação se foque “de volta às ‘coisas em si’” (cf. Husserl, 2001: 168). Isto é, este método permite “voltar” a uma análise mais pura e mais verdadeira da forma como a experiência se dá à consciência.³³ Poderá ser argumentado, todavia, que este método apenas dará acesso à experiência individual de cada um, mas não poderá retirar da análise fenomenológica regras gerais e universais sobre a consciência humana ou do próprio mundo. Será que tal serve como crítica à limitação do escopo da fenomenologia?

³¹Husserl aceita a doutrina de Brentano inovando-a num aspeto central: a análise das diversas formas de como é que o objeto da relação intencional é dado ou pode ser apreendido.

³²Fazemos notar que, contudo, esta distinção é alvo de muita controvérsia dentro da literatura especializada em Husserl (cf. Smith, 2007: 304-11).

³³Heidegger vai, contudo, interpretar criticamente a redução husserliana de outra forma, declarando que: “[...] para Husserl, a redução fenomenológica [...] é o método de levar a visão fenomenológica da atitude natural do ser humano, cuja vida está envolvida no mundo das coisas e das pessoas, de volta à vida transcendental da consciência [...] Para nós, redução fenomenológica significa levar a visão fenomenológica de volta à apreensão de um ser [...] para a compreensão do ser deste ser” (Heidegger, 1982: 21).

Husserl responderá que não: além da primeira redução (“redução fenomenológica”), é necessário executar uma segunda redução: “a redução eidética” – o ato de colocar entre parêntesis (novamente) todas as considerações contingentes e acidentais da experiência individual, focando-se na (ou intuindo a) natureza essencial dos objetos (cf. Husserl, 1983: sec. 2).

Para além disso, devemos usar a imaginação para perceber que tipo de propriedades do objeto são essenciais para definir a ontologia desse objeto, isto é, que propriedade, estando em falta, coloque a própria existência do objeto em causa – alcançamos este processo através da intuição eidética (cf. Husserl, 2001: 292-294).³⁴

Finalmente, faltará explicitar o conceito de “intencionalidade” que, como já fizemos referência, influencia Husserl através do seu professor, Franz Brentano, que definia o conceito da seguinte forma:

Todo fenómeno mental é caracterizado pelo que os Escolásticos da Idade Média chamavam de inexistência intencional (ou mental) de um objeto, e o que poderíamos chamar, embora não de forma totalmente inequívoca, de referência a um conteúdo, a direção a um objeto (que não deve ser entendido aqui como significando uma coisa) ou objetividades imanente. Todo fenómeno mental inclui algo como objeto em si mesmo, embora nem todos o façam da mesma maneira. Na representação, algo é representado, no julgamento, algo é afirmado ou negado, no amor amado, no ódio odiado, no desejo desejado e assim por diante (Brentano, 1995: 68).

Contudo, é importante notar que Husserl se afasta de Brentano: para o filósofo alemão, o objeto de um ato intencional é externo ao próprio ato (cf. Husserl, 2001: 126-127), pelo que devemos considerar a intencionalidade³⁵ não como uma relação, mas, diferentemente, uma propriedade intrínseca dos atos intencionais (cf. Smith & Thomasson, 2005).

IV.1.2. A Importância do Corpo e a Crítica ao Cérebro

O filósofo francês Maurice Merleau-Ponty é muito relevante para a abordagem fenomenológica ao estudo da mente humana. Tendo lido Husserl, focou a sua investigação na importância do corpo de um ponto

³⁴Curiosamente, alguns autores argumentam que este método não se diferencia substancialmente do método analítico da análise conceptual: especificamente, das experiências de pensamento imaginadas (cf. Zahavi, 2003: 38-39).

³⁵ Husserl distingue entre a matéria intencional (significado) de um ato consciente e sua qualidade intencional (cf. Husserl, 2001: 119-122).

de vista da primeira pessoa, descrevendo essa experiência específica como “corpo vivido” em contraste com o “corpo objetivo”.

Em que se baseia essa distinção? No fundo, a diferença encontra-se fundamentada no método utilizado para observar o corpo humano: por um lado, podemos observar o corpo de uma perspectiva de terceira pessoa, analisando a sua fisiologia, anatomia e organização empírica; por outro, e esta é a principal inovação filosófica de Merleau-Ponty, podemos experimentar o nosso corpo da perspectiva da primeira pessoa, focando a nossa atenção na experiência corporal – da sua conexão com o mundo – da forma como “aparece à consciência” (cf. Merleau-Ponty, 1989).

Mais relevante, este “corpo vivido” era, para o filósofo francês, a solução para o problema da consciência, sendo central para compreender este fenómeno complexo. O corpo é o fundamento-base dos restantes conteúdos mentais, assim como as outras pessoas (intersubjetividade) ou o próprio mundo: sem o corpo, não poderíamos sequer ter qualquer experiência.

A conceção de corpo de Merleau-Ponty surge como crítica à visão husserliana de intencionalidade que teria sido levantada também por Martin Heidegger: para ambos, a intencionalidade em Husserl parece caracterizar-se apenas como ato mental ou consciente. Todavia, esta visão deixa de fora o aspeto mais fundamental da intencionalidade para estes autores:

A conceção usual de intencionalidade [...] interpreta erroneamente a estrutura de autodirecionamento para [...] um ego ou sujeito, cuja esfera das experiências intencionais lhe devem pertencer [...] O modo de ser de nós mesmos, o *Dasein*, é essencialmente tal que esse ser, na medida em que é, reside com o que existe. A ideia de um sujeito que tem experiências intencionais meramente dentro da sua própria esfera e ainda não está fora dele, mas encapsulado em si mesmo, é uma absurdidade (Heidegger, 1982: 63-64).

Assim, para Heidegger, o “estar-no-mundo” é mais fundamental ou primitivo do que a conceção de intencionalidade dos atos conscientes descrita por Husserl. Todavia, é Merleau-Ponty que introduz explicitamente a importância do conceito de “corpo”, defendendo o aspeto fundamental da “intencionalidade motora” (cf. Merleau-Ponty, 1989: 105).

Nesta conceção, as atividades corporais, e não somente os atos conscientes, são incluídos na intencionalidade – o corpo direciona-se para o mundo estando aberto a diversas possibilidades:

[...] é o corpo que ‘captura’ e ‘compreende’ o movimento. A aquisição de um hábito é efetivamente a compreensão de um significado, mas é o entendimento motor ou um significado motor [...] [sc. e é] o corpo que ‘compreende’ (Merleau-Ponty, 1989: 143-144).

Desta forma, o filósofo francês procura afirmar que nós somos o nosso corpo, que a nossa experiência vivida acontece pelo corpo e que, fundamentalmente, a distinção cartesiana³⁶ entre mente e corpo está equivocada porque confunde, da mesma forma, a distinção entre sujeito e objeto (cf. Merleau-Ponty, 1989: xii).

O fenómeno da percepção ganha nesta perspetiva filosófica uma nova interpretação. Em vez de existir um sujeito que assume o papel de mero espectador passivo dos conteúdos percetivos, o sujeito existe situado no mundo e em ligação com o seu corpo – a percepção e a ação coincidem: “todo o hábito percetivo é ainda um hábito motor” (Merleau-Ponty, 1989: 153).

Para o filósofo francês, a tradição filosófica até então teria ignorado a existência paradoxal³⁷ do sujeito humano: somos parte – coexistensivamente – do mundo, mas, ao mesmo tempo, somos constituídos pelo mesmo (cf. Merleau-Ponty, 1989: 453). Ora, de forma interessante, está aqui pressuposta uma crítica feroz à Neurociência tradicional: esta erra no seu centro de investigação ao focar-se primariamente no estudo do cérebro e, apenas secundariamente, no estudo da relação do cérebro com o restante corpo.

De que forma é que a fenomenologia impacta a investigação da Neurociência? Cruzando os princípios-base da fenomenologia com o objetivo de que orientem a investigação empírica, alguns autores defenderam uma abordagem neurofenomenológica ao estudo da mente (cf. Petitot *et al.*, 1999; Lutz & Thompson, 2003; Rainville & Price, 2003).

Para o neurofenomenólogo, analisar a consciência no cérebro – Neurociência tradicional – seria como procurar o valor do euro numa nota de cinco euros: podemos ampliar ao microscópio qualquer parte da nota que não encontraremos o seu valor. Antes, o valor está na nota inteira e na sua interação com regras económico-financeiras de mercado específicas, e é apenas aceite como valendo alguma coisa se estiver inserida no ambiente económico de troca de valores (cf. Noë, 2009: 3-4). Nas palavras de Alva Noë:

³⁶ Curiosamente, Merleau-Ponty irá, na sua obra mais tardia, assumir que a sua grande obra *Fenomenologia da Percepção* não tinha conseguido libertar-se do cartesianismo dado que o ponto inicial de investigação pressupõe a existência de um sujeito a realizar a mesma e que a dicotomia sujeito-objeto teria influenciado ainda demasiado as suas ideias iniciais (cf. Merleau-Ponty, 1989: 179).

³⁷ No mesmo sentido, afirma: “Dentro e fora são inseparáveis. O mundo está totalmente dentro e eu estou totalmente fora de mim” (Merleau-Ponty, 1989: 407).

De certa forma, o problema é que estamos à procura da consciência onde ela não está. Devemos procurá-la onde está. A consciência não é algo que acontece dentro de nós. [sc. Antes] É algo que fazemos. Melhor: é algo que alcançamos. A consciência é mais como dançar do que como a digestão (Noë: 2009: xii).

A mente consciente, assim como os restantes estados mentais, devem ser considerados da mesma maneira. Veremos, na secção das objeções, de que forma esta forma de expor o problema da mente consciente pode ser problemática.

Analisemos as principais ideias do fundador da Abordagem da Neurofenomenologia, Francisco Varela.

IV.1.3. Os 4E's e a Abordagem Neurofenomenológica de Francisco Varela

O pensador pioneiro deste tipo de metodologia é, sem grandes dúvidas, Francisco Varela. Com formação filosófica e médica, procurou durante a sua atividade académica desenvolver uma metodologia de trabalho que fosse mais eficaz que as praticadas até então, criticando os pressupostos cartesianos – que eram ignorados pela própria prática neurocientífica. A primeira tese negativa, que é avançada por um neurofenomenólogo (TNN – Tese Negativa da Neurofenomenologia) e que se divide em duas subteses (1) e (2), é a seguinte:

(TNN)

(1) A fenomenologia por si só não é completa o suficiente para resolver os grandes problemas da mente – tome-se em consideração as tentativas de Husserl, Merleau-Ponty e Heidegger, por exemplo, em que os próprios apontam falhas ao seu sistema de pensamento, exactamente pela existência de um limite metodológico;

(2) A Neurociência, abstratamente considerada, não é suficiente para estudar os grandes problemas da mente – note-se os problemas que a ciência contemporânea tem a lidar com a existência da consciência na evolução, da subjetividade e até da doença mental.

Esta primeira tese, que é essencialmente negativa, é facilmente justificada pela própria atividade de fenomenólogos e de cientistas. Aliás, o fenómeno da consciência apenas tem entrado na investigação científica numa tentativa muito rudimentar, classificando-a segundo padrões de ativação neuronal onde

conseguimos ‘apenas’ identificar algumas estruturas e formações cerebrais necessárias para que uma pessoa possa de facto estar consciente – os chamados ‘correlatos neuronais da consciência’ (NCC).

Mesmo assim, é duvidoso que esta conceção de consciência seja a que de facto interesse aos filósofos em geral e, particularmente, aos fenomenólogos. Existe, pela existência de uma certa ignorância epistémica por parte de grande parte dos neurocientistas, uma atitude bastante pouco humilde em relação à dificuldade que o fenómeno da consciência – isto é, o fenómeno da experiência subjetiva e de saber porque ela existe de facto quando poderia não existir – levanta à explicação científica do mundo (cf. Chalmers, 1996: capítulo 1).

Uma segunda tese, de carácter positivo, é então avançada pela neurofenomenologia (TNP – Tese Positiva da Neurofenomenologia) e divide-se em três ideias fundamentais:

(TNP)

(3) O programa de investigação de fenomenologia é necessário, mas não suficiente;

(4) O programa de investigação oferecido pela Neurociência é necessário, mas não suficiente;

(5) Considerando os programas numa perspetiva multidisciplinar, poderemos encontrar um projeto de investigação que seja suficiente para lidar com os problemas difíceis da mente (proposta da Abordagem da Neurofenomenologia).

Vejamos de que forma esta tese positiva é justificada. A ideia (3) é afirmada por grande parte da história da fenomenologia e dos seus protagonistas que observem falhas nas suas obras tardias às suas teses principais (e.g. Husserl ou Merleau-Ponty). A ideia (4) é também afirmada pela história da Ciência, principalmente pela dificuldade da investigação contemporânea sobre o cérebro e a consciência (como vimos no capítulo 2). Por fim, a ideia (5) é essencialmente a proposta da Neurofenomenologia: coligar os dois métodos que parecem, na sua essência, dissemelhantes, num único método que consiga oferecer um programa de investigação coerente e que, de forma rigorosa e científica, possa criar progresso.

De modo a focar a nossa atenção e não alargar o escopo do tema, apresentaremos de seguida a análise do fenómeno da percepção e de uma leitura neurofenomenológica da mesma.

Um primeiro pressuposto – e inovação – de uma aplicação da fenomenologia à Neurociência é o seguinte: a percepção – e poder-se-á colocar qualquer processo mental no lugar deste – é corpórea (*embodied*). Tal significa o seguinte: contra grande parte da Neurociência atual, a investigação da percepção pelas técnicas neurocientíficas do estudo do cérebro não é suficiente para alcançar uma

explicação total do fenómeno. Ou seja: dizer que para compreender a percepção bastará estudar o cérebro no laboratório de Neurociência é cometer uma falácia. Esta falácia, além de ser logicamente problemática – o que interessa à fenomenologia – é cientificamente problemática – o que interessa à Neurociência.

Deste modo, um neurofenomenólogo defende que a percepção é um processo corpóreo. Tal significa simplesmente aceitar que o *design* funcional e estrutural do nosso corpo forma ou influencia a maneira como percebemos o mundo – esta lição pode ser encontrada também em investigações da Neurociência e.g. o trabalho de António Damásio.

Shaun Gallagher resume assim o erro da visão tradicional da percepção: “Pensar que a percepção é algo que acontece apenas no cérebro é ignorar a contribuição dos préprocessos sensoriais corpóreos” (Gallagher, 2008: 164). Ademais, o filósofo oferece um breve exemplo com a percepção auditiva: “No caso da sensação auditiva, por exemplo, a forma e a localização das orelhas determina a direção da informação, amplificando ou filtrando *inputs* específicos” (Gallagher, 2008: 164).

Obviamente, muitos destes aspetos corpóreos operam sob o nosso inconsciente, de forma automática, sendo irrelevantes para a nossa percepção consciente. Mas estes processos determinam efetivamente de que forma é que percebemos do mundo – o ritmo cardíaco, a tensão muscular, a pressão sanguínea, a respiração, etc. É assim necessário dar atenção ao carácter holístico que os processos mentais devem pressupor nas nossas investigações.

Como veremos com mais detalhe no capítulo 7, Georg Northoff (2018) mostra que existe um processo de acoplamento (*coupling*) entre a atividade neuronal do cérebro, o ritmo cardíaco e a atividade do estômago: um cérebro saudável tenta alinhar as fases da sua atividade espontânea com os batimentos cardíacos e, por outro lado, a sua amplitude é acoplada ao início da fase do estômago. Muitas doenças do foro psiquiátrico podem ser descritas como uma falta de acoplamento entre a atividade cerebral do sujeito e a captação dos *inputs* do mundo exterior (cf. Northoff, 2018: 201-202).

Um segundo pressuposto é assumir que a percepção é situada no ambiente (*embedded*). A ideia de que a nossa cognição é situada foi já enunciada por pragmatistas e fenomenólogos. Por exemplo, como vimos em Heidegger, o tipo de ser que é capaz de ter relações intencionais com o mundo é um ser que já está no mundo – claro, este estar é mais básico do que estar localizado espacialmente no mundo. Nas palavras de Gallagher:

Para se adaptar a análise heideggeriana de cognição corpórea, pode-se dizer que o organismo simplesmente se encontra profundamente situado num ambiente apresentado como uma possibilidade em vez de outro. Pelo contrário, é parte da própria natureza da

existência humana estar no ambiente, esta é uma de suas características necessárias (Gallagher, 2008: 165).

Merleau-Ponty, como vimos, também defendeu algo semelhante: o que encontramos no processo de percepção não é um sujeito passivo que observa os objetos em seu redor (veremos, no Capítulo 8, uma proposta avançada recentemente em Neurociência cuja tese essencial se baseia exatamente nesta ideia). Como é sugerido pelo filósofo francês, todo o tipo de reflexão que possamos fazer de nós próprios e da nossa atividade no mundo tende a ser situada no mundo, e não é apenas uma mera abstração introspectiva ou mental.

Veja-se o seguinte exemplo prático: alguém não concebe primeiramente o espaço através do qual precisa de guiar a sua mão para agarrar este livro que possui nas mãos. Nem a forma de o agarrar é algum tipo de representação do mesmo: o próprio agarrar não seria possível se a sua mão não estivesse “já situada numa linha de ação” (Gallagher, 2008: 166).

O terceiro pressuposto é o seguinte: a nossa percepção, ou mais em geral, a nossa mente, é expandida (*extended*). Gallagher indica o trabalho de David Chalmers e Andy Clark afirmando que “a cognição envolve a manipulação de elementos portadores de informações do sistema, parte da qual está ‘lá fora’ no mundo de modo que a cognição não acontece ‘apenas na cabeça’” (Gallagher, 2008: 169).

O artigo onde é apresentado este terceiro pressuposto abre com a seguinte afirmação:

A tendência geral do pensamento humano está altamente influenciada por suportes ‘ambientais’. Considere-se por exemplo o uso de uma caneta e de um papel para fazer uma conta longa de multiplicação, ou a parafernália geral da linguagem, livros, diagramas e cultura. Em todos estes casos, o cérebro individual executa algumas operações, enquanto que outras [*sc.* operações] estão delegadas à manipulação de meios externos (Clark & Chalmers, 1998: 7).

Um exemplo ilustrativo pode ser oferecido a partir da prática cirúrgica: quando um cirurgião entra na sala de operações, a própria sala e o equipamento médico dão assistência aos processos cognitivos que estão envolvidos na cirurgia. A forma como os instrumentos cirúrgicos estão dispostos, a sua ordem e posição na mesa de operações, tudo isto permite ao cirurgião concentrar-se mais no que está a fazer do que em pensar sobre os próximos passos da cirurgia. Estes “arranjos ambientais” permitem ao cirurgião executar o procedimento eficazmente de forma intuitiva sem ter que parar para refletir no próximo passo.

Neste sentido, a perícia e o *know-how* do cirurgião são constituídos em parte pelo ambiente onde trabalha. O cirurgião consegue fazer melhor o seu trabalho por causa da forma como a sala e os instrumentos estão colocados de forma específica.

Conclui-se assim que o nosso pensamento e a nossa percepção não acontecem apenas “dentro da nossa cabeça”, mas de uma forma que é estendida ao mundo.

Finalmente, o quarto pressuposto é assumir que a percepção é enativa (*enactive*) – termo cunhado por Evan Thompson e Francisco Varela (Varela, Thompson & Rosch, 1991). Tal significa que a experiência não é determinada somente pelos estados neuronais que são ativados por estímulos sensoriais. Antes, esta depende também das habilidades sensoriomotores do observador: “A percepção não é algo que nos acontece, ou que acontece em nós. É, antes, algo que fazemos” (Noë, 2004: 1).

A ação, para os teóricos enativistas, não ocorre no cérebro. Em vez disso, a unidade da percepção é o organismo tido como um todo e que age no ambiente. Assim, a visão não é uma representação que emerge numa rede de neurónios particular, mas antes “um modo de exploração do ambiente com base num entendimento implícito das regularidades sensoriomotores” (Noë, 2004: 29-30).

Concluimos, assim, os pressupostos assumidos por um neurofenomenólogo e as suas consequências para o estudo da mente em geral. O que um neurofenomenólogo propõe e defende é que a investigação de fenómenos subjetivos não deve ser reduzida ao estudo do cérebro, mas que deve expandir-se e considerar a totalidade do corpo e da sua relação fixada num ambiente cuja interação é constante e influencia todos os processos conscientes e inconscientes.

Francisco Varela tentou formular uma nova abordagem para resolver os problemas difíceis da mente sem deixar de aceitar que a experiência da primeira pessoa é um fenómeno irreduzível.

Varela define deste modo o projeto da Neurofenomenologia:

[...] é o que designo pela missão de casar a ciência cognitiva moderna com uma abordagem disciplinada à experiência humana, enquadrando-me assim na linhagem da fenomenologia de tradição continental (Varela, 1996: 330).

O que é então a Neurofenomenologia e como pode resolver o problema difícil da consciência?

O filósofo e cientista cognitivo irá apresentar duas teses fundamentais:

(6) A fenomenologia não se reduz ao introspecionismo (algo que veremos ser uma crítica levantada ao projeto em causa);

(7) O ponto central de fenomenologia é suspender a experiência habitual do mundo e catalisar investigações novas (cf. I.1.1, este capítulo);

Para conseguirmos (7), é necessário aplicar uma série de passos metodológicos específicos que permitem garantir uma abordagem eficaz e útil. Essas características estão resumidas na seguinte Figura 4.1.:

Redução Fenomenológica	
Aspectos de método	Características resultantes do exame
<i>Atitude</i>	<i>Colocar crenças entre parenteses, suspender crenças</i>
<i>Intuição</i>	<i>Intimidade, evidência imediata</i>
<i>Invariantes</i>	<i>Inscrições, intersubjetividade</i>
<i>Treino</i>	<i>Estabilidade, pragmático</i>

Figura 4.1.: Redução fenomenológica (adaptado de Varela, 1996: 338).

Vejamos, sucintamente, os vários aspetos do método que permite alcançar a redução fenomenológica, elemento essencial de (7):

- **Atitude: redução** – o ponto de partida é o ponto de partida necessário; a ideia é colocar em suspenso todas as nossas crenças atuais sobre o mundo; o objetivo é libertar-nos do fluxo habitual da nossa mente e conseguirmos “desligar” a nossa presença situada no mundo;
- **Intimidade: intuição** – o resultado da redução é tornar mais vivo o nosso campo de experiência; devemos aplicar também variações ao tipicamente pensado, de modo a abrir e a dar espaço a novas experiências;
- **Descrição: invariantes** – ficar apenas na redução e intuição seria condenar o método a um acesso apenas privado; antes, o que foi obtido pela intuição deve ser inscrito ou traduzido de

forma a ser comunicado, usualmente pela linguagem ou outros sistemas simbólicos (como desenhos ou fórmulas); o objetivo é formular descrições públicas e acessíveis a terceiros;

- Treino: estabilidade – a pessoa tem de estar treinada para este tipo de prática; tal como um velocista tem de treinar para correr abaixo dos 10 segundos os 100 metros, a pessoa que queira prosseguir este tipo de análise deverá treinar as competências necessárias para assegurar que o método é eficaz e pode proporcionar progresso. Esta parece ser a etapa mais complicada para ser aplicada num programa de investigação seriamente científico, visto que teria de implicar o treino e formação de certas técnicas junto dos cientistas e filósofos da mente (cf. Varela, 1996: 336-338).

Estes são os passos que uma investigação neurofenomenológica deve assumir segundo Varela. Mas esta abordagem não se fundamenta demasiado no fenómeno da introspeção? Como veremos, esta será uma crítica levantada a esta abordagem: a introspeção poderá não ser a melhor forma de metodologicamente investigar os aspectos da mente consciente humana.

O aspecto interessante na abordagem de Varela é a defesa de uma redefinição do conceito tradicional de introspeção usado normalmente em Filosofia. Para o filósofo e biólogo chileno, bastará pensar na matemática para perceber o que é a introspeção – e as evidências que podemos alcançar pela intuição:

[...] em última análise, o peso de uma prova matemática é a sua natureza convincente, a imediatez com que a evidência se impõe em nós, além da corrente lógica e simbólica do raciocínio. Esta é a natureza da evidência intuitiva (Varela, 1996: 339).

Estamos, portanto, perante um redirecionar do conceito de intuição que poderá superar as críticas ao método da intuição tradicional.

Existe, na linha fenomenológica desta abordagem neurofenomenológica, um pressuposto fundamental neste projeto: a distinção dualista entre sujeito e objeto é falsa. Como tal, a consciência é definida pelo neurofenomenólogo não como privada ou um fenómeno meramente interno, mas tendo, no fim, uma existência externa no mundo.

Assim, a investigação neurofenomenológica não é uma viagem privada, visto que se destina aos outros pela validação intersubjetiva. Contra a lacuna de Chalmers, concluiu-se que a divisão entre a primeira e a terceira pessoa é, afinal, falsa.

A novidade da proposta de Varela reside numa abordagem da primeira pessoa ponderada e rigorosa – pelos passos metodológicos indicados – e que integra elementos de validação da neurobiologia, e não apenas informação de caráter heurístico.

O que acrescenta esta metodologia à Ciência Cognitiva em geral, e à Neurociência em particular? Esta proposta pode fornecer ideias interessantes sobre a estrutura e a natureza da vida mental, mas a sua maior força é que o faz tornando a nossa experiência reconhecível. Este segundo ponto é, certamente, o mais difícil para convencer a comunidade científica que tem sido influenciada por uma abordagem de inclinação funcionalista ou computacionalista.

Contudo, urge perguntar o que, para Varela, é necessário fazer para tornar a Neurofenomenologia numa prática científica.

Segundo o filósofo, é necessário aplicar duas ideias fundamentais: em primeiro lugar, é preciso reaprender a dominar a capacidade da descrição fenomenológica que foi esquecida durante a segunda metade do século XX e o início do século XXI. Não parece haver razão para que tal domínio seja diferente da aquisição de qualquer *know-how*, como aprender a tocar um instrumento ou falar uma língua nova. Qualquer pessoa que aprende algo novo traz mudanças para a sua visão quotidiana do mundo (cf. Varela, 1996: 346).

Esta é a grande tese de Varela: uma aprendizagem metódica e sustentada implica transformação, tal como tudo o que se aprende de forma sustentada. Com este treino, poderemos ter acesso a aspetos da experiência que não estão acessíveis por norma e pela nossa psicologia situada do diário rotineiro da vida. O objetivo final é superar os nossos hábitos automáticos da introspeção e experiência do quotidiano de modo a alcançarmos novos conhecimentos do nosso conteúdo mental.

O segundo desafio implica a transformação do estilo e valores da própria comunidade científica. A não ser que aceitemos que é necessária uma mudança radical na forma de fazer ciência, não poderemos esperar grandes avanços e progressos que contrariem o ciclo de falhanços que a consciência e a subjetividade têm causado à ciência contemporânea (cf. Nagel, 1974; Chalmers, 1996; Searle, 1997; Nöe, 2009). Tal implica deixar de lado uma imagem específica de como a ciência é feita, e questionar a forma como treinamos os nossos cientistas e estudantes algo que, como veremos de seguida, poderá ser um ponto problemático desta posição metodológica.

Mas mais importante do que estas duas indicações, a Abordagem da Neurofenomenologia poderá ajudar a tornar mais claro o próprio papel do cérebro na atividade consciente:

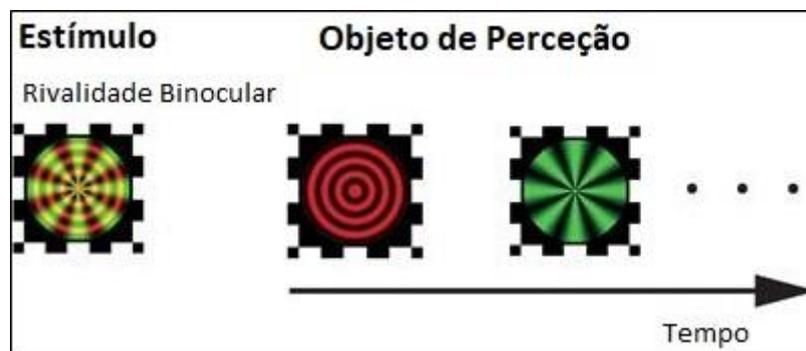
Para avançar no entendimento da consciência, é necessário abandonar o microfoco neural interno [...] O *locus* da consciência é a vida dinâmica de toda a pessoa ou animal conectado ao meio ambiente. De facto, é somente quando adotamos essa perspectiva holística da vida ativa da pessoa ou animal que podemos começar a entender a contribuição do cérebro para a experiência consciente (Noë, 2009: xiii).

Como veremos nas objeções, não é de todo claro o que tal poderá significar, quer em termos de metodologia específica, quer em termos de considerações metafisológicas explícitas. Mas antes de avançar com essa argumentação, vejamos de que forma uma investigação neurofenomenológica se sucede utilizando os pressupostos agora descritos.

IV.1.4. Um exemplo de Investigação Neurofenomenológica: o caso da “Rivalidade Binocular”

Vejamos um exemplo prático da aplicação deste método³⁸ a um fenómeno mental que procura mostrar como as teses da Neurofenomenologia podem ser aplicadas empiricamente sem ignorar os aspetos fenomenológicos da consciência.

Diego Cosmelli, um dos últimos estudantes de doutoramento de Varela, fez uma demonstração do método agora disposto, aplicando-o ao fenómeno da “rivalidade binocular” (cf. Cosmelli *et al.*, 2004: 128-140). A rivalidade binocular é o que acontece quando cada olho vê uma imagem diferente ao mesmo tempo:



³⁸ Um dos primeiros estudos que aplica diretamente o método neurofenomenológico de Varela de forma rigorosa, segundo os padrões neurocientíficos, pode ser encontrado em Lutz, 2002: 133–167.

Figura 4.2.: Exemplo do desenho experimental e estímulos (adaptado de Ye *et al.*, 2019).

Através de uma disposição espacial, as duas imagens são mantidas em separado: a imagem em formato circular é projetada apenas para o olho esquerdo, e a imagem em formato espiral para o olho direito. O que acontece é que uma das imagens irá sobrepor-se à percepção visual por uns segundos, fazendo desaparecer a outra; mas, logo a seguir, a que estava a desaparecer volta e faz a primeira desaparecer também. Sentimos, assim, a alteração das imagens que se mantêm perfeitamente estáticas na sua projeção.

O interessante desta experiência é que os sujeitos começam a perceber que podem, de certa forma, influenciar a alternância das duas imagens através da fixação de um ponto particular no centro das duas imagens, ou retendo mentalmente uma das imagens que, quando parece desaparecer, pode ser novamente percebida totalmente. Contudo, este domínio é difícil de ser mantido pois não existe qualquer padrão regular, mas apenas uma alternância espontânea e imprevisível.

Este fenómeno é conhecido dos psicólogos e cientistas da visão: em vez da nossa cognição formar uma terceira imagem estável a partir das duas imagens, o que acontece apenas é uma alternância perceptual entre as imagens (cf. Thompson, 2015: 21-22).

De facto, não precisamos de um laboratório de ponta para fazer isto: basta enrolar uma folha A4 de papel, olhar para o buraco que é criado no centro durante alguns segundos, e depois colocar a outra mão a 2 cm do outro olho e conseguiremos ver que em alguns segundos a mão parece ter um buraco no meio. Ora, o buraco aparece na mão porque cada olho tem diferentes estímulos, e os dois estímulos competem para ter o domínio perceptual. Embora o estímulo visual não mude, a nossa consciência visual desse estímulo muda; sabemos, então, que a mudança acontece em nós, e não fora de nós.

Podemos agora perguntar: “qual é a diferença no nosso cérebro quando dizemos que vemos apenas uma imagem, e quando dizemos que não vemos imagem nenhuma (embora a imagem ainda esteja presente a afetar o nosso olho e sistema visual)?”

Este fenómeno parece ser excelente para refletir sobre os correlatos neuronais da consciência. Isto porque diferencia-se entre conteúdos visuais conscientes e inconscientes. Assim, parece oferecer uma forma de dissociar a atividade neuronal correspondente diretamente à percepção visual consciente da atividade neuronal correspondente à percepção visual inconsciente – isto porque o estímulo continua presente (cf. Cosmelli & Thompson, 2007).

O que neurocientistas (e.g. Nikos Logothetis) descobriram ao estudar o cérebro de sujeitos que estavam a fazer a experiência foi o seguinte: não existe um ponto específico no cérebro – ou em qualquer

das suas partes – onde a experiência visual acontece. Antes, esta envolve variadas partes do cérebro, desde as partes frontais da área parietal, do córtex visual primário, córtex inferotemporal, entre outras (cf. Logothetis *et al.*, 1996; Logothetis, 1999: 74).

Até agora, apenas foi aplicado o estudo objetivo da terceira pessoa da Neurociência. Como incorporar então a perspectiva da primeira pessoa e da fenomenologia?

Cosmelli pediu aos participantes da experiência para descreverem o que viam durante o estímulo binocular. As suas descrições descreviam a experiência como sendo um mar de ondas; havia períodos de dominância de uma das imagens, mas rapidamente começava a alternância (como uma onda do mar); a mudança começava em diferentes pontos da imagem dominante (cf. Thompson, 2015: 29).

Ora, antes de ter em conta estes relatos privados, os neurocientistas apenas pediam às pessoas para pressionar um botão quando vissem a alteração da imagem dominante; tal dava um ponto fixo de referência no tempo, definindo o estado cerebral médio quando o participante vê uma imagem e o estado cerebral médio quando deixa de a ver. O investigador contrasta depois estas medidas médias para perceber o que é específico da percepção consciente da imagem.

Contudo, apesar de tal estudo ter proporcionado descobertas importantes, ele ignora por completo a fenomenologia da experiência, deixando de lado importantes informações que podem explicar melhor o fenómeno. Aplicando o método da Neurofenomenologia, Cosmelli concluiu que a alternância entre as duas imagens na ‘rivalidade binocular’ corresponde a ondas de oscilações sincrónicas no cérebro.

A forma como a percepção consciente varia em tempo corresponde à forma como a sincronia neuronal varia no tempo. O surgimento de um novo momento de percepção corresponde à formação de um novo padrão de sincronia neuronal e a diminuição desse momento perceptual e a sua substituição por um outro corresponde à dissolução do padrão de sincronia e à sua substituição por outro. Tanto na experiência perceptiva como na atividade cerebral, há um ritmo de momentos distintos que alternam constantemente, mas que nunca são exatamente os mesmos.

Uma consequência filosófica importante destes resultados é relatada por Logothetis da seguinte forma:

Os resultados até o momento sugerem fortemente que a consciência visual não pode ser considerada o produto final de uma série hierárquica de estágios de processamento. Em vez disso, envolve todas as vias visuais, bem como áreas parietais frontais, que estão envolvidas no processamento cognitivo superior. A atividade de uma minoria significativa de neurónios

reflete o que é visto conscientemente, mesmo nos níveis mais baixos que observamos, V1 e V2; é apenas a proporção de neurónios ativos que aumenta em níveis mais elevados no caminho (Logothetis, 1999: 74).

Ora, esta experiência mostra que uma das teses mais impregnadas da Neurociência contemporânea – a tese de que estados mentais particulares acontecem em zonas particulares e bem-definidas do cérebro – se encontra equivocada, dado que não existe um único *locus* onde a consciência visual aconteça no cérebro; antes, esta é distribuída por várias zonas relevantes do mesmo.

Deste modo, compreendemos perfeitamente como é que o programa de investigação da Neurofenomenologia funciona. Nas palavras de Thompson:

[sc. A] Neurofenomenologia combina o estudo cuidadoso da experiência “de dentro” com as investigações do cérebro e o comportamento “de fora”. Utiliza descrições da experiência direta [sc. dos sujeitos] para orientar o estudo dos processos cerebrais relevantes para a consciência (Thompson, 2015: 28).

Assim, tendo em consideração a fenomenologia da percepção, o fenómeno foi melhor explicado e compreendido cientificamente. Este método poderá ser aplicado a todos os fenómenos tipicamente trabalhados pelos filósofos e pode, de facto, causar progresso na área.

IV.1.5. O Eu (*Self*) em Evan Thompson

Evan Thompson é um dos investigadores mais ligados à investigação utilizando os métodos da Neurofenomenologia. No seu *Waking, Dreaming, Being* (2015), analisou vários fenómenos à luz do método deixado pelo seu mestre, Francisco Varela.

Especificamente, Thompson irá tentar cruzar as práticas da primeira pessoa das religiões e filosofias contemplativas da Ásia como o Yoga, Vedanta ou o Budismo com os estudos da Neurociência contemporânea da terceira pessoa, com dois objetivos principais:

- (i) Limitar o escopo das asserções metafísicas das primeiras (que argumentam erroneamente, por exemplo, que a consciência existe fora de um corpo e do cérebro);

- (ii) Orientar o método da Neurociência tradicional tentando extrair mais conhecimento de forma rigorosa e sistemática – acabamos, assim, com mais uma instanciação do projeto da Neurofenomenologia.

Um dos conceitos filosóficos mais analisados nesta obra é o conceito do Eu (*Self*): Thompson argumenta a favor da tese de que o Eu é, ao contrário de várias teorias tradicionais da Filosofia, um processo fluído e não uma entidade estática nem uma entidade ilusória:

[...] o Eu é um processo, não uma coisa ou uma entidade. O Eu não é algo exterior à experiência, oculto no cérebro ou em algum domínio imaterial. É um processo experiencial que está sujeito a mudanças constantes (Thompson, 2015: xxxi).

Igualmente importante é compreender a crítica ao programa reducionista que apresentámos no capítulo anterior. Para Thompson – e a Neurofenomenologia em geral –, a consciência como experiência subjetiva não está relacionada necessária ou unicamente ao cérebro nem às diversas funções neurobiológicas, mas somente de forma contingente: a consciência não se reduz à atividade cerebral. O argumento é o seguinte: dado que a consciência é subjetiva, não pode ser reduzida a algo não-subjetivo como o mero disparar de neurónios:

[...] dado que os modelos científicos do mundo são sempre destilações da nossa experiência consciente como observadores, não faz sentido tentar reduzir a consciência a um ou outro dos modelos científicos, incluindo os modelos da atividade cerebral (Thompson, 2015: 101).

Contudo, tal não significa que a consciência não seja um fenómeno natural e que a mesma não varie segundo determinadas funções neurobiológicas. Como faremos notar nas objeções, não se compreende exatamente em que posição poderemos colocar a teoria da consciência ou do Eu proposta por Thompson.

O grande objetivo do filósofo é demonstrar que a tese de que o Eu é uma ilusão – historicamente associada ao Budismo, no Oriente, e a David Hume, no Ocidente – é falsa. Vários investigadores contemporâneos defenderam versões diferentes desta tese: Daniel Dennett (1986), Owen Flanagan (1996) ou Thomas Metzinger (2009). No seu âmago, a posição do “não-eu” (*no-self view*) defende a negação de que existe um algo ou entidade a que podemos chamar de “eu”, sendo este, antes, uma

mera ilusão criada pela nossa neurobiologia por razões evolutivas. Nas palavras de Metzinger no seu *The Ego Tunnel*.

[...] Não existe um Eu. Ao contrário do que a maioria das pessoas acredita, ninguém nunca foi ou teve um Eu [...] considerando o melhor conhecimento atual, não existe nada, nenhuma entidade indivisível, que somos nós, nem no cérebro nem em algum domínio metafísico além deste mundo (Metzinger, 2009: 1).

Thompson irá denominar a esta tese de “neuro-niilismo” (Thompson, 2015: 322), tentando contrariá-la através de uma defesa enativista – ou mesmo neurofenomenológica – do Eu, que é considerado como um processo e não uma entidade que, de certa forma, emerge de variados processos psicológicos e neurofisiológicos, mas também sociais e intersubjetivos a que chama de “produção do Eu” (*I-making*) (cf. Thompson, 2015: 325).

Para o filósofo, há duas componentes essenciais deste processo:

- (i) Auto-designação (*self-designating*): cada indivíduo constrói-se como um sujeito individual distinto e dissemelhante dos outros indivíduos (cf. Thompson, 2015: 344);
- (ii) Auto-projecção (*self-projection*): as faculdades da memória e prospecção expandem o sentido do Eu individual para abranger o passado e o futuro imaginado, originando a sensação (*sense / feeling*) de que existe um sujeito duradouro, unitário e que executa as ações do sujeito (cf. Thompson, 2015: 347-350).

Em relação a (ii), Thompson indica que essa componente específica está dependente de zonas particulares do cérebro como o lobo frontal, associado ao planeamento, as áreas temporal-parietais mediais, associadas à memória, sendo que existe uma justaposição entre esta componente e a *Default Mode Network* (RMP)³⁹ (cf. Thompson, 2015: 349).

Para Thompson, estas duas componentes ajudam a criar mentalmente um Eu fluído que diminui quando estamos a adormecer e desaparece quando entramos em sono profundo (sem sonhos), mas que volta a reaparecer quando sonhamos e estamos acordados. Para o filósofo, o facto de o Eu ser uma

³⁹ Veremos a importância desta rede particular no Capítulo 7 quando analisarmos o conceito de percepção de uma perspectiva neurofilosófica não-reduitiva.

construção mental não implica que seja uma ilusão, ao contrário de, por exemplo, o que defende Metzinger. Mas não será tal uma contradição?

A resposta de Thompson é clara. O facto de o Eu ser uma projeção mental não faz deste uma ilusão sem existência. Antes, o Eu deve ser visto como uma imagem no espelho: a imagem depende da existência do espelho sendo a sua base, mas a sua existência (da imagem) não é a mesma que a do espelho nem é composta pela natureza do espelho. Pelo contrário, a imagem apenas existe em relação a um observador que a percebe: apesar da imagem do espelho ser dependente de um observador, tal não significa que seja uma ilusão subjetiva (cf. Thompson, 2015: 364-365).

Thompson admite, contudo, que há um elemento ilusório no processo que dá origem ao Eu:

A ilusão – ou a decepção (*delusion*) – está em considerar o Eu como possuindo uma existência independente, como considerar que a imagem do espelho existe realmente no espelho. Note que a imagem como tal não é uma ilusão; antes, é considerar que a imagem existe no espelho que é a ilusão. Da mesma forma, não é a aparência do Eu como tal que é a ilusão; é considerar o Eu como existindo independentemente, essa é a ilusão (Thompson, 2015: 365).

A estratégia de Thompson para refutar a concepção “neuro-niilista” do Eu como mera ilusão passa por uma analogia interessante da história da Ciência: o átomo era definido como uma entidade indivisível nas concepções antigas (por exemplo, em Lucrecio). Ora, o facto de termos descoberto que o átomo era divisível em subpartes atómicas – como eletrões, neutrões, positrões, etc. – não implica que os átomos não existam, ou seja, que são ilusórios: apenas requer que a definição antiga de átomo seja revista para permitir uma nova definição.

O plano de Thompson para o Eu é o mesmo: o facto de descobrirmos que não existe uma substância ou entidade que existe permanentemente não implica negar a existência do Eu, mas, diferentemente, leva à sua redefinição, o que permite descrever o mesmo como um processo e não com uma coisa. Assim, a grande lição para o estudo do Eu desta abordagem neurofenomenológica é avançar com uma nova definição do Eu que não se foque meramente no conteúdo do que é ser um Eu, mas, pelo contrário, e ecoando as teses orientais, que tome o Eu como um veículo que “carrega” esse mesmo conteúdo, isto é, o sentido de ser tal Eu (cf. Thompson, 2015: capítulo 10).

IV.1.6. Três lições pragmáticas para uma Neurofenomenologia rigorosa

Para concluir a análise desta abordagem, iremos apresentar um artigo muito relevante para compreender o alcance e as diferenças que é necessário introduzir na investigação empírica da mente humana em relação a abordagens mais tradicionais da Ciência Cognitiva e da Neurociência em geral.

No artigo “Lições metodológicas em neurofenomenologia: revisão de um estudo de base e recomendações para abordagens de pesquisa”, Patricia Bockelman, Lauren Reinerman-Jones e Shaun Gallagher apresentam três lições que é necessário ter em conta se se quiser avançar com uma investigação fundamentalmente neurofenomenológica que se diferencie das abordagens mais tradicionais de investigação.

IV.1.6.1. Reformulação da Ciência

A primeira lição é que a Neurofenomenologia deve fazer uso de um “modelo mental compartilhado” (MMC) devido à natureza essencialmente interdisciplinar desta abordagem. Esse modelo passa por requerer a colaboração de investigadores especializados em diferentes abordagens de modo que os dados objetivos oferecidos pela perspectiva da terceira pessoa não sejam assumidos sem o contexto complexo de onde são recolhidos, analisados e interpretados (cf. Van Rijnsoever & Hessels, 2011).

Mais importante, é necessário combater um vício metodológico que acontece usualmente na ciência tradicional em que uma equipa multidisciplinar trabalha: o facto de não haver cruzamento de conhecimento, linguagem e métodos entre os próprios investigadores. Pelo contrário, a Abordagem da Neurofenomenologia requer que todos compreendam o processo e o trabalho de cada membro do grupo de investigação em causa (cf. Bockelman, Morrow & Fiore, 2013).

Esta perspectiva procura que cada investigador faça uso de um léxico partilhado através de técnicas específicas que é necessário colocar em prática para que todos os investigadores de diferentes áreas de especialização possam estar no mesmo plano conceptual – algo que contraria a forma usual de fazer ciência, em que cada investigador tem a sua especialização cuja idiosincrasia metodológica e conceptual não permite um diálogo eficaz com os restantes especialistas (cf. Cannon-Bowers *et al.*, 1993).

Nas palavras dos autores:

Logisticamente, [sc. a implementação do (MMC)] leva tempo. Agendar reuniões regulares para *brainstorming* e ensino, bem como criar um glossário e um documento de estrutura conceptual, são práticas e ferramentas simples para a implementação bem-sucedida da neurofenomenologia. O resultado não deve ser apenas uma colaboração produtiva no desenvolvimento experimental, mas também levar a resultados sintetizados, e não partes individuais na interpretação da perspectiva de cada domínio. Em vez disso, os resultados são a totalidade ou uma grande parte do puzzle (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 6).

Assim, a primeira recomendação metodológica da Neurofenomenologia é alterar de forma substancial a pragmática da prática científica. É claro, é necessário também compreender os resultados da aplicação de um método como (MMC) em termos pragmáticos, algo que, dada a juventude desta abordagem, está ainda em falta. Como veremos na reflexão crítica, estas ideias poderão parecer úteis à primeira leitura, mas poderão ser inconsequentes para uma sólida defesa metodológica das relações entre a Filosofia e a Neurociência.

IV.1.6.2. Melhor Ciência

A segunda lição pragmática que a Abordagem da Neurofenomenologia tem de considerar é a forma como, muitas vezes, o conceito de “interdisciplinaridade” é usado: com pouco rigor científico, sem considerações empíricas claras e com consequências teóricas e práticas liliputianas.

Pelo contrário, esta abordagem deve ter “elevados padrões de *design* e execução experimental para alcançar controlo variável, fiabilidade, generalização e replicação dos resultados” (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 6). Esta consideração advém do problema de trabalhar numa equipa multidisciplinar em que cada investigador apenas se foca na sua área particularizada de especialização: se não houver um diálogo interlinguístico entre os diferentes membros numa linguagem comum, não se conseguirá notar claramente qual é a visão ou quadro geral da investigação em causa.

Grande parte das abordagens neurofenomenológicas que observámos fizeram realmente uso de (i) medições neurológicas e fisiológicas para avaliar a resposta física da experiência do paciente; (ii) questionários de psicologia para avaliar fatores diferentes da experiência do paciente (e.g. demografia,

peculiaridades, respostas cognitivas, etc.) e (iii) entrevistas fenomenológicas que ligavam a experiência à prática e uso linguísticos de determinados conceitos para descrever essa mesma experiência.

Ora, esta abordagem, embora seja neurofenomenológica na sua essência, dado que utiliza vários métodos para entender um determinado fenômeno, poderá não ser uma metodologia de excelência, dado que, apesar de haver uma abordagem multidisciplinar presente em (i), (ii) e (iii), falta o mais relevante que uma abordagem da Neurofenomenologia deve oferecer: o cruzamento, afinação, balanço e troca entre as metodologias (i), (ii) e (iii), de modo a haver uma influência constante que otimize cada parte da investigação para que a interpretação dos resultados da experiência não seja limitada às abordagens reducionistas tradicionais:

A interação entre o projeto fenomenológico e algumas técnicas da ciência cognitiva pode gerar uma metodologia integradora única. Como tal, a neurofenomenologia deve tirar o melhor das práticas da ciência cognitiva, contribuindo ao mesmo tempo com técnicas únicas que não fazem parte, usualmente, da caixa de ferramentas da ciência cognitiva (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 7).

Assim, é preciso ter em consideração que é necessário integrar a fenomenologia de forma rigorosa na prática da melhor ciência que se possa apresentar, tendo sempre em foco o rigor experimental e o objetivo de criar estudos empíricos que possam ser replicáveis: no fundo, é preciso aliar a fenomenologia às considerações da melhor ciência que pode ser praticada.

IV.1.6.3. O Foco no Entrevistador

A terceira lição metodológica particulariza uma parte substancial da abordagem da Neurofenomenologia: a entrevista fenomenológica. Usualmente, as entrevistas que fazem uso da primeira pessoa (e.g. Psicologia) para descrever alguma experiência exigem um foco peculiar na habilidade do indivíduo entrevistado. Contudo, nesta abordagem, o papel – e o treino específico – mais relevante está no entrevistador: é este que deve ser treinado para conseguir guiar o entrevistado a uma descrição mais precisa ou relato experiencial de algum fenômeno subjetivo.

A grande diferença com as formas tradicionais utilizadas na Ciência Cognitiva é a relação entrevistador-entrevistado: convencionalmente, há uma linha que separa exactamente um do outro,

sendo que a influência do primeiro sobre o segundo é vista como uma má prática que levará a resultados enviesados. Todavia, na abordagem neurofenomenológica, o contrário acontece: a linha atenua-se, e o entrevistador participa ativamente na reflexão e articulação da experiência do entrevistado (cf. Bockelman, Reinerman-Jones & Gallagher, 2013: 7).

É importante notar uma diferença interessante entre diversos autores da abordagem neurofenomenológica. Varela e Shear, por exemplo, defendem que esta abordagem apenas funciona com pessoas treinadas na arte da introspeção – no sentido fenomenológico – que podem oferecer uma elevada precisão relevante para alcançar resultados úteis (cf. Varela & Shear, 1999). No sentido contrário, Petitmengin usa informação de participantes sem treino (cf. Petitmengin, 2006).

Para os autores, o que estas perspectivas opostas demonstram é que existe um contínuo – e não algo fundamentalmente binário – entre os participantes altamente treinados e aqueles sem qualquer treino ou habilidade introspectiva. Esta informação convoca a ideia de que é necessário adaptar o nível do entrevistador consoante o nível do entrevistado: para entrevistados com menos capacidade introspectiva, deve ser alocado um entrevistador mais competente e treinado; pelo contrário, para entrevistados com alta capacidade introspectiva, um entrevistador menos treinado poderá ser suficiente.

Nas palavras dos autores:

Se o entrevistador estiver a trabalhar com monges budistas, não precisará de muito treino e poderá simplesmente dizer ao participante o foco do estudo. Por outro lado, se o mesmo entrevistador estiver a trabalhar com estudantes de licenciatura em qualquer universidade do Ocidente, pode ser necessário extrair uma coleção de ferramentas e técnicas para dar ao participante a capacidade de aceder aos pensamentos e sentimentos experienciados (Bockelman, Reinerman-Jones & Gallagher, 2013: 7).

Embora a Psicologia Cognitiva tradicional faça também uso de entrevistas, o objetivo das mesmas não é aceder à descrição da própria experiência do sujeito entrevistado. No entanto, para a abordagem da Neurofenomenologia, a finalidade do entrevistador não é ser meramente um espetador passivo e objetivo da perspectiva da terceira pessoa (como sugere Dennett, 1991), mas é auxiliar o entrevistado a “abrir-se para a sua própria experiência” (cf. Bockelman, Reinerman-Jones & Gallagher, 2013: 7).

Para tal, evita-se uma “falsa” objetividade que a psicologia tradicional tende a procurar na sua metodologia:

A ciência cognitiva tradicional geralmente tenta forçar o participante a um molde objetivo de estruturas conceituais pré-projectadas. Por outro lado, as interações da entrevista permitem uma mudança de direção que permite que o entrevistador participe temporariamente da subjetividade do participante (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 8).

Assim, concluímos a análise da abordagem da Neurofenomenologia com especificações precisas da forma como esta deve ser aplicada na investigação de diferentes fenómenos sem perder o rigor e qualidade do método científico. É compreendida a tentativa metodológica de cruzar a perspetiva da fenomenologia com o método científico – algo que, como veremos, poderá ser um ponto fraco desta abordagem devido às suas suposições metafilosóficas.

Nas palavras de um dos fundadores da abordagem naturalista da fenomenologia defendida pela neurofenomenologia, esta abordagem cruza a fenomenologia naturalizada de Husserl com uma teoria científica da cognição que preenche a lacuna entre a “mente fenomenológica e o cérebro” (cf. Petitot *et al.*, 1999: xiii). Se esse projeto virá a ser bem-sucedido, é uma questão que será analisada de seguida nas objeções.

IV.2. Reflexão Crítica:

Depois de termos refletido sobre as várias dimensões da abordagem da Neurofenomenologia, focaremos agora a nossa atenção nas principais críticas e problemas desta abordagem metodológica.

IV.2.1. A Objeção da “Lacuna Explicativa” (*Explanatory Gap*)

Uma primeira objeção que pode ser levantada à abordagem da Neurofenomenologia é o problema filosófico que esta perspectiva procura resolver: a lacuna explicativa entre a ciência objetiva e a existência da consciência subjetiva (cf. Levine, 1983).

Este problema particular – a lacuna explicativa – foi apresentado em detalhe por David Chalmers na sua obra *The Conscious Mind* (1996). Chalmers formulou o famoso problema difícil da consciência: como é que podemos estudar um fenómeno que parece ser irreduzível a qualquer lei da natureza? Como é que podemos ligar uma perspectiva da primeira pessoa com uma perspectiva da terceira pessoa?

O problema surge exatamente por se assumir que a ciência, como empreendimento que recusa a subjetividade e afirma a objetividade, não parece conseguir lidar com a existência da mente consciente: as descrições físicas não caracterizam adequadamente tudo o que pode ser aplicado à experiência ou consciência fenomenal (cf. Nagel, 1974; Jackson, 1982).

Ora, para um neurofenomenólogo, esta tese é um total equívoco porque confunde certas categorias. Assim, irá recusar a mesma dado que comete uma falácia. Especificamente, recusa a ideia de que os predicados “objetivo/subjetivo” estejam a ser aplicados no mesmo sujeito. O argumento⁴⁰ pode ser formulado da seguinte forma:

(8) A ciência é objetiva;

(9) A consciência é subjetiva;

(10) Logo, não pode haver uma ciência da consciência.

Este argumento, que é apresentado por parte da comunidade científica e por críticos da abordagem naturalista, comete um erro lógico: a falácia da ambiguidade. Este tipo de falácia ocorre quando existe

⁴⁰Argumento e explicação em Searle (2015a).

uma falta de rigor e clareza no uso de certas palavras. Essas palavras podem ser pouco claras por serem ambíguas – e possuírem vários significados – ou por serem vagas – e não possuírem um significado específico. Ora, no argumento que conclui (10) os conceitos que devem fazer soar o alarme são “objetivo” e “subjetivo” dado que ambos se referem a domínios diferentes, apesar de não parecer sem uma análise mais detalhada.

Analisemos assim as premissas: a premissa (8) refere-se ao domínio da epistemologia – é assumido que o método científico é objetivo no sentido em que é imparcial, não é influenciado pelas crenças do cientista e respeita os resultados das experiências científicas. Mas a premissa (9) refere-se a outro domínio, nomeadamente ao domínio ontológico: é claro – pelo menos para os defensores da irredutibilidade da consciência – que a consciência é subjetiva. Mas tal não implica que não possa ser investigada objetivamente. Tal acontece porque podemos referir esses atributos a domínios diferentes de investigação: o domínio epistemológico e ontológico.

Veja-se a seguinte afirmação epistemologicamente objetiva: “Heidegger nasceu em 1889” – tal não depende do sujeito que formula a afirmação. Esta proposição é verdadeira se for afirmada por um contemporâneo de Heidegger, ou por nós mesmos neste preciso ano. A particularidade situacional e psicológica do sujeito que afirma a frase não é relevante para se saber o valor de verdade da mesma – isto é, a verdade da proposição é independente do sujeito que a afirma.

Pelo contrário, uma afirmação epistemologicamente subjetiva pode ser algo como: “Heidegger é o maior filósofo de sempre” – a veracidade desta afirmação depende claramente de quem a formula. Para muitos, a filosofia heideggeriana não passará de uns aforismos pseudoprofundos, pouco claros e sem rigor, que nada dizem sobre o mundo real. Para outros, Heidegger chegou onde nem a melhor ciência poderia ter chegado. E não há nenhum problema neste debate: de facto, o valor da proposição é subjetivo e depende de quem a afirma.

O mesmo acontece no domínio ontológico: os gatos, as bolas de basquetebol e as peças de xadrez têm uma ontologia objetiva – subsistem por si mesmos, e não parecem precisar de outra coisa para existir. Pelo contrário, as dores, a consciência ou a alegria parecem ter uma ontologia subjetiva – dependem de outra coisa (nomeadamente, de alguém que as tenha, pessoas ou animais) para existirem. É por este motivo que não observamos dores (e somente dores, sem o sujeito que as possui) por aí.

Deste modo, o argumento pode ser reformulado assegurando a sua validade da seguinte forma:

(11) A ciência é (epistemologicamente) objetiva;

(12) A consciência é (ontologicamente) subjetiva;

(13) Logo, podemos ter uma ciência (epistemologicamente objetiva) da consciência (ontologicamente subjetiva).⁴¹

Estando este pressuposto assumido, abrimos caminho para apresentarmos uma primeira abordagem a uma ciência da consciência.

Ora, Chalmers parece fazer notar que a ciência actual não pode lidar com a consciência e que é necessário um “ingrediente extra” (cf. Chalmers, 1995). Como já fizemos notar, a existência desta lacuna é tida pela Neurofenomenologia como falsa, sendo esta abordagem uma solução para a própria lacuna.

Como Chalmers, Varela aceita que a consciência é irreduzível. Mas contra Chalmers, o neurofenomenólogo não acredita que é necessário um “ingrediente extra” para resolver a lacuna:

[...] em vez de encontrar “ingredientes extras” para explicar como a consciência emerge da matéria e do cérebro, a minha proposta reformula a questão de encontrar pontes significativas entre os dois domínios fenomenais irreduzíveis. Nesse sentido específico, a neurofenomenologia é uma solução potencial para o problema difícil da consciência, lançando uma luz totalmente diferente sobre o que significa “difícil” [...] O que é necessário são precisamente as estruturas de conexão fornecidas pela neurofenomenologia, pois ambas são imediatamente pertinentes à experiência (pela sua própria natureza) e, ao mesmo tempo, suficientemente intersubjetivas para servir como contrapartes construtivas para análises externas (Varela, 1996: 340-341).

Deste modo, é necessário, para Varela, redefinir a forma como a ciência ilumina a experiência consciente: não há somente uma direção unilateral entre as duas partes; pelo contrário, é necessário perceber que essa relação pode ser bilateral – algo que é profundamente ignorado pela comunidade científica, segundo o filósofo chileno.

O que pode a intuição fenomenológica fornecer à ciência? Esta pode ter dois contributos essenciais: (i) sem o aspeto fenomenológico, a qualidade da experiência desaparece simplesmente ou é tomada como um enigma misterioso inultrapassável; (ii) com o aspeto fenomenológico, podemos

⁴¹ Poderá negar-se este raciocínio defendendo um ceticismo profundo em relação a qualquer tipo de ontologia positiva, defendendo que nada possui uma existência ontológica subjetiva, como por exemplo o idealismo de Berkeley em relação ao mundo externo. Embora tal seja possível, não é essa a linha seguida nem pela Abordagem da Neurofenomenologia, nem pela abordagem que achamos ter maior potencial para refletir a relação metodológica entre a Filosofia e a Neurociência (cf. Capítulo 5).

influenciar a própria investigação empírica e fornecer constrangimentos ou inovações à mesma (cf. Varela, 1996: 343).

Assim sendo, a lacuna explicativa não parece ser um obstáculo sólido contra esta abordagem particular. Veremos, de seguida, os pressupostos metafilosóficos desta abordagem. Antes, iremos analisar a objeção que põe em causa um dos fundamentos principais desta perspectiva: a validade da introspeção.

IV.2.2. A Objeção da “Introspeção”

Daniel Dennett apresenta uma estranha obsessão que a história da Filosofia parece possuir em relação a várias perguntas. Descartes questionou-se sobre como é que seria possível para um sujeito compreender se um pensamento que possui é verdadeiro ou falso, sonho ou realidade? De seguida, Kant evoluiu a questão para uma outra: como é que é possível que algo seja um pensamento de um sujeito? Isto é: quais são as condições de possibilidade de existir experiência de todo (cf. Dennett, 2001)?

Ora, para Dennett, estas questões pressupõem uma tese fundamental que é, no mínimo, controversa, e, no máximo, problemática: os dados da primeira pessoa (pensamentos, experiência, sensações, etc.) são de algum modo misteriosos. Assim, o filósofo americano acha que devemos trocar a primeira e a segunda questões, por uma terceira questão, oferecida por Alan Turing: como poderíamos construir um robot que tivesse pensamentos, que aprendesse através da ‘experiência’ – interagindo com o mundo – e usasse o que aprendeu da mesma forma que o fazemos?

Dennett afirma, assim:

A grande contribuição de Turing foi mostrar que a questão kantiana pode ser refeita por uma questão de engenharia. Turing mostrou-nos como podemos trocar a perspectiva da primeira pessoa de Descartes e Kant pela perspectiva da terceira pessoa das ciências naturais e responder a todas as perguntas! (Dennett, 2001: 455)

Deste modo, o que é pressuposto é a compreensão de que a ciência contemporânea já faz uso de uma metodologia que consegue conciliar as duas perspectivas que parecem opostas: através do método da heterofenomenologia.

A heterofenomenologia pode ser concebida como o caminho neutro entre a objetividade das ciências físicas – e da insistência das mesmas no ponto de vista da terceira pessoa – e um método de descrição fenomenológica que poderia (em princípio) fazer justiça às experiências mais privadas e subjetivas, sem ter de abandonar os princípios metodológicos da ciência (cf. Dennett, 1993: 72).

Mas por que razão a fenomenologia, só por si, não é suficiente? Para Dennett, há duas grandes razões que justificam as dúvidas sobre a eficácia problemática da mesma:

(14) Falso Positivo: algumas crenças que os sujeitos possuem sobre os seus estados mentais são provavelmente falsas e, portanto, o que precisa de uma explicação nesses casos é uma etiologia da falsa crença (a existência de um campo visual uniforme e coerente, por exemplo);

(15) Falso Negativo: alguns dos nossos estados psicológicos não são identificados por nós e, portanto, muito do que acontece no nosso cérebro não é conscientemente detetado.

Facilmente podemos pensar em casos de (14): as ilusões perceptivas (cf. Capítulo 7) são um exemplo comum do nosso dia-a-dia em que formamos uma ideia de algo que, na realidade, não existe da forma como conscientemente a interpretamos ou representamos. O mesmo acontece com (15): grande parte da nossa atividade mental é inconsciente, como por exemplo, os aspetos externos da perceção que a nossa visão ignora em detrimento de outros aspetos que o organismo acha mais relevante focar – apesar de tal influenciar a nossa visão consciente do mundo.

Assim, a heterofenomenologia incorpora uma neutralidade interessante: caracteriza o mundo heterofenomenológico de uma pessoa, sem fazer qualquer julgamento, e depois procura investigar o que poderá explicar a existência dessas crenças – o foco é retirar do sujeito uma autoridade epistémica às suas crenças e estados mentais.

Será a heterofenomenologia uma ferramenta que poderá ser usada pela neurofenomenologia? Apesar de ser um método que parece aproximar as duas perspetivas, muitos consideram que a heterofenomenologia acaba por deixar as próprias experiências conscientes de fora, focando-se apenas nos seus relatos verbais e análises linguísticas das mesmas. Os defensores deste tipo de método são até acusados de pretender eliminar o próprio objeto que supostamente querem investigar e explicar, como por exemplo, Daniel Dennett (1991) no seu *Consciousness Explained* ou Patrícia Churchland (1986), como vimos, no seu *Neurophilosophy*.

Mas mais relevante que este ponto, a heterofenomenologia de Dennett critica a noção de “introspeção” avançada pela neurofenomenologia devido a (14) e (15). Para Dennett, os dados alcançados pela introspeção sobre a consciência têm o mesmo valor que uma ficção artística: é totalmente imune a refutação – não podemos estar errados numa ficção porque o valor de verdade da mesma não é extensivo ao valor de verdade do mundo natural⁴² (cf. Dennett, 1991: 81-94).

Mas será que (15) é de facto real? Algumas evidências apontam para tal: grande parte das pessoas estão enganadas sobre a natureza da sua experiência alcançada pela visão periférica. Grande parte dos indivíduos afirma experienciar uma grande e generalizada clareza visual no arco da visão perceptiva. Todavia, ao contrário do que relatam – baseados na sua introspeção – grande parte dos indivíduos possuem apenas clareza visual numa região muito restrita do arco em questão, experienciando de facto somente uma parte reduzida do campo de visão (cf. Blackmore, 2002; Schwitzgebel, 2011).

Ademais, Dennett ataca em geral a fenomenologia por não haver uma única orientação metodológica entre os seus defensores: “A fenomenologia não conseguiu encontrar um método único estabelecido com o qual todos pudessem concordar” (Dennett, 1991: 44). Esta crítica, contudo, não parece ser sólida dado que o mesmo poderia ser afirmada em relação à ciência: há diversas técnicas e métodos considerando as diversas ciências naturais, humanas ou sociais. Tal descrição metodológica pluralista não parece ser uma objeção demolidora contra o facto de que a fenomenologia tem vindo a ter diversas especificações ou instanciações, tal como a diversidade de técnicas e formas de fazer ciência não implicam que esta não tenha valor.

Efetivamente, muitas pessoas justificam determinado comportamento apelando exatamente a um pensar introspetivo. Ora, investigações em Psicologia Social têm demonstrado que as explicações que as pessoas oferecem sobre o seu próprio comportamento pouco têm a ver com confiar na introspeção, mas antes procuram espontaneamente confluir o seu comportamento como uma hipótese explanatória que justifique o seu comportamento nas circunstâncias observadas.

Além disso, é muitas vezes demonstrado que as pessoas estão erradas, visto que o relato introspetivo dado parece ter uma função no todo das características externas da experiência, características que estão sob controlo dos cientistas. Na visão deste estudo, o que passa por relato introspetivo é, na realidade, apenas a expressão espontânea de alguém a teorizar sobre os motivos e razões que teve para

⁴²Dennett é, neste ponto, filosoficamente ingénuo: há um intenso e rico debate sobre o valor normativo e epistémico das ficções que parece ignorar totalmente, como por exemplo em Lewis (1978: 37-46), Levinstein (2007: 64-5) ou Sainsbury (2010).

ter um determinado comportamento, sendo que tal irá formar uma teoria que será baseada exatamente nas mesmas evidências externas que são acessíveis ao público em geral (cf. Nisbett & Wilson, 1977).

Ora, será que a crítica de Dennett à introspeção faz sentido? Para Varela, a resposta é claramente negativa dado que os críticos da introspeção confundem dois pressupostos que deveriam ser distinguidos e não confundidos entre si. É necessário distinguir entre o aspecto fenomenológico propriamente dito e o aspecto do “introspeccionismo”: os dois partilham a ideia de que a capacidade reflexiva é útil, mas por razões dissemelhantes.

Para a fenomenologia, o introspeccionismo é somente parte da “atitude natural” estando orientado por pressupostos que devem ser colocados em causa através da suspensão do juízo. Mas para a abordagem neurofenomenológica, a introspeção deve ser entendida não neste primeiro sentido, mas num segundo: não como uma introspeção acrítica – que é o tipo de introspeção que se encontra no introspeccionismo e criticada por Dennett –, mas, antes, como forma de suspender o juízo rápido e enviesado sobre as crenças das próprias pessoas (cf. Varela, 1996: 338).

Ora, é por isso que esta objeção não parece ser demolidora: Dennett está a criticar um aspecto ou uma definição particular de “introspeção” que não é a assumida pela Neurofenomenologia: o introspeccionismo pressupõe, por exemplo, uma dualidade sujeito-objeto que é falsa para estes autores. De certa forma, esta objeção não funciona por duas razões: (i) trata-se de uma falácia do espantinho e (ii) o próprio Varela concordaria com a mesma exactamente por causa de (i): o ataque ao tipo de introspeção que Dennett critica é aceite pela abordagem neurofenomenológica.

Como já fizemos notar inicialmente neste capítulo, a abordagem neurofenomenológica faz uso de uma redefinição de “introspeção”: esta não é um mero “olhar para dentro” enviesado de um sujeito passivo para um objeto. Antes, a introspeção é vista como uma habilidade primária que é constantemente utilizada no nosso dia-a-dia, mas que precisa de ser refinada:

[...] [*sc.* A Introspeção] é, pelo contrário, uma habilidade humana básica que opera constantemente na vida quotidiana, e que tem sido amplamente discutida nos estudos da criatividade. Pense em matemática: em última análise, o peso de uma prova [*sc.* formal] é sua natureza convincente, o imediatismo das evidências que nos são impostas, além das cadeias lógicas do raciocínio simbólico. Essa é a natureza da evidência intuitiva: nascida não do argumento, mas do estabelecimento de uma clareza que é totalmente convincente (Varela, 1996: 339).

Assim, a Neurofenomenologia procura redefinir o pressuposto de que a introspeção funciona como uma viagem privada à mente de cada um: antes, a introspeção deve ser entendida como uma atividade mental feita por um ser corpóreo, embutido no mundo de interações com objetos desse mesmo mundo e com outros indivíduos e que é reproduzida intersubjetivamente.

É partindo da redefinição desta noção que, para Varela, a lacuna apresentada por Chalmers deixa de fazer sentido: a oposição entre as perspectivas da primeira *versus* terceira pessoa é falsa dado que a linha que separa as duas abordagens não é entre “rigor” e “falta de rigor” mas, antes, se há “uma base metodológica clara que leva a uma validação comunitária e a um conhecimento compartilhado” (Varela, 1999: 340), que é exatamente o que a abordagem da Neurofenomenologia procura defender.⁴³

Mas não estaremos perto de uma abordagem também ela reducionista no sentido exposto no capítulo anterior? Não estaremos a instrumentalizar o conceito de experiência da primeira pessoa de forma a que possa ser investigado por instrumentos científicos particulares? Essa é a objeção que analisaremos de seguida.

IV.2.3. A Objeção da “Acusação Reducionista”

Uma objeção que pode ser levantada ao programa da Neurofenomenologia é a acusação de que o mesmo parece demasiado próximo do reducionismo que apresentámos anteriormente no seguinte aspeto: esta abordagem parece aceitar que as questões da mente consciente podem ser resolvidas ou, pelo menos, auxiliadas, pela ciência empírica, encontrando-se numa posição anti-isolacionista e, portanto, naturalista.

Esta acusação é interessante dado que surge antes do próprio desenvolvimento da abordagem neurofenomenológica: Husserl já tinha sido acusado de ser um “positivista” no sentido de atribuir demasiado peso à experiência. E de certa forma, Husserl aceitava esta acusação: tanto os positivistas lógicos, como os empiricistas britânicos e, finalmente, os fenomenólogos, aceitavam que o que deve ser analisado é aquilo que é “dado” ou oferecido pela experiência (cf. Overgaard, Gilbert & Burwood, 2013: 77). O próprio Husserl afirmou que, caso o positivismo fosse algo considerado como o pressuposto

⁴³Para alguns autores, a diferença entre o conceito clássico de introspeção e a sua redefinição não é de todo clara: “Não encontro aqui [*sc.* nas explicações das diferenças entre um e outro] nada que revele como a redução fenomenológica difere da introspeção como tal” (Bayne, 2004: 352).

central de todas as ciências, então os fenomenólogos seriam os “verdadeiros positivistas” (cf. Husserl, 1983: 39).

É claro que o conceito de “experiência” na fenomenologia difere substancialmente da abordagem positivista (cf. IV.1.1.). Mas será a proposta de Varela uma instanciação de uma teoria reducionista que procura conectar a experiência alcançada pela análise fenomenológica e os factos neurofisiológicos do sujeito? Para o filósofo chileno, a resposta é negativa: mesmo que se alcance uma correlação sólida entre as duas perspectivas, esses correlatos não são avançados *a priori* como uma defesa de um pressuposto fisicalista, mas, antes, como uma análise que toma em conta as diferenças e semelhanças de cada sujeito e da sua experiência quando analisada empiricamente (cf. Varela, 1996: 344).

Para Varela, esta posição surge não de um compromisso filosófico que se alinha numa conceção reducionista (cf. Capítulo 3) ou isolacionista (cf. Capítulo 2) da relação entre a Filosofia e a Neurociência, mas, diferentemente, de um compromisso pragmático. Mas não será a Neurofenomenologia uma posição anticientífica se for considerado o foco dado por esta abordagem à análise introspectiva⁴⁴ do próprio sujeito? Vejamos as palavras do fundador desta abordagem:

Mas [sc. a Neurofenomenologia] não é uma traição à ciência: [sc. antes] é uma extensão e complemento necessários. ciência e experiência restringem e modificam-se como uma dança. É aqui que reside o potencial de transformação. É também a chave para as dificuldades que essa posição [sc. a subjetividade da experiência] encontrou na comunidade científica. Exige que deixemos para trás uma certa imagem de como a ciência é feita e questionemos um estilo de formação em ciência que faz parte do próprio tecido da nossa identidade cultural (Varela, 1996: 347).

No mesmo sentido, Gallagher aponta para a ideia de que deve ser a ciência que se deve abrir à introdução da perspectiva fenomenológica e não deve ser dogmática ou recusar a sua utilidade epistémica:

Os investigadores não devem aceitar a eliminação dos processos da primeira pessoa e experiências únicas devido à rigidez dos métodos científicos; antes, a ciência pode expandir os seus métodos para incluir práticas fenomenológicas que capturam, investigam e explicam

⁴⁴ A contradição é assumida pela tese de que a introspeção, definida tradicionalmente como processo meramente subjetivo está em contradição direta com a tese objetiva do método científico de que o conhecimento, tanto do mundo como do próprio sujeito, deve ser imparcial do sujeito e de qualquer forma introspectiva à partida.

o que pode ser irreduzível a processos estritamente neurais (Bockelman, Reinerman-Jones & Gallagher, 2013: 3).

Ora, fica claro que estamos longe, metafilosoficamente, da abordagem reducionista analisada no capítulo anterior. Esta abordagem procura ainda defender a validade da experiência subjetiva sendo que esta não se reduz meramente à atividade neuronal de uma zona particular do cérebro como defende a perspectiva reducionista.

A Neurofenomenologia poderá ser vista (como veremos de seguida) como um reducionismo numa metafilosofia isolacionista. Contudo, dentro de uma metafilosofia que aceita o papel da ciência na descoberta ou explicitação da consciência fenomenal, esta abordagem deve ser caracterizada como sendo não reducionista: a solução do problema difícil da consciência é oferecer uma resposta que relacione a experiência e o físico sem que, contudo, uma se reduza a outro e vice-versa.⁴⁵

Mas esta objecção levanta um pressuposto interessante: mas, afinal, qual é a posição metafilosófica da abordagem neurofenomenológica? Analisemos essas considerações de seguida.

IV.2.4. A Objecção da “Metafilosofia Inconclusiva”

De forma interessante, mas metafilosoficamente dissemelhante, os defensores da Abordagem da Neurofenomenologia aceitam, num certo sentido, a falácia mereológica (FM) da Neurociência que foi avançada e analisada no segundo capítulo pela abordagem isolacionista: a consciência ou a mente não acontecem nem estão presentes ou localizados particularmente no cérebro. Alva Noë, um dos proponentes da abordagem da Neurofenomenologia, explicita da seguinte forma esta tese:

Para compreender a consciência – o facto de pensarmos e sentirmos e que um mundo surge para nós – precisamos de olhar para um sistema maior em que o cérebro é apenas um elemento desse sistema. A consciência não é algo que o cérebro possa alcançar [*sc.* ou produzir] por si só. A consciência requer a operação conjunta do cérebro, corpo e mundo. De facto, a consciência é uma conquista de todo o animal no seu contexto ambiental. Nego, enfim, que [*sc.* o leitor] seja o seu cérebro [...]. Ter uma mente [...] requer mais que um

⁴⁵Veremos de seguida de que forma esta tese poderá, na sua essência, ser contraditória.

cérebro. Cérebros não têm mente; [somente] pessoas (e outros animais) têm mentes (Noë, 2009: 10).

Ora, a semelhança entre a falácia mereológica da Neurociência avançada por Peter Hacker e a citação de Alva Noë é muito grande, chegando mesmo a aludir, igualmente, a Wittgenstein. Poderíamos, então, esperar que a mesma metafísica fosse avançada pelos neurofenomenólogos.

Todavia, e de forma interessante, a realidade é precisamente a oposta. A Abordagem da Neurofenomenologia não aceita as definições metafísicas da abordagem isolacionista em relação à Filosofia e à Ciência. Pelo contrário, tenta conciliar ambas de forma equilibrada como uma “dança”, na metáfora avançada por Varela.

Noë, por exemplo, afirma que “os objetivos da filosofia e da ciência não são diferentes: [ambas] procuram ter compreensão dos problemas que nos interessam” (Noë, 2009: 6). Ademais, Noë enquadra-se perfeitamente na visão naturalista de Quine, afirmando que “a ciência e o estilo humanístico de pensar devem alicerçar-se um no outro” (Noë, 2009: xv).

Além disso, Noë chega mesmo a criticar a concepção isolacionista de forma explícita ao afirmar que “a ideia de que a ciência e a filosofia, ou as humanidades em geral, são esferas separadas com os seus próprios critérios é, ela mesma, uma ideologia questionável, um resquício entusiasmado do início da Idade Moderna” (Noë, 2009: xv).

Finalmente, Noë afirma claramente a sua posição anti-isolacionista, afirmando que “a filosofia e a ciência partilham um objetivo comum: compreensão. A ciência e a filosofia devem trabalhar juntas para avançar essa compreensão” (Noë, 2009: xv).

Ora, mas será esta metafísica naturalista exclusiva de um proponente da Abordagem da Neurofenomenologia? Francisco Varela, o fundador desta abordagem, toma uma posição semelhante a Noë ao afirmar que “As questões empíricas devem ser guiadas pelas evidências [*sc.* da perspectiva] da primeira pessoa” (Varela, 1996: 344).

Ademais, afirma também que “a novidade da minha proposta é que as abordagens disciplinadas da primeira pessoa devem ser um elemento integrante da validação de uma proposta neurobiológica, e não apenas informações coincidentes ou heurísticas” (Varela, 1996: 344).

Fica, portanto, claro que uma objecção isolacionista à Neurofenomenologia perde qualquer força dado que esta abordagem tem, como pressuposto, uma interação nítida entre a investigação empírica e

conceptual. Se a abordagem reducionista⁴⁶ procura reduzir a segunda à primeira, a Abordagem da Neurofenomenologia, como perspectiva anti-reducionista, procura aceitar ambas como válidas, necessárias e fundamentais para se resolver os problemas difíceis da mente, como o problema da consciência.

Porém, tal poderá ser um erro porque, como daremos conta, embora possa parecer inicialmente que esta abordagem possui uma metafilosofia clara e explícita, quando analisada em detalhe, ficaremos com a possibilidade de que esta é inconclusiva.

Vejamos, assim, mais sistematicamente, por que razão a proposta metodológica da Neurofenomenologia pode ser interpretada como confusa. Os neurofenomenólogos avançam com o seguinte princípio metodológico: uma “metodologia de restrições recíprocas” (MRR) (cf. Varela, 1996: 343; Van Gelder, 1999: 246). Mas de que se trata especificamente esta metodologia particular? Alguns autores desta abordagem afirmam que é equivalente ao conceito de “Equilíbrio Refletido” (ER)⁴⁷ usado em Filosofia Moral.

Por exemplo, Thompson, Noë e Pessoa afirmam que estas duas noções, (MRR) e (ER) são “intermutáveis” entre si (Thompson, Noë & Pessoa, 1999: 195). Contudo, Varela, ao contrário destes autores, distingue claramente que (MRR) é diferente de (ER), dado que o primeiro é muito mais “preciso e exigente” que o segundo (cf. Varela, 1996: 344).

Há dois problemas evidentes nesta discordância: (i) se (MRC) e (ER) são realmente diferentes, como aponta Varela, ficamos sem perceber porquê, dado que o filósofo chileno não apresenta um argumento explícito nem sequer uma ideia implícita nas suas reflexões. Apenas afirma que há uma diferença, não explicitando qual é essa diferença, usando concretamente dois termos demasiado vagos e ambiguos como são “preciso” e “exigente”. Sem tal explicitação, o princípio metodológico perde força e utilidade.

Um segundo problema surge se aceitarmos, como fazem Thompson ou Noë, que (MRR) e (ER) são semelhantes. Como aponta Bayne (2004), (ER) é um princípio epistémico, e não um princípio explanatório ou explicativo: em (ER), conetar num equilíbrio refletido dois elementos de um conjunto de

⁴⁶ Alva Noë critica a abordagem reducionista da seguinte forma: “A neurociência cartesiana não tem qualquer evidência para suportar o seu pressuposto básico de que a experiência consciente é esgotada na sua totalidade pelos fenómenos neuronais” (Noë, 1999: 263).

⁴⁷ O equilíbrio refletido (ER) é uma metodologia avançada por John Rawls em Filosofia Política e Moral que sustenta um cruzamento entre as intuições morais particulares, por um lado, e os princípios gerais que deveriam justificar essas intuições (cf. Rawls, 2001: 29-32). No caso aplicado à relação entre a Filosofia e a Neurociência, a versão do (ER) pretende cruzar os princípios filosóficos gerais sobre a mente e o cérebro com os dados e factos particulares sobre a mente e o cérebro.

dados diferentes, sejam os dados neurofisiológicos e os dados da primeira pessoa, não é afirmar nenhuma relação explanatória – e muito menos ontológica – entre esses elementos (cf. Bayne, 2004: 356).

Seria necessário, para tal, avançar com “princípios-ponte”, como observámos no capítulo anterior. Mas como também já fizemos notar, a posição antireducionista da Abordagem da Neurofenomenologia não aceita o programa reducionista e, por consequência, a existência de tais leis de conexão, entrando assim num dilema difícil de ser resolvido.

Mas as contradições continuam entre os neurofenomenólogos. Como vimos, a Neurofenomenologia não deve ser encarada somente como uma “heurística” da Neurociência segundo Varela, dado que oferece muito mais que tal mero complemento. Outros proponentes desta abordagem, como Antonio Lutz, afirmam, em sentido oposto, que a vantagem desta proposta é exatamente a sua capacidade heurística de guiar a investigação empírica para alcançar novos resultados (cf. Lutz, 2002).

É interessante notar por que motivo o filósofo chileno não quer reduzir a sua abordagem a uma proposta meramente heurística: porque, no fundo, tal seria ficar somente nos correlatos neuronais da consciência. É certo de que os dados da primeira pessoa, ao guiar a investigação empírica, podem trazer novidades ou perspectivas inovadoras. Mas urge fazer notar que esses resultados inovadores serão sempre resultados empíricos: poderá descobrir-se que zonas do cérebro participam em processos mentais que eram tidos até então como localizados em zonas específicas do cérebro, por exemplo.

Ora, é exactamente por isto que Varela não quer que se pense a Neurofenomenologia como uma ferramenta heurística: neste sentido, ela apenas trará inovação ou avanço no conhecimento dos correlatos neuronais da consciência. Mas é exactamente isso que neurofenomenólogos como Lutz parecem querer alcançar com esta abordagem. O que levanta uma segunda contradição entre os defensores desta abordagem que parece indicar algo mais relevante: estamos perante uma proposta demasiado ambígua ou inconclusiva.

Uma forma de salvar o princípio (MRR) poderá ser uma defesa de que este funciona como princípio causal, seguindo o artigo de Varela e Thompson (2001). Neste artigo, os filósofos defendem que há dois tipos de causalidade no cérebro: (i) causalidade “ascendente”, ou local-global; (ii) causalidade “descendente”, ou global-local. Estes dois tipos de causalidade explicam as “relações causal-explicativas recíprocas entre eventos neurais e eventos conscientes quando estes últimos são concebidos como parâmetros de ordem da dinâmica cerebral em larga escala” (Thompson & Varela, 2001: 421).

Ora, chegamos a mais uma contradição que requer vários passos para ser compreendida no seu todo (cf. Bayne, 2004: 358-359):

- (a) Seguindo os autores, (i) e (ii) formam uma relação de causalidade entre o neuronal e o fenomenal⁴⁸ – tal é uma premissa assumida necessariamente *a priori*;
- (b) de (a), os autores parecem defender uma identificação entre o neuronal e o fenomenal;
- (c) Mas os autores são antireducionistas – não aceitam identificações reducionistas;
- (d) Ademais, as relações de causalidade / identidade entre o neuronal e o fenomenal coloca em causa a posição enativista (cf. IV.1.3.) dos autores – o fenomenal não está causalmente relacionado nem com o neuronal, nem com nenhuma parte do corpo da pessoa, mas com o animal inteiro;
- (e) Mas (b) contraria (c) e (d);

Assim, percebemos que (a) é falso e, mais importante, compreendemos que a proposta da Neurofenomenologia se encontra fundamentada em contradições (cf. (e)) que dificilmente podem ser resolvidas sem que uma das partes da abordagem (e.g. (b), (c) ou (d)) fique para trás. No fundo, estamos perante uma proposta que, à superfície, parece ter mérito e uma intuição metodológica muito interessante e relevante para as relações entre a Filosofia e a Neurociência.

Todavia, quando analisada em detalhe, percebemos que há uma incoerência metafilosófica difícil de ser ultrapassada. Além disso, as próprias posições particulares de alguns neurofenomenólogos parecem incoerentes ou, pelo menos, filosoficamente ambíguas.

Por exemplo: não se compreende que teoria metafísica da consciência surge do programa neurofenomenológico. Thompson defende que a consciência é um “fenómeno natural” cuja “complexidade cognitiva aumenta em função da crescente complexidade dos seres vivos” (Thompson, 2015: 103), mas é sabido que o filósofo canadiano não defende uma visão emergentista, nem pampsiquista, nem dualista ou que considere como um mistério. Apenas temos uma indicação de que “o físico não significa mais não-mental [...] o ser físico e o ser experiencial implicam-se ou derivam de algo que é neutro entre eles” (Thompson, 2015: 105), algo que é também afirmado por outros pampsiquistas como Galen Strawson ou Philip Goff.

Ora, esta afirmação parece declaradamente pampsiquista, ou uma defesa de um monismo neutro. No entanto, não encontramos nos escritos de Thompson uma defesa explícita de nenhuma

⁴⁸Conceito referente à experiência consciente subjetiva, baseada no conceito de “fenómeno” analisado na primeira parte do capítulo (cf. IV.1.1.)

dessas teorias metafísicas sobre a consciência. Ademais, não é de todo claro de que forma a introdução do princípio da causalidade recíproca que encontramos em (MRR) não seja uma forma estranha e ambígua de dualismo como é, aliás, afirmado por um dos defensores principais da Abordagem da Neurofenomenologia: “A introdução da causalidade recíproca pode ser entendida como uma expressão subtil de dualismo” (Lutz, 2002: 150).

Finalmente, e para concluirmos este capítulo, fazemos notar que a tentativa de solucionar a lacuna explicativa (cf. II.2.1) pode também ser colocada em causa. Kirchhoff e Hutto (2016) argumentam que a proposta neurofenomenológica baseada no conceito de “causalidade recíproca” falha por um motivo principal: o facto da grande descoberta desta abordagem – o facto do processo neurofisiológico ser sincrónico, heurístico, global – não demonstra que a consciência está relacionada causalmente com essas mesmas propriedades.

Antes, para estes autores, a Abordagem da Neurofenomenologia não se consegue ver livre da lacuna explicativa ao permitir ou defender ainda a distinção metafísica – cartesiana – entre o mental e o físico, e ao defender que estas duas partes têm uma existência irreduzível e que devem ser consideradas de igual forma na investigação sobre a mente consciente.

A única forma de combater esta lacuna é advogar a tese que defende que os níveis neuronais e fenomenais de descrição são, no fundo, níveis diferentes de descrever o mesmo fenómeno. Os autores acabam por propor o que chamam de “Cognição Enativa Radical” (cf. Hutto & Myin, 2012): mantendo a convicção de que o corpo e a sua interação com o mundo são fundamentais para compreender a mente humana consciente, rejeita-se as categorias metafísicas do fenomenal e do físico, dado que ambas são descrições – diferentes – do mesmo fenómeno e avança-se para a defesa de uma tese estrita de identidade que, ao contrário da proposta analisada no capítulo 3, incorpora no seu fundamento a ideia de que essa identidade é baseada na atividade dinâmica e corpórea da cognição.

É importante notar que esta posição está em oposição ao reducionismo fisicalista dado que não se assume que *todos* os fenómenos mentais podem ser descritos simplesmente pela linguagem ou vocabulário da física: “O fenomenal pode ser apenas o físico descrito diferentemente – sob uma aparência ou modo de apresentação diferente” (Kirchhoff & Hutto, 2016: 353). Desta forma, a ideia dos autores é evitar que o problema da consciência se coloque em primeiro lugar não precisando de lhe oferecer uma resposta – sendo esse o erro principal da Abordagem da Neurofenomenologia – mas, antes, e de forma bastante *wittgensteiniana*, defendendo que, com esta estratégia, se dissolve a existência do problema à partida.

Desta forma, compreendemos as várias razões para sermos céticos em relação à aceitação de uma metodologia baseada nos princípios da abordagem neurofenomenológica.⁴⁹ Embora haja um progresso metafilosófico relevante entre a abordagem isolacionista e a abordagem reducionista, é necessário detalhar com maior precisão de que forma é que efetivamente a relação metodológica entre a Filosofia e a Neurociência pode ocorrer de forma cogente, algo que apresentaremos no seguinte capítulo.

⁴⁹ Há uma objecção que não é metodológica na sua essência, mas levanta uma questão pertinente. Baseado nos fundamentos da teoria computacional (e.g. Minsky, 1998), pode ser argumentado que alargar o foco de investigação do cérebro para o corpo é não compreender o problema à partida: um computador pode controlar outros subsistemas. Por analogia, um cérebro pode controlar outras partes do corpo – tal não implica qualquer avanço para uma possível solução. Claro, esta objecção só terá relevância se se assumir que o cérebro funciona como um computador, um pressuposto filosófico que não é assumido pelos neurofenomenólogos em geral e que pode ser colocado em causa (cf. Gouveia & Neiva, 2017: 180-207).

IV.3. Conclusão:

Neste quarto capítulo, analisámos ao detalhe a Abordagem da Neurofenomenologia. Começámos por explorar os fundamentos teóricos e a suas bases filosóficas, tendo dado conta da importância inicial da fenomenologia de Husserl e dos seus postulados teóricos. De seguida, analisámos a importante crítica de Merleau-Ponty a Husserl defendendo que o corpo, tido como elemento da subjetividade, deve ter maior relevância conceptual e empírica na compreensão da mente consciente. Avançámos analisando a influência da teoria dos 4E's na primeira proposta relevante da Neurofenomenologia avançada por Francisco Varela: examinámos ao detalhe os seus pressupostos metodológicos, assim como os pressupostos metafilosóficos que esta abordagem exige.

Para demonstrar a aplicação empírica desta abordagem, apresentámos o caso da rivalidade binocular para demonstrar de que forma a Abordagem da Neurofenomenologia tem impacto na própria forma como a investigação empírica em Neurociência pode ser atualizada incorporando componentes da perspetiva da primeira pessoa que parece recusar à partida. No mesmo sentido, apresentámos as reflexões neurofenomenológicas sobre o Eu (*self*) de Evan Thompson.

Finalmente, avançámos com a análise das variadas objeções que podem ser levantadas a esta abordagem, nomeadamente a sua tentativa de “fechar” a lacuna explicativa famosamente argumentada em Filosofia da Mente; o uso problemático do conceito de introspeção; a objeção de que esta abordagem se aproxima demasiado dos projetos reducionistas apresentados no terceiro capítulo; e, por fim, a objeção de que estamos perante uma metafilosofia no mínimo inconclusiva e no máximo incoerente.⁵⁰

Embora esta abordagem seja um avanço interessante às abordagens anteriores, será necessário aprimorar com maior particularidade metodológica as relações entre a Filosofia e a Neurociência algo que, como iremos analisar de seguida, é proposto pela Abordagem da Neurofilosofia Não-Redutiva.

⁵⁰ Este capítulo adapta e expande partes mínimas do capítulo “Uma Solução para o Problema da Mente Consciente?” (2018) da obra *Reflexões Filosóficas: Arte, Mente e Justiça* pela Editora Húmus.

V. NEUROFILOSOFIA NÃO-REDUTIVA

Neste quinto capítulo, procuraremos analisar a última abordagem da relação entre a Filosofia e a Neurociência nesta investigação. Depois de termos efetuado uma análise à (i) abordagem que nega essa relação totalmente (abordagem isolacionista), (ii) a abordagem reducionista que defende a redução da primeira à segunda e, finalmente, (iii) a abordagem neurofenomenológica que defende um cruzamento particular entre a fenomenologia e a Neurociência, iremos, por último, avançar com a análise de uma abordagem que tentará conciliar com mais precisão e detalhe metafilosófico as duas disciplinas numa tentativa de fundar uma metodologia que possa realmente potenciar algum progresso efetivo nas investigações da mente consciente.

Veremos de que forma esta abordagem a que chamaremos de “Neurofilosofia Não-Redutiva” se enquadra e se opõe em relação às abordagens já apresentadas em (i), (ii) e (iii). No fim, iremos compreender de que forma esta abordagem final faz uso de várias teses e intuições das três abordagens anteriores, podendo ser, por isso, uma abordagem inovadora.

Examinaremos, para tal, o maior defensor desta abordagem: o neurocientista, filósofo e psiquiatra Georg Northoff, focando a nossa atenção no seu *Minding the Brain* (2014). Faremos notar de que forma esta abordagem metodológica difere da ideia reducionista de pensar a mente como restringida ao cérebro, sendo uma abordagem caracterizada como “*brain-reduced*”, procura encarar o pressuposto oposto: aceitar o rigor neurocientífico do estudo do cérebro, consentindo uma abordagem “*brain-based*”, mas não de um modo unilateral – daí falar-se numa Neurofilosofia Não-Redutiva, em oposição à Neurofilosofia Redutiva apresentada no terceiro capítulo.

Finalmente, veremos de que forma esta perspetiva faz efetivamente uso de ferramentas e técnicas das duas disciplinas sem colocar em causa uma em detrimento da outra, sendo, no fim, mais uma resposta positiva ao problema metodológico aqui analisado.

V.1. Pressupostos da Abordagem Não-Redutiva

V.1.1. Equilíbrio Metodológico

A primeira consideração avançada por um defensor da Abordagem da Neurofilosofia Não-Redutiva é a ideia de que devemos separar duas ideias fundamentais acerca do reducionismo: (i) o seu sentido metafísico e (ii) o seu sentido metodológico.

Em (i), esta perspetiva pode ser tida como reducionista caso o padrão de comparação seja uma filosofia dualista como o cartesianismo: dado que se assume que a mente humana é um fenómeno natural que não está fora da interação entre o cérebro, o corpo e o ambiente, esta abordagem pode ser vista como reducionista no sentido de negar um dualismo ontológico fundamental entre o mental e o neuronal/corpóreo.

Contudo, se for visto como (ii), esta posição pode ser considerada como antireducionista: dado que nega a tese da abordagem reducionista que procura reduzir metodologicamente a Filosofia à Neurociência, a posição de Northoff, neste sentido metodológico, é vista como totalmente oposta. Nesta posição, não procuramos reduzir nem o método da Filosofia à Neurociência, nem o oposto: estaremos perante um equilíbrio metodológico neutro que aceita que ambas as disciplinas são fundamentais para contribuir para a exploração rigorosa da mente consciente.

Dessa primeira premissa, surge uma segunda: a separação metodológica entre o conceito de “disciplina” e de “domínio”. O primeiro envolve, usualmente no seu uso, um cruzamento entre o assunto ou tema da disciplina e os métodos utilizados para investigar esse assunto: por exemplo, a disciplina da Neurociência investiga de que forma o cérebro percebe o mundo externo através de métodos empíricos particulares.

Contudo, o segundo conceito, de domínio, poderá ser mais útil neste contexto: este conceito limita-se ao conteúdo, problema ou assunto que se quer investigar sem, todavia, especificar os procedimentos metodológicos que devemos utilizar na investigação:

[...] o conceito de domínio diz respeito ao campo ou área de investigação e, portanto, problemas ou questões particulares, independentes e distintas do tipo de estratégia metodológica por meio da qual essas questões são investigadas [...] domínio diz respeito ao assunto, enquanto o conceito de disciplina refere-se mais ao método aplicado para investigar esse determinado assunto (Northoff, 2014a: 13).

Esta estratégia metodológica procura alcançar uma forma neutra de cruzar os problemas e conceitos da Filosofia (e.g. a mente) com os problemas e conceitos da Neurociência (e.g. o cérebro) de modo que haja, como veremos, um cruzamento constante entre as duas disciplinas.

Deste modo, através da exclusão de uma metodologia preferencial assumida à partida, esta estratégia procura uma independência neutral que poderá fazer sentido: deixamos de considerar, *a priori*, um isolamento metodológico entre as duas disciplinas. Pelo contrário, passemos a estar perante uma posição de continuidade entre a Filosofia e a Ciência. Assim, poderemos considerar esta posição como uma abordagem naturalista no sentido explorado anteriormente.

Esta nova abordagem acabará por fazer uma defesa de um “naturalismo cooperativista” (cf. Goldman, 1999: 24-50): a relação entre a Filosofia e a Neurociência deve ser vista como uma cooperação intermetodológica entre as duas disciplinas (Huaping & Xiaoming, 2007). Esta posição encontra-se no sentido oposto de um “naturalismo de substituição” que detalhámos no capítulo terceiro: a ideia de que os métodos da Filosofia devem ser tendencialmente substituídos pelos métodos da Neurociência (cf. Northoff, 2014a: 64-66).

Vejamos como é que a Abordagem da Neurofilosofia Não-Redutiva difere da versão redutiva.

V.1.2. Estratégias Metodológicas da Neurofilosofia

De que forma a Neurofilosofia Redutiva (NR) funciona em oposição à Neurofilosofia Não-Redutiva? (NNR) Como vimos, a metodologia unilateral da primeira implica uma forma de monismo metodológico cuja base está no método observacional e experimental da ciência. Deste modo, (NR) é vista como uma abstração teórica dos dados empíricos descobertos pelo método observacional-experimental: tais abstrações são descritas como conceitos que serão usados na investigação empírica.

Como fizemos notar, esta abordagem aceita a rejeição avançada por Quine da dicotomia sintético/analítico, transformando os conceitos ditos filosóficos em meras abstrações dos dados científicos. Tal implica a substituição dos conceitos trabalhados pelos filósofos por outros alegadamente científicos. Assim, os conceitos passam a fazer parte do domínio empírico.

Como consegue um defensor da (NR) ligar conceitos a factos de forma sistemática? Patricia Churchland (cf. Churchland, 1986) sugere uma estratégia pragmática chamada “bootstrapping”. Tradicionalmente, a investigação filosófica começa com uma definição de um conceito particular, sendo

que estas definições são pontos de partida para discutir as consequências dessas definições – estamos perante uma metodologia do tipo isolacionista nesta forma particular de fazer filosofia (cf. Capítulo 2; Northoff, 2014a: 113).

Churchland altera o ponto de partida no seu método: em vez de se iniciar uma investigação com definições claras e *a priori* de conceitos de um fenómeno particular, a filósofa sugere que comecemos com casos empíricos típicos e representativos em que todos os investigadores acordam representarem ou refletirem o fenómeno em questão.

Por exemplo, podemos começar por estudar pacientes com graves lesões cerebrais. Estudamos estes doentes de forma empírica: tal aumenta o nosso conhecimento dos mecanismos neuronais subjacentes às suas lesões. Tal poderá também trazer algum desenvolvimento sobre como devemos definir certo conceito (como a consciência).

Assim, os conceitos não são pressupostos na investigação à partida: a definição do conceito é inferida dos dados empíricos – o conceito já não é visto como um “input”, mas, antes, como um “output” (cf. Northoff, 2014a: 114). Vejamos um exemplo oferecido pela filósofa sobre o fenómeno do cérebro dividido (*split-brain*):

Um princípio-padrão, ilustrado pelos resultados do cérebro dividido, é que a definição do fenómeno a ser explicado se desenvolve com descobertas experimentais. Nos estágios iniciais da investigação científica a qualquer problema, a definição precisa do fenómeno é dificultada precisamente porque não se sabe o suficiente para permitir uma definição precisa. Uma estratégia pragmática é começar a estudar os casos que se concorda serem exemplos óbvios do fenómeno. Alimentada por esse acordo, a caracterização aproximada e provisória pode alavancar os primeiros estágios da ciência, com aprimoramentos na definição do fenómeno a emergir à medida que os factos ao redor se tornam claros (Churchland, 2002b: 309).

Podemos, assim, colocar em figura a unilateralidade deste método da seguinte forma:

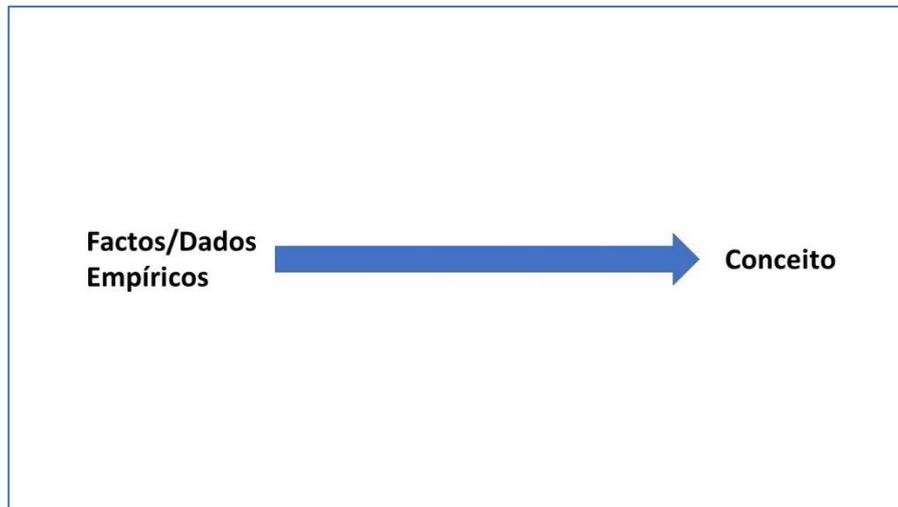


Figura 5.1.: Estratégia metodológica de Churchland – conceito como um “output” da investigação empírica.

É importante fazer notar a ideia defendida nesta estratégia metodológica: a plausibilidade empírica tem o maior peso da equação, sendo a plausibilidade conceptual menos relevante.

É difícil perceber em que sentido não existem conceitos que pressuponham os factos científicos descritos pelas investigações neurocientíficas – nem que seja implicitamente. Esta crítica aponta para a possível circularidade desta metodologia: os conceitos podem já estar pressupostos pelos dados, em vez de ser inferidos dos mesmos (cf. Northoff, 2014a: 114). Já levantámos esta crítica anteriormente quando analisámos a abordagem isolacionista no segundo capítulo (cf. II.1.2.): os conceitos antecedem a experiência (cf. Bennett & Hacker, 2003: 4), e não o contrário.

É neste ponto que entra John Searle, sugerindo que devemos adaptar os conceitos e as suas definições aos dados e factos empíricos, em vez de inferi-los. Uma segunda estratégia utilizada por filósofos como Searle é a estratégia da “adaptação unilateral dos conceitos aos factos”: nesta estratégia, compreende-se que a primeira estratégia falha, porque não se apercebe que os próprios factos e dados empíricos já pressupõem certos sentidos e significados dos conceitos que são usados na própria investigação empírica a efetuar. Assim, para superar esta circularidade, Searle sugere que adaptemos os conceitos e as suas definições aos dados e factos em vez de inferi-los (cf. Northoff, 2014a: 115).

Porém, será que a ideia de Searle faz sentido? Apliquemos a mesma ao conceito de “mente”: muitos filósofos defendem que este conceito se refere a algo não-físico. Se aceitarmos que o conceptual tem primazia sobre o empírico, e se definirmos *a priori* o conceito de mente como não-física, condenamos

à partida qualquer possibilidade de haver uma investigação científica do fenómeno em questão dado que a sua definição exclui essa possibilidade excluindo uma investigação da perspectiva da terceira pessoa: impossibilita-se, assim, uma neurociência da consciência (cf. Northoff, 2014a: 115).

Todavia, poderá ser defendido que tal estratégia contradiz recentes investigações que demonstram que mecanismos neuronais específicos poderão estar relacionados com a experiência subjetiva. O que fazer? Não podemos mudar os factos: podemos, contudo, adaptar e modificar os nossos conceitos para acomodarem esses factos (cf. Northoff, 2004a: 114-116).

A ideia de Searle é que temos de redefinir o conceito, por exemplo, de “físico” de forma a associar a perspectiva da primeira e terceira pessoa. Bastará, por exemplo, livrarmo-nos dos conceitos cartesianos de “físico” e “mental” e redefini-los de forma diferente (cf. Searle, 2004a: 82-83).

Adaptando desta forma definições de conceitos filosóficos às descobertas neurocientíficas, esses conceitos revistos podem abrir portas para conseguirmos ligar os conceitos filosóficos aos dados da Neurociência. Segundo Searle, fazendo este trabalho unilateral de adaptação dos conceitos aos factos conseguir-se-á que a Filosofia se torne uma propedêutica conceptual da Neurociência (cf. Searle, 1999: 2070).

Admitindo assim que definições adequadas de conceitos são necessárias para tornar acessível a própria investigação científica, os conceitos passam a ser o “input” – em vez do “output” acima descrito – para a observação e experimentação:

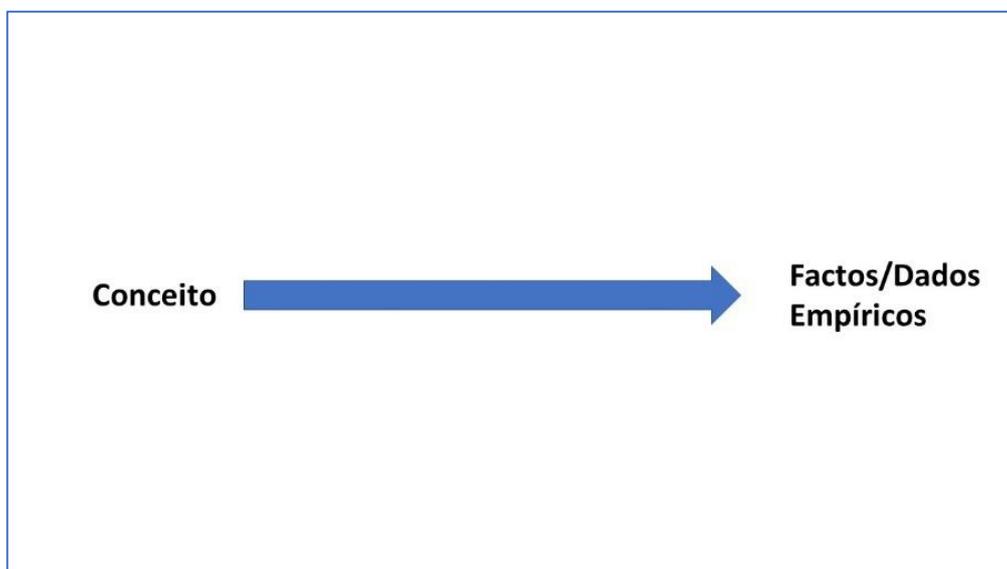


Figura 5.2.: Estratégia da adaptação unilateral dos conceitos aos factos – conceito como um “input” da investigação empírica.

Nesta estratégia, dá-se importância à plausibilidade conceptual dos conceitos, mas estes também têm de se adequar e adaptar aos factos empíricos, sendo a plausibilidade empírica dos mesmos também relevante.

De certa forma, a estratégia de Churchland parece ser oposta à estratégia de Searle: enquanto a primeira considera somente relevante a plausibilidade empírica e considera o conceito como um mero “output” do empírico, Searle faz notar que a plausibilidade conceptual do conceito também é relevante, considerando-o mais como um “input” do empírico que deve ser adaptado. Haverá alguma forma de compatibilizar estas duas estratégias que parecem opostas entre si?

A proposta da Abordagem da Neurofilosofia Não-Redutiva é responder positivamente: nesta perspectiva, o primeiro passo da investigação neurofilosófica é o “input conceptual”, onde um conceito filosófico original é tido como “input” – no sentido de Searle – para desenvolver critérios operacionais para investigação empírica subsequente.

O segundo passo é a investigação empírica em si, cujo dados e factos descobertos podem ser comparados e combinados com a definição original do conceito. Tal passo leva ao “output conceptual”, onde o conceito é redefinido considerando a sua plausibilidade empírica.

Finalmente, o conceito já redefinido pode ser reinserido no seu espaço filosófico original – o domínio epistemológico, ontológico ou ético – de forma a que se investigue as suas implicações para os respetivos problemas filosóficos. Este passo leva a um “ciclo de reentrada conceptual” – onde é comparado o “output conceptual” com o “input conceptual” – que origina verdadeiros conceitos neurofilosóficos em vez de conceitos apenas neurocientíficos ou filosóficos:

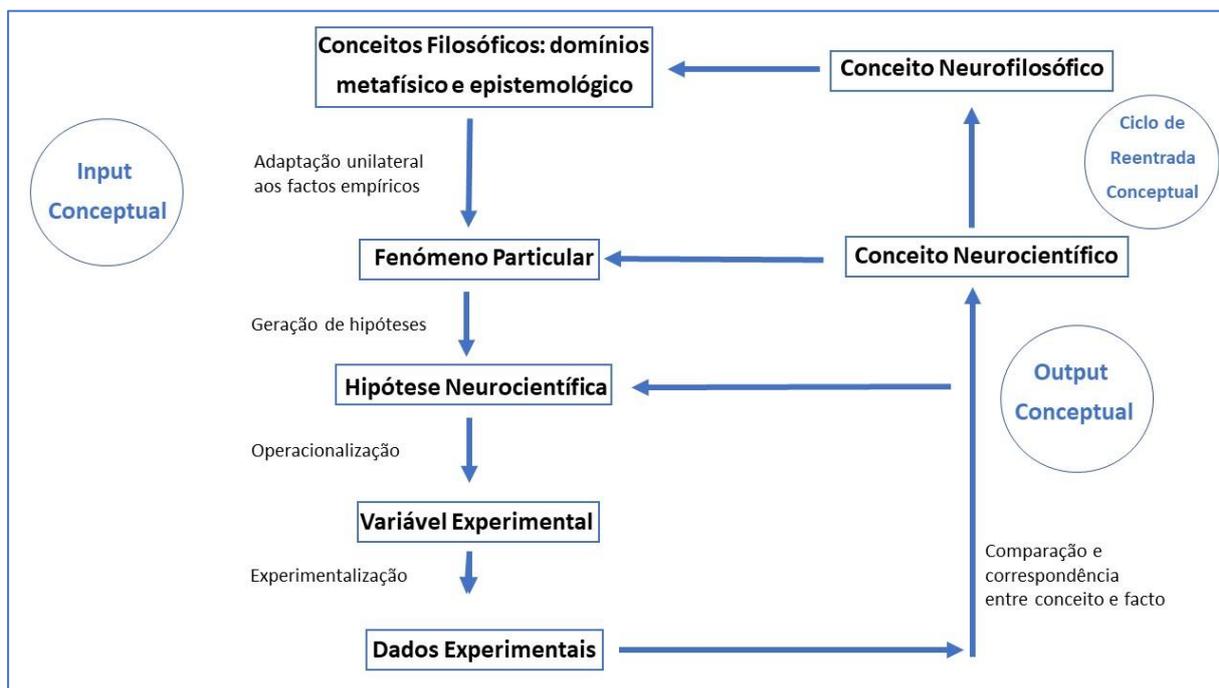


Figura 5.3.: Estratégia da Neurofilosofia Não-Redutiva (cf. Northoff, 2014a: 119).

Deste modo, esta nova estratégia irá requerer uma contínua troca iterativa entre os conceitos e os dados que se afetam continuamente, modificam-se e redefinem o conceito a ser investigado. Estamos, assim, perante uma estratégia que incorpora a estratégia de Churchland e a estratégia de Searle: ao contrário da primeira, os conceitos filosóficos são ainda considerados; ao contrário da segunda, essa relação não é unilateral, mas bilateral e constante.

Esta estratégia metodológica torna possível que a Neurofilosofia se torne uma disciplina distinta e separada tanto da Neurociência como da Filosofia, com um método específico, que é diferente do método conceptual-linguístico da Filosofia e diferente do método observacional-experimental da Neurociência sendo, antes, uma forma de fundir as duas metodologias numa nova metafilosofia (cf. Northoff, 2014a: 119-120).

Vejamos um exemplo simples de como é que esta metodologia pode ser aplicada na prática a conceitos da Filosofia, como por exemplo o conceito de “consciência”. Imagine-se que um neurocientista quer fazer um estudo imagiológico funcional da consciência e procura construir uma experiência correspondente. Para tal, deve possuir alguma definição de consciência, mesmo que seja apenas uma

definição operacional inicial. Este tipo de definição operacional oferece algum critério pelo qual o neurocientista pode, por momentos, distinguir entre a presença e a ausência de consciência.

Note-se que, sem esta definição operacional e o seu critério, o neurocientista seria incapaz de fazer sentido e interpretar os factos e os mecanismos neuronais específicos associados às características fenomenais particulares da consciência. Ao continuar a investigação, poderá perceber que a definição (inicial) não consegue incorporar todos os factos e poderá reconsiderar a definição (e o critério) inicial, tornando assim a investigação iterativa no sentido anteriormente indicado (cf. Northoff, 2014a: 117).

Esta sucessiva “dança” entre os conceitos – e a sua plausibilidade conceptual – e os factos empíricos – e a sua plausibilidade empírica – é a marca da Abordagem da Neurofilosofia Não-Redutiva. Vejamos que princípios metodológicos estão presentes nesta metodologia.

V.1.2.1. Princípio da Assimetria

A Abordagem da Neurofilosofia Não-Redutiva faz uso de um princípio metodológico relevante: o princípio da assimetria entre as condições lógicas e as condições do mundo natural. O conjunto das possibilidades lógicas é o plano de fundo onde o mundo natural se encontra limitado, sendo uma parte reduzida desse plano:

As condições lógicas referem-se a todos os mundos possíveis, isto é, logicamente concebíveis, incluído os mundos naturais e não naturais, com apenas o primeiro subjacente às nossas leis físicas e biológicas. As condições naturais, pelo contrário, referem-se apenas ao mundo natural e, portanto, às respetivas leis físicas e biológicas. Como as condições lógicas abrangem mundos naturais e não naturais, necessariamente incluem também as condições naturais (Northoff, 2004a: 95).

Ora, dado que a relação entre o mundo lógico e o mundo natural é assimétrica, qualquer tentativa de reduzir as condições lógicas às condições do mundo natural falhará, como analisado por McCauley (2001): qualquer tentativa de eliminar completamente as condições lógicas e, portanto, as teorias filosóficas, em favor das condições do mundo natural e, portanto, das hipóteses neurocientíficas, estará condenada ao insucesso (cf. McCauley, 2001: 439–440).

A principal razão é exatamente a relação assimétrica entre os dois planos: se fosse simétrica, a redução ou eliminação poderia ser filosoficamente sustentável. Mas dado que é, pelo contrário, assimétrica, o caso da eliminação mantém-se uma impossibilidade *a priori*. (Northoff, 2004a: 95). Mas que implicações surgem com esta assimetria? Vejamos as duas principais:

- (a) As inferências diretas das condições lógicas para as condições naturais são inexecutáveis: “Como as condições lógicas incluem uma gama mais ampla de condições do que as condições naturais, a inferência direta do primeiro para o segundo pode confundir mundos não naturais, isto é, logicamente concebíveis com o mundo natural, ou seja, real.” (Northoff, 2004a: 96);
- (b) Na mesma linha, inferências diretas entre as condições naturais e as condições lógicas são igualmente inexecutáveis: “Como as condições naturais incluem uma gama menor de condições do que as condições lógicas, a inferência direta do primeiro para o segundo pode equiparar falsamente os mundos não naturais ao mundo natural” (Northoff, 2004a: 96).

Ambas as inferências (a) e (b) representam “falácias condicionais” dado que se confunde as diferentes relações possíveis. Por exemplo, em relação a (b), este tipo de falácia é cometido por Churchland – e a Abordagem Reducionista em geral – em relação ao problema mente-cérebro: a tese de que este intrincado problema possa ser resolvido apenas apelando à ciência empírica é exatamente confundir as condições lógicas com as naturais:

[...] o problema mente-cérebro refere-se a condições lógicas, incluindo mundos naturais e não naturais. Em contraste, a neurociência refere-se a condições naturais, incluindo apenas o mundo natural. A aplicação direta e a inferência de descobertas empíricas na neurociência para o problema mente-cérebro podem, assim, falsamente equiparar mundos não naturais ao mundo natural. Tal, no entanto, pode levar a conclusões falsas, uma vez que mundos não naturais incluem uma gama mais ampla de condições do que o mundo natural. O neurocientista [*sc.* tradicional], portanto, comete uma “falácia condicional” quando aplica diretamente e deduz de suas descobertas empíricas ao problema mente-cérebro (Northoff, 2004a: 97).

Contudo, o filósofo tradicional também comete este erro na sua versão (a): a solução do problema mente-cérebro pode, segundo esta abordagem, necessitar somente de uma solução lógica que não se aplique necessariamente ao mundo natural investigado pela Neurociência. O problema é que “as condições lógicas podem não necessariamente ‘corresponder’ às condições naturais” (Northoff, 2004a: 97). Assim, soluções meramente filosóficas (i.e. apelando à lógica) podem levar a confusões entre o mundo natural e o mundo não natural.

Apesar desta assimetria, existe uma sobreposição entre o mundo natural e lógico, dado que o mundo lógico inclui, como vimos, tanto o mundo natural como o mundo não natural. É exactamente esta sobreposição entre os dois mundos que é necessário conseguir conjugar numa investigação neurofilosófica: a metodologia da Neurofilosofia Não-Redutiva encontra-se no limiar dessa sobreposição, ou seja, entre as teorias filosóficas e as hipóteses neurocientíficas (cf. Northoff, 2004a: 97).

Portanto, o que está em causa é recusar uma preferência unilateral pelo mundo natural (i.e. os factos *a posteriori*) ou pelo mundo lógico (i.e. os conceitos *a priori*). Pelo contrário, procura-se o cruzamento desses dois mundos com o objetivo de criar conceitos verdadeiramente neurofilosóficos de modo a não cometermos os erros indicados anteriormente.

V.1.2.2. Princípio de Bidirecionalidade

O segundo princípio avançado pela Abordagem da Neurofilosofia Não-Redutiva é o princípio da bidirecionalidade: a ideia de que é essencial cruzar as teorias da Filosofia com as hipóteses da Neurociência, isto é, as condições lógicas com as condições do mundo natural.

Esta justaposição entre as duas condições pode parecer favorável a um projeto reducionista do tipo apresentado no terceiro capítulo pela Abordagem do Reduccionismo: dado que há uma sobreposição entre o mundo lógico e o mundo natural, fará sentido pensar reduzir a parte lógica que se sobrepõe às condições naturais ao próprio mundo natural (cf. McCauley, 2001: 441).

Northoff concorda com McCauley na ideia de que há uma possível sobreposição dos dois mundos. Todavia, discorda que, dada essa sobreposição, se possa pensar um mapa de cruzamentos interteóricos completo e fechado: para o filósofo alemão, esse mapa terá de ser, pelo contrário, incompleto, sendo que apenas poderia ser completo caso houvesse uma relação unilateral entre o mundo lógico e o mundo natural: mas tal não é o caso, sendo exactamente essa a tese da bidirecionalidade (cf. Northoff, 2004a: 97-98).

O princípio da bidirecionalidade permite dois passos relevantes para a Abordagem da Neurofilosofia Não-Redutiva:

- (i) O trabalho filosófico pode ser cruzado com o trabalho neurocientífico permitindo um critério de “plausibilidade empírica” desse trabalho filosófico: caso essa plausibilidade empírica seja positiva (i.e. estiver de acordo com os dados e factos recolhidos pela evidência empírica), estaremos perante uma teoria promissora e sólida; caso essa plausibilidade empírica for negativa, duas consequências podem acontecer: (i.1) podemos assumir que o trabalho filosófico em questão não poderá ser cruzado com o trabalho neurocientífico (i.e. existe uma lacuna intransponível) e que, por causa disso, não poderemos desenvolver qualquer hipótese verdadeiramente neurofilosófica; ou (i.2) podemos optar por modificar o trabalho filosófico anterior para acomodar o trabalho neurocientífico através de uma reconceptualização dos conceitos filosóficos iniciais;
- (ii) O trabalho neurocientífico pode ser cruzado com o trabalho filosófico permitindo um critério de “plausibilidade conceptual” desse trabalho neurocientífico: o último deve ser logicamente consistente em relação às implicações e pressupostos epistémicos e ontológicos (cf. Northoff, 2004a: 98).

Deste modo, o princípio da bidirecionalidade permite uma dança iterativa constante entre os dois mundos, isto é, o lógico e o natural, o filosófico e o neurocientífico. É necessário detalhar, contudo, com mais precisão, de que forma se relacionam os dois mundos. O próximo princípio, da “circularidade transdisciplinar”, procura detalhar essa mesma relação, que passaremos a analisar de seguida.

V.1.2.3. Princípio de Circularidade Transdisciplinar

O princípio da “circularidade transdisciplinar” refere-se exactamente à dança iterativa referida acima entre o trabalho filosófico e o trabalho neurocientífico com o objetivo de desenvolver-se uma hipótese verdadeiramente neurofilosófica.

É importante fazer notar que, como vimos anteriormente, cruzamentos isomórficos ponto a ponto não podem ser alcançados; em alternativa, podem ser encontrados cruzamentos indiretos que são alcançados através dessa dança iterativa constante entre os dois mundos (cf. Figura 5.3). Northoff faz uso, curiosamente, neste ponto, do conceito avançado por Francisco Varela: a “circularidade disciplinada” (Varela, 1998: 342). Mas esta circularidade específica refere-se a que conjuntos, exatamente?

O próximo passo estratégico é conjugar o trabalho filosófico com “proposições teóricas” (PT) ligadas às condições lógicas e, por isso, independentes do mundo natural actual, relacionadas com o epistémico e o ontológico, e o trabalho neurocientífico conjugado com as “observações” (O) ligadas às condições naturais e, por isso, fazendo parte do mundo natural actual.

As observações particulares (O) podem ligar-se a proposições teóricas (PT) particulares e, ao mesmo tempo, excluir outras (PT): a possibilidade ou impossibilidade de se conseguir inferir (O) de (PT) é analisada pela “implicação empírica” do trabalho filosófico comparando-o com as descobertas empíricas da Neurociência para garantir uma “plausibilidade empírica” relevante.

O mesmo deve acontecer na “explicação teórica”: deve-se testar a “plausibilidade conceptual” dos pressupostos epistémicos e ontológicos assumidos necessariamente pelo trabalho neurocientífico. Assim, a comparação entre (PT) e (O) considerando a plausibilidade conceptual e empírica deve ser vista como uma condição necessária da possibilidade de conciliar o próprio trabalho filosófico com o trabalho neurocientífico (cf. Northoff, 2004a: 102-103).

De seguida, deve-se proceder à “homogeneização”. No caso de existir uma plausibilidade conceptual inconsistente, o trabalho neurocientífico deve proceder à modificação e correção das próprias proposições teóricas ou das teorias ontológicas ou epistémicas assumidas de modo a que haja um processo de “homogeneização” entre as duas disciplinas.

No mesmo sentido, no caso de existir uma plausibilidade empírica inconsistente, o trabalho filosófico deve proceder à modificação e correção das suas “implicações empíricas”. Com este passo, garantimos que há uma mútua homogeneização entre as duas disciplinas garantindo a plausibilidade conceptual e empírica de uma hipótese a investigar.

Este passo garante igualmente que os dois níveis de investigação, (PT) e (O), se vão modificando e orientando (cf. Figura 5.3.) um ao outro, conseguindo-se assim, outra condição necessária à possibilidade de conciliar o trabalho filosófico com o trabalho neurocientífico (cf. Northoff, 2004a: 103).

Por fim, deve-se investigar as consequências das teorias modificadas no estágio anterior:

- (a) A hipótese neurocientífica pode (i) permanecer independente das pressuposições ontológicas/epistêmicas modificadas ou (ii) ter que ser modificada para ser compatível com as pressuposições ontológicas/epistêmicas modificadas; caso (ii), podemos estar perante o desenvolvimento de uma hipótese verdadeiramente neurofilosófica;
- (b) A teoria filosófica pode (i) permanecer independente da hipótese empírica modificada ou (ii) ter que ser modificada para ser compatível com a hipótese empírica modificada; caso (ii), podemos estar perante o desenvolvimento de uma hipótese verdadeiramente neurofilosófica (cf. Northoff, 2004a: 102-103).

Assim, neste último estágio, pode-se comparar os resultados obtidos (a) e (b) levando-o a uma “comparação interdisciplinar” entre o trabalho filosófico e o trabalho neurocientífico: (PT) e (O) não são apenas modificadas ou corrigidas, mas, neste processo, vão influenciando-se mutuamente para novos (PT*) e (O*). Este último estágio do princípio torna-se assim também ele uma condição necessária à possibilidade de conciliar o trabalho filosófico com o trabalho neurocientífico (cf. Northoff, 2004a: 104).

Vejamos de que forma podem ser ilustrados os variados passos metodológicos com um exemplo do problema da personalidade investigado por Derek Parfit (1989). Suponha-se que Parfit avança com uma teoria filosófica sobre a personalidade que assume, implicitamente, pressupostos empíricos sobre o cérebro. Todavia, esses pressupostos empíricos estão errados quando comparados com a literatura neurocientífica atual. Assim, seguindo a estratégia apresentada, devemos considerar, modificar e corrigir os pressupostos empíricos da suposta teoria de Parfit, o que poderá levar à modificação da própria teoria filosófica. Especificamente, devemos prosseguir com os seguintes passos:

- (1) Compreender as “implicações empíricas” da teoria de Parfit; suponha-se que o filósofo sugere que há uma relação linear entre as células cerebrais e as funções cognitivas, pelo que implica uma relação um-por-um entre os dois níveis (i.e. células cerebrais e funções cognitivas);
- (2) Comparando (1) às investigações actuais em Neurociência, perceberemos que (1) deve ser revisto e modificado dado que estamos perante uma inconsistência e, portanto, está em falta a sua “plausibilidade empírica”;
- (3) Depois de atualizar (1), alcançamos uma homogeneização de (1) que permite modificar a implicação empírica anterior para pressupostos empíricos válidos: de que a relação entre as

células cerebrais e funções cognitivas não é uma relação de um-por-um mas, antes, de um-por-muitos ou muitos-por-um;

(4) Esta alteração pode, de seguida, ser confrontada com suposta teoria filosófica da personalidade de Parfit que precisará de ser atualizada;

(5) Finalmente, o processo de (1) a (4), isto é, a interação constante entre a hipótese empírica do funcionamento do cérebro e a teoria filosófica da personalidade levará a um aumento da plausibilidade tanto conceptual como empírica da teoria (cf. Northoff, 2001; Northoff, 2004a: 104;).

Compreendemos, com esta ilustração do problema da identidade pessoal, de que forma podemos aplicar o método da Neurofilosofia Não-Redutiva. Apesar de ser logicamente consistente, a suposta teoria da personalidade de Parfit tem problemas a nível de inconsistências empíricas óbvias. Como tal, uma hipótese verdadeiramente neurofilosófica da personalidade terá de corrigir essas inconsistências caso queira ser bem-sucedida. É a natureza desta hipótese neurofilosófica que analisaremos de seguida.

V.1.3. A Natureza e Metafilosofia da Hipótese Neurofilosófica

A Abordagem da Neurofilosofia Não-Redutiva avança com uma definição particular de hipótese neurofilosófica: seguindo determinados princípios metodológicos (cf. princípio da assimetria, princípio de bidirecionalidade e princípio de circularidade transdisciplinar) e relacionando iterativamente os dois níveis de investigação (cf. Figura 5.3.), alcançamos uma ligação muito precisa entre o trabalho filosófico e o trabalho neurocientífico (cf. Northoff, 2004a: 113) De que forma se pode refutar uma hipótese neurofilosófica neste sentido?

Existem três formas objetivas de o fazer, baseadas no que apresentámos:

- (a) Apontar uma inconsistência lógica particular que ponha em causa a validade interna da hipótese neurofilosófica;
- (b) Apontar uma inconsistência empírica particular (como foi feito no exemplo de Parfit) que ponha em causa a validade externa da hipótese neurofilosófica;

- (c) Apontar uma inconsistência transdisciplinar que ponha em causa a ligação entre as duas disciplinas, isto é, uma hipótese não está de acordo com os princípios metodológicos apresentados (cf. Northoff, 2004a: 113).

A natureza da hipótese neurofilosófica é assim diferente tanto do trabalho filosófico como do trabalho neurocientífico: o primeiro só pode ser falsificado apontando alguma inconsistência lógica, sendo que uma inconsistência empírica será totalmente irrelevante. Já o segundo só pode ser falsificado apontando alguma inconsistência empírica, sendo que uma inconsistência lógica será totalmente irrelevante, dado que o conceito de “hipótese” é alargado ao sentido usual que é usado para descrever uma hipótese empírica:

Devido à inclusão de pressupostos ontológicos/epistémicos, isto é, dos aspetos teóricos, o significado do termo ‘hipótese’ é ampliado no caso de ‘hipótese neurofilosófica’, em comparação com o empírico, ou seja, hipótese neurocientífica. Esse significado ampliado reflete-se na necessidade de considerar os dois tipos de falsificação, a ‘falsificação empírica e lógica’ (Northoff, 2004a: 114).

Deste modo, por razões semelhantes, devemos distinguir uma hipótese neurofilosófica do trabalho filosófico dado que o segundo ignora refutações empíricas.

Deve-se também considerar uma segunda diferença entre os dois empreendimentos: a hipótese neurofilosófica não pressupõe ou predefine os seus conceitos iniciais, mas, como vimos, mantém-nos sujeitos a alteração, modificação ou correção consoante a informação recebida pela investigação empírica (cf. Praetorius, 2000: 30; Northoff, 2004a: 115).

Mas qual é a natureza das asserções neurofilosóficas? Serão analíticas, sintéticas, um misto delas, ou nenhuma delas? Northoff argumenta a favor da posição que rejeita a existência de proposições “puramente” analíticas e “puramente” sintéticas: dado que assume a importância tanto da plausibilidade conceptual (analítica) e da plausibilidade empírica (sintética), rejeita-se a posição extrema dos dois tipos de asserções.

Consequentemente, a própria distinção metafilosófica entre o trabalho filosófico e o trabalho neurocientífico, isto é, entre a Filosofia e a Neurociência fica limitada na Abordagem da Neurofilosofia Não-Redutiva. É curioso que também Northoff cite, neste momento, Quine, quando afirma que

Carnap reconheceu que é capaz de preservar um padrão duplo para questões ontológicas e hipóteses científicas apenas assumindo uma distinção absoluta entre analítico e sintético; não será necessário afirmar novamente que essa é uma distinção à qual resisto (Quine, 1951: 43).

Assim, podemos ter uma ampla gama de hipóteses neurofilosóficas que, no seu âmago, incluem considerações sintéticas e analíticas, podendo variar no grau de atenção e de fundamento de cada componente: algumas hipóteses poderão ser tendencialmente mais analíticas e outras mais sintéticas. Todavia, em ambos os casos, a componente contrária será sempre considerada e influenciará, num segundo nível, a plausibilidade conceptual e empírica de determinada hipótese neurofilosófica.

Portanto, a natureza da hipótese neurofilosófica deve ser vista como um “misto” que inclui tanto o *a priori* como o *a posteriori*, sendo que a mesma se encontra num balanço muito particular entre os dois: considerando-o de igual forma, mas não reduzindo um ao outro em nenhum aspeto do processo metodológico. O mesmo acontece com a necessidade ou contingência ontológica dessas hipóteses, que deverão ser consideradas como diferenças de grau, e não como dicotomias extremas (cf. Northoff, 2004a: 119).

Vejam, de seguida, qual o ponto de partida de uma hipótese na Abordagem da Neurofilosofia Não-Redutiva.

V.1.3.1. Uma Perspetiva Baseada e não Reduzida ao Cérebro

A Abordagem da Neurofilosofia Não-Redutiva, que é uma alternativa à Neurofilosofia Redutiva, procura misturar as duas estratégias: reconsiderando as definições conceptuais – a sua plausibilidade conceptual – e ao mesmo tempo procurando estar de acordo com os factos empíricos – plausibilidade empírica. Como vimos, esta metodologia procura assim acomodar as duas conceções numa só: considerando o “input” de Searle com o “output” de Churchland.

Uma novidade estratégica desta abordagem é o seu ponto de partida. Northoff faz notar que grande parte da investigação em Filosofia da Mente e Neurofilosofia pressupõe um ponto de partida da abordagem baseado na mente: apesar de haver diferentes teorias com diferentes caracterizações ontológicas do mental (físico, não-físico, imaterial, emergente, superveniente, etc.), todas estas teorias assumem a mente e os estados mentais como ponto de partida da investigação.

Mesmo as abordagens mais radicais como a Neurofilosofia Redutiva têm no seu âmbito estratégico o projeto de reduzir a mente ao cérebro ou, com mais rigor, reduzir uma teoria do mental a uma teoria do cérebro. O mesmo acontece com a abordagem isolacionista: é a análise do discurso sobre o mental e as suas relações com outros conceitos que tem prevalência.

Nesta perspectiva de abordagem baseada no mental, a questão central que se levanta é a seguinte: como é que a mente e os estados mentais estão relacionados com o cérebro e os estados físicos?

A diferença entre as abordagens mentalistas e fiscalistas é o valor atributivo dado ao mental: nas primeiras, o mental tem um papel positivo cuja estratégia metodológica procura perceber de que forma se pode inferir a existência da mente e a sua relação com os estados mentais. Mas se quisermos negar este papel positivo, como fazem as abordagens de pendor mais fiscalista, a mente continua a ter um papel (negativo) inicial que conduz a investigação: embora ontologicamente as teorias fiscalistas possam ser abordagens baseadas no cérebro/corpo, metodologicamente, ambas as estratégias são abordagens baseadas no papel inicial da mente (cf. Northoff, 2014a: 213-15).

Ora, uma novidade desta nova abordagem é não colocar como ponto de partida a mente, nem a sua consideração positiva ou negativa: antes, o ponto de partida inicial será o cérebro, não no sentido ontológico ou metafísico, mas no sentido metodológico. Em vez da questão central da investigação em Neurofilosofia ser de que forma o mental se relaciona com o físico, a direção de investigação é a oposta: de que forma o físico e as características neuronais e corpóreas estão relacionados com os diversos estados mentais? (cf. Northoff, 2014a: 217).

Mas não será esta inovação uma manobra falaciosa? O conceito de “cérebro” já pressupõe um domínio particular, o empírico, assim como uma metodologia particular para o estudar: as estratégias experimentais da Neurociência. Ora, é neste exato ponto que esta abordagem tem uma forma subtil de resolver o problema (se é ou não bem-sucedida, veremos adiante): o conceito de cérebro, nesta abordagem, não é um conceito meramente científico ou empírico, mas é um conceito neurofilosófico. Que tal significa? Que não se pressupõe nenhum domínio particular *a priori* para o estudar ou investigar; antes, o conceito de cérebro poderá ser associado a diversos domínios (no sentido exposto acima), como sejam o empírico, o epistémico ou o ontológico (cf. Northoff, 2014a: 242).

Desta forma, há um importante desenvolvimento na forma de ver a disciplina da Neurociência: esta deve transformar-se numa “Neurociência Crítica”. Esta perspectiva do trabalho neurocientífico assume que a explicação meramente neuronal não é condição necessária ou suficiente para acomodar

os níveis não-neuronais: é necessário, antes, que as explicações do nível neuronal sejam enquadradas e contextualizadas no seu respetivo contexto.

De que forma é que esta ideia abstrata pode ser colocada em prática? Veja-se, por exemplo, o uso de ressonâncias magnéticas funcionais para estudar o Eu (*se/ñ*) em zonas particulares do cérebro: se não for “encontrado” o Eu nos estudos da atividade cerebral em diferentes zonas do cérebro, não podemos automaticamente inferir que o Eu não existe – tal seria cometer um erro, uma interpretação excessiva ou, logicamente, uma falácia do apelo à ignorância, sendo uma mera confusão entre níveis que não podem ser confundidos.

É necessário, antes, considerar, neste exemplo ilustrativo, informações adicionais que complementem as observações empíricas dado que a aquisição dos dados neuronais, apesar de ser metodologicamente independente, não é filosoficamente neutra, dependendo do seu contexto, da interpretação e da sua conceptualização: é necessário contextualizar os dados obtidos pelas ferramentas e técnicas de investigação neurocientíficas, sendo que a explicação meramente neuronal passa, nesta Abordagem da Neurofilosofia Não-Redutiva, a ser uma condição somente necessária, mas não suficiente, para o estudo do cérebro e das características mentais (cf. Northoff, 2014a: 310).

V.1.3.2. Neurociência da Primeira Pessoa

Como analisámos anteriormente, existe uma lacuna fortemente vincada entre a perspetiva da primeira pessoa, habitualmente trabalhada pela Filosofia tradicional e a perspetiva da terceira pessoa, habitualmente trabalhada pela Neurociência tradicional. Notámos, no capítulo anterior, uma tentativa estratégica de resolver essa lacuna pela conjugação do método neurocientífico com princípios da fenomenologia.

A Abordagem da Neurofilosofia Não-Redutiva, seguindo uma intuição estratégica semelhante, irá procurar fazer uma defesa de uma Neurociência diferente da forma como é usualmente praticada. Para tal, irá fundar uma forma de investigar neurocientificamente o cérebro e os estados mentais – uma Neurociência da Primeira Pessoa: “(...) definimos Neurociência da Primeira Pessoa como uma investigação de estados neuronais sob orientação e orientada para os estados mentais” (Northoff & Heinzl, 2006: 1).

Mas que estratégia é usada para cruzar as duas perspetivas em debate? Eis a definição metodológica desta nova forma de fazer Neurociência:

Neurociência da Primeira Pessoa usa métodos para o exame sistemático e a avaliação dos estados mentais em si mesmos e no seu conteúdo, como experienciado na perspectiva de primeira pessoa, e de que forma se vincula a dados sobre estados neuronais, obtidos na perspectiva de terceira pessoa (Northoff & Heinzl, 2006: 2).

Seguindo esta perspectiva, percebemos finalmente que a mesma se encontra em oposição à forma tradicional de praticar Neurociência, a que podemos chamar de “Neurociência da Terceira Pessoa”, exatamente porque a investigação dos estados neuronais ocorre somente através de métodos empíricos que se enquadram na perspectiva da terceira pessoa. Ademais, esta perspectiva não permite o estudo dos estados mentais na sua subjetividade dado que recusa a irredutibilidade da mesma.

No sentido contrário, a “Neurociência da Primeira Pessoa” procura desenvolver métodos precisos que formem uma união de ligação entre os dados da primeira pessoa e os dados da terceira pessoa. Compreendemos, nesta forma de expor o programa da Neurofilosofia Não-Redutiva, que estamos perante uma atitude ou estratégia inicial muito semelhante à Abordagem da Neurofenomenologia investigada no capítulo anterior: será necessário, contudo, perceber as diferenças entre estas metodologias.

Vejamus de que forma se pode programar uma investigação numa “Neurociência da Primeira Pessoa”. Northoff e Heinzl (2006) criaram um estudo focado nas emoções, dado que as mesmas são inerentemente subjetivas por definição.

Convencionalmente, os estudos que envolvem emoções fazem uso de um instrumento de nome “Sistema Internacional de Imagens Afetivas” (*International Affective Picture System; IAPS*)⁵¹ cujo objetivo é permitir ao investigador classificar, de forma objetiva, as emoções em três categorias: positivas, negativas ou neutras. Ora, este instrumento parece ser um método da perspectiva da terceira pessoa: o investigador pode comparar, depois dessa categorização, os correlatos neuronais dessas emoções obtidos através de uma ressonância magnética funcional, por exemplo.

Todavia, como já foi dito, as emoções são particularmente subjetivas e a própria forma de classificar as mesmas em positivo, negativo ou neutro é baseada numa asserção subjetiva e contínua que pode ser definida como um instrumento da perspectiva da primeira pessoa, pelo que esta forma de

⁵¹ O IAPS é constituído por uma base de dados de imagens que ativam uma resposta emocional particular. O objetivo do instrumento é encontrar um padrão de reação que permita formar uma linha de base que pode ser utilizada universalmente em diversas experiências e, tendo por base diversos públicos, dado que se usa o mesmo estímulo (e.g. as mesmas imagens) nos diferentes sujeitos de estudo.

classificar as emoções em conjunção com dados obtidos pela fMRI pode ser entendida como forma de conjugar os dados da primeira com a terceira pessoa (cf. Northoff & Heinzl, 2006: 4).

De que forma se pode integrar os resultados alcançados com o instrumento de categorização das emoções acima descrito? Uma forma de ligar esses resultados à subjetividade é criar uma “escala visual contínua entre 1 a 9 de valências emocionais” (Northoff & Heinzl, 2006: 4) permitindo cruzar esses dados com correlatos neuronais obtidos pela ressonância magnética. Mas de que forma se pode retirar uma relação direta? Eis a explicação dos autores:

Essa relação direta é possível, uma vez que a experiência emocional foi transformada através de uma escala visual analógica em valores numéricos que podem ser correlacionados com os valores resultantes da medição da fMRI. É preciso considerar que os valores numéricos não são idênticos à própria experiência emocional. Numa próxima etapa, pode comparar-se os resultados da fMRI de ambas as abordagens, a categórica ou de terceira pessoa versus a análise paramétrica ou de primeira pessoa. Se ambas as análises revelarem regiões diferentes, os correlatos neuronais da categorização de emoções na terceira pessoa podem ser separados da experiência na primeira pessoa (paramétrica). Como tal, o neuronal correlaciona-se especificamente à experiência da primeira pessoa e, portanto, os estados mentais podem ser revelados e separados daqueles associados à observação em terceira pessoa (Northoff & Heinzl, 2006: 4-5).

Que resultados foram obtidos nesta investigação? Os autores indicam que as duas diferentes análises (fMRI, ligada à terceira pessoa e a escala visual contínua, ligada à primeira pessoa) revelaram resultados diferentes: enquanto a primeira indicou atividade em várias regiões do córtex cerebral, a segunda indicou atividade somente em partes específicas do cérebro, como estruturas da linha média cortical (CMS), no córtex pré-frontal orbitomedial (OMPFC) e dorsomedial (DMPFC) e o córtex medial parietal (MPC) (cf. Northoff & Heinzl, 2006: 5).

Deste modo, os autores concluem que as estruturas da linha média cortical são especialmente relevantes para a experiência subjetiva das emoções. É esta informação adicional que permite supostamente descobrir conhecimento do cérebro relevante para o estudo da subjetividade que estaria em falta caso a perspectiva da primeira pessoa fosse ignorada.

Se os autores são bem-sucedidos na sua tentativa de unificar uma metodologia que une duas perspectivas fundamentalmente diferentes será uma questão que analisaremos na segunda parte do capítulo dado que poder-se-á estar a confundir níveis que necessitam de ser diferenciados.

V.1.4.1. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o caso do Eu (*Self*)

A natureza do Eu (*self*) tem uma longa história na investigação filosófica sobre este conceito tão relevante (cf. Northoff, 2004b; Zhu & Han, 2008). Historicamente, William James distinguiu entre um “eu físico”, um “eu mental” e um “eu espiritual”. Essas distinções parecem reaparecer nas recentes discussões em Neurociência: Damásio (1999) e Panksepp (2003) sugerem um “proto-eu” nos domínios sensorial e motor que se assemelha à descrição de James do “eu físico”; na mesma linha, o que foi descrito como o “eu mínimo” (Gallagher, 2000) ou o “eu central ou mental” (Damásio, 1999) pode corresponder ao conceito de “eu mental” de James; finalmente, o “eu autobiográfico” de Damásio (1999) e o “eu narrativo” de Gallagher (2000) dependem fortemente da ligação de eventos passados, presentes e futuros com algumas semelhanças com o “eu espiritual” de James (cf. Han & Northoff, 2009: 203-204).

Investigações recentes de neuroimagem dos correlatos neuronais do Eu realçam o papel das estruturas da linha média cortical (CMS) no processamento do Eu (Northoff & Bermpohl, 2004; Northoff *et al.*, 2006). Dado que o Eu de cada indivíduo se desenvolve num contexto sociocultural específico, poderá sofrer fortes modulações desses mesmos contextos sociais e valores culturais, originando uma forma específica de interação eficiente com outros indivíduos em ambientes sociais.

De facto, psicólogos sociais e culturais têm demonstrado que existem diferenças culturais no processamento do Eu e da relação do outro com o próprio Eu (Markus & Kitayama, 1991; Zhu & Han, 2008). As descobertas da Psicologia Social e Cultural levantam questões interessantes sobre se a representação neuronal do Eu e os substratos neuronais do processamento do Eu são moldados por contextos socioculturais ou se são independentes destes.

Nas investigações tradicionais com ressonância magnética funcional, o “Eu” e o “Outro” (ou não-Eu) eram vistos como categorias distintas e, portanto, frequentemente contrastavam na pesquisa funcional de neuroimagem. Ora, Northoff propõe que a forma como o “Outro” é percebido poderá fazer parte da mesma função de autorelação que funciona num contínuo entre o Eu e o Não-Eu que são os limites extremos desse espetro, sendo que um indivíduo (i.e. outro) dos mesmos contextos socioculturais fará parte dessa mesma função contínua estando presente algures na zona intermédia do espetro podendo ser descrita como “parecido com o Eu, mas não Eu” ou um “Quase-Eu”.

Este espetro é relevante, porque as diferentes posições (Eu, Não-Eu, Quase-Eu, etc.) do espetro estão presentes (ou ausentes) nas mesmas regiões do cérebro, sendo que variam, como veremos, no grau de ativação dessas zonas, permitindo que o organismo se defina a si mesmo em relação a membros

do mesmo e diferente grupo social ou cultural, o que lhe permite prever comportamentos, desenvolver emoções ou sentimentos específicos (e.g. empatia) e antecipar ambientes com alto nível de complexidade.

Este “mapa interno” permite oferecer uma vantagem seletiva durante a evolução pelo que faz sentido que o mecanismo do Eu seja altamente desenvolvido em seres humanos e espécies geneticamente próximas do mesmo.

O importante deste exemplo é o seguinte: a proposta inicial de distinguir o Eu do Não-Eu, como fazem as investigações tradicionais, enviesando os próprios estudos científicos e originando resultados contraditórios, necessita de ser atualizado à luz dos indícios que faremos notar de seguida. Porém, essa atualização necessita ainda de uma reconceptualização que inclua uma plausibilidade conceptual relevante. Esta forma de balançar entre o empírico e o conceptual iterativamente é a marca da Abordagem da Neurofilosofia Não-Redutiva. Vejamos de seguida os detalhes da discussão.

Há dois tipos de resultados contraditórios em relação a zonas específicas do cérebro relacionadas com o processamento do Eu:

- (a) Vários estudos de imagética relatam uma sobreposição entre o processamento do Eu e o processamento do Não-Eu, especialmente no córtex préfrontal dorsomedial (DMPFC) e no córtex préfrontal medial ventral (VMPFC) (Schmitz *et al.*, 2004; Platek *et al.*, 2004; Seger *et.* 2004; Beer *et al.*, 2006).
- (b) No entanto, foi observada dissociação neuronal entre o Eu e Não-Eu nas mesmas regiões de (a), mas também em outras regiões corticais como o córtex préfrontal, parietal e temporal lateral (Craig, 2002; Kelley *et al.*, 2002; Ochsner *et al.*, 2005).

Ora, de que forma se podem integrar estes factos empíricos contraditórios num esquema conceptual consistente? Os estudos de (a) e (b) podem ser conjugados da seguinte forma coerente: se o conceito de Eu enquanto diferente do Não-Eu for tido como um espectro contínuo, e não como uma bifurcação absoluta e independente, poderemos tornar consistente a suposta incoerência empírica.

Em relação à atividade cerebral, Northoff elucida que (a) e (b) podem ser interpretados da seguinte forma: quanto mais o Não-Eu for identificado com ou próximo do Eu, maior será a semelhança das respostas neuronais entre (VMPFC) e (DMPFC) e as restantes regiões do cérebro. Por exemplo: um indivíduo português caucasiano irá reagir de forma mais semelhante ao ver-se a si mesmo ao espelho

do que a ver um outro indivíduo alemão caucasiano. Contudo, se perceber um indivíduo africano de raça negra, a ativação nas zonas indicadas será menor.

Três estudos de Jason P. Mitchell, da Universidade de Harvard, apontam para este resultado, descobrindo que quanto mais semelhante os sujeitos da experiência julgavam a sua face em relação à de outros, maior ativação era observada na (VMPFC), sugerindo que a mesma possa contribuir como evidência para a nova conceptualização de conceber o Eu enquanto Não-Eu como um espectro contínuo de dois extremos. Quanto maior a semelhança, maior a ativação; quanto maior a dissemelhança, menor a ativação (cf. Mitchell *et al.*, 2005a; 2005b; 2006).

Conceptualmente, incorporando estes dados empíricos, passamos de uma dicotomia explícita para um contínuo explícito entre o Eu e Não-Eu. Esta diferença conceptual pode ser historicamente situada. Segundos os autores, trata-se da forma como a Filosofia Ocidental e Oriental interpretam a existência do Eu: na segunda, há uma crença de que o Eu está fortemente associado às relações humanas, aos contextos sociais e aos contextos culturais (cf. Markus & Kitayama, 1991; Zhu & Han, 2008). Porém, na primeira, há uma crença na independência do Eu, que se mantém invariável em contextos sociais e culturais.

Ora, será que essas diferentes formas de interpretar o Eu e a sua relação com os outros influencia de alguma forma os substratos neuronais ligados ao processamento do Eu? Para descobrir a resposta, investigou-se dois grupos culturais: um grupo ocidental (de língua inglesa) e um grupo oriental (de língua chinesa). O objetivo da tarefa era julgarem-se a si mesmos e a uma pessoa próxima, a progenitora de cada indivíduo.

A atividade neuronal de autoreferência foi localizada no córtex préfrontal medio (MPFC) e no córtex cingulado anterior (ACC), zonas que também eram ativadas no grupo oriental quando confrontados com um Eu próximo (a progenitora). Curiosamente, essas zonas não mostraram um aumento de ativação com o grupo ocidental. Pode ser concluído que os orientais fazem uso da (MPFC) para se representarem a si mesmos e à progenitora, enquanto que os ocidentais usam essa zona exclusivamente para se representarem a si mesmos.

Um segundo estudo aponta para a mesma tese: Sui, Liu e Han (2009) registaram potenciais de eventos relacionados com eletroencefalografia (EEG) a indivíduos britânicos e chineses quando julgavam a sua face em contraste com o julgamento da face de um familiar. A primeira evidência captada foi a de que, em ambos os grupos, havia uma resposta mais rápida à percepção da sua própria face em relação à percepção da face de um familiar (algo que seria esperado). Contudo, de forma interessante, o grupo chinês demorou mais tempo a discernir a diferença entre a avaliação da sua face em relação à face de

um familiar, sendo que o grupo britânico foi muito mais rápido a discernir entre a sua face e a face de um familiar.

Novamente, tal aponta para favorecer uma reconceptualização a favor do Eu como um espectro contínuo e contra a perspectiva de uma dicotomia vincada entre o Eu e o Não-Eu. Assim, há uma relação empírico-conceptual que deve ser considerada no seu conjunto.

V.1.4.2. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o Eu e a COVID-19

Esta abordagem pode ser aplicada a todos os fenómenos investigados tanto pela Filosofia, Psicologia ou Neurociência como outras disciplinas como a Inteligência Artificial. Veja-se, por exemplo, a aplicação desta abordagem em relação ao impacto da nova COVID-19 no nosso sentido do Eu (*self*) e da sua relação com os outros e o mundo.

Northoff investigou esse impacto fazendo notar que este vírus afeta o nosso Eu (*self*) de três formas: (i) o medo de sermos infetados pelo mesmo; (ii) o medo de infetarmos outras pessoas (principalmente familiares); (iii) o medo da perda das relações sociais (cf. Scalabrini *et al.*, 2020: 46).

Como vimos, o Eu está muito ligado à relação com o mundo e os outros: devido às variadas medidas de prevenção associadas à luta contra a COVID-19, existe uma perda substancial da nossa subjetividade, principalmente devido ao medo existencial que aparece como resposta emocional às adversidades levantadas pela pandemia.

De forma interessante, os autores do estudo indicam que essas emoções devem ser enquadradas numa perspectiva anticartesiana, não se reduzindo meramente a uma investigação cerebral para as compreender na sua totalidade, mas, antes, redefinido o Eu como uma entidade ou instância que está em relação com o corpo e o mundo:

Podemos ver que, mesmo em animais, existe um sistema subcortical de medo-ansiedade básico. Esse mesmo sistema recebe em particular diversos *inputs* do corpo – estamos, portanto, intimamente conectados ao nosso corpo e essa relação pode se transformar em medo e ansiedade se ocorrerem discrepâncias entre estímulos ínter e êxtero-receptivos. O medo e a ansiedade podem, portanto, ser vistos como estratégias do cérebro para se reequilibrar e, por fim, “normalizar” os seus *inputs* ínter e êxtra-receptivos – o cérebro

estabelece uma relação com o corpo, manifestado no medo-ansiedade. Mais profundamente, diz que mesmo uma emoção tão básica quanto o medo não é isolada e “fechada” no cérebro ou no organismo. Em vez disso, os dados mostram que o medo é intrinsecamente relacional, algo constituído neuro-vegetativamente com base nos *inputs* interoceptivos do corpo. Tal, por sua vez, estabelece relações com o meio ambiente e o mundo – tomadas num sentido mais geral, as emoções como medo-ansiedade são, portanto, sobre o equilíbrio entre mundo, corpo e cérebro e, portanto, são genuinamente relacionais e neuro-ecológicas (Scalabrini *et al.*, 2020: 47).

É importante notar que a ansiedade pode ser definida como uma antecipação preditiva do organismo (cf. Grupe & Nitschke, 2013: 488-501) que pode ser útil em várias ocasiões, mas que pode ser negativa se a sua frequência ultrapassar as situações em que fará sentido para o organismo processar essa resposta emocional.

As teorias do processamento preditivo (cf. Capítulo 8) podem ser úteis para explicar este fenómeno. Estas teorias propõem que o cérebro gera modelos antecipados tanto das condições externas (e.g. estímulos visuais ou auditivos) como as condições internas do próprio organismo (e.g. ritmo cardíaco) (cf. Seth, Suzuki, & Critchley, 2012; Seth, 2013). Quando o *input* antecipado é diferente do *input* factualmente recebido, o cérebro gera uma correção do modelo anterior tentando reduzir ao máximo o erro de predição.

É esperado que, com uma condição alarmante como a COVID-19, o organismo crie um modelo antecipativo que aumente o ritmo cardíaco. Contudo, se houver uma falha na expectativa do aumento do ritmo cardíaco, sinais de “cima para baixo” (*top-down signals*) levarão a esse aumento para corresponder ao modelo preditivo criado inicialmente (cf. Barrett & Simmons, 2015: 244-247).

Dado que a ansiedade é uma resposta emocional caracterizada por uma indeterminação ou incerteza constante, levará à produção de expectativas de excitação elevadas que não serão atualizadas ou correspondidas pela resposta não-preditiva (i.e. real) contrária, como seja a ausência do aumento do ritmo cardíaco (cf. Paulus, Feinstein & Khalsa, 2019: 98).

Não havendo correspondência entre o modelo antecipado ou preditivo e a resposta do organismo, estaremos perante um erro preditivo de amplitude elevada, fazendo com que essa mesma resposta (somática) do organismo seja demasiado ampliada para tentar corresponder ao modelo criado anteriormente, resultando exatamente nessa resposta emocional de ansiedade e medo acrescida (cf. Scalabrini *et al.*, 2020: 47).

Northoff acrescenta uma parte importante desta explicação: a estrutura espaçotemporal do organismo – que dá atenção à estrutura temporal da atividade neuronal em conjunção com os padrões espaciais do cérebro (Northoff, 2016a) – irá representar o *a priori* na fórmula de Bayes do modelo preditivo (cf. Capítulo 8).

Assim, tendo em conta a estrutura espaçotemporal, percebemos de que forma os *inputs* somáticos (i.e. da atividade corporal) deverão também ser integrados (cf. Northoff, 2014b: 372) pelo que a sua não-integração levará a um erro preditivo elevado, que poderá estar conectado com os próprios sintomas de ansiedade elevada por dois mecanismos principais: (i) o equilíbrio exteroceptivo anormal na relação mundo-corpo-cérebro e (ii) o equilíbrio interoceptivo anormal na relação cérebro-corpo-mundo.

Dado que, como vimos, na Abordagem da Neurofilosofia Não-Redutiva o Eu é social, embutido no mundo e no Outro, o sentido de medo e ansiedade causado pela COVID-19 afeta a nossa relação com o Eu diminuindo a nossa relação espaçotemporal com o outro: o espectro do “Eu/Não-Eu” (cf. I.1.4.1.) desloca-se mais para a direita do mesmo, isto é, para um desligar mais enfatizado com o outro o que, por sua vez, causa uma diminuição do nosso sentido do Eu. Em conjunção com a ameaça de morte de cada um e dos outros, o organismo entra em estados profundos de ansiedade “existencial” como descrito pela filosofia existencialista, sendo esta tese apoiada por diversos estudos empíricos (cf. Scalabrini *et al.*, 2020: 50), como indicado pelos investigadores:

A ameaça do coronavírus apresenta uma dupla ameaça – o colapso da nossa subjetividade, devido ao risco do nosso corpo ser infetado, e o colapso de nossa intersubjetividade, pelo isolamento social de cada um em relação aos outros. A dupla ameaça da subjectividade e intersubjetividade fornece uma dimensão profunda às nossas emoções, incluindo os nossos medos e ansiedades, resultando em medo existencial. Essa conexão íntima do medo, do Eu (*self*) e da existência deve-se à maneira como o nosso cérebro e sua atividade espontânea são organizados, ou seja, o primeiro estabiliza-se alinhando-se (*alignment*) ao corpo e ao mundo da mesma maneira que estabilizamos os nossos movimentos durante uma dança, alinhando-os ao ritmo da música. Esse mesmo alinhamento do nosso cérebro com o corpo e o mundo, ou seja, o alinhamento espaçotemporal, está ameaçado, se não mesmo interrompido pela crise do Coronavírus – do mesmo modo, quando se dança erráticamente ao ignorar o ritmo da música, tornamo-nos existencialmente ameaçados e erráticos quando o cérebro e, no fundo, o alinhamento neuroecológico espaçotemporal do Eu (*self*) ao mundo está ameaçado pelo vírus e pelas medidas de defesa do distanciamento físico e social (Scalabrini *et al.*, 2020: 53).

Desta forma, compreendemos como é que o ecletismo metodológico da Abordagem da Neurofilosofia Não-Redutiva pode servir para analisar vários conceitos e temas de uma perspectiva interdisciplinar. Vejamos, de seguida, um último exemplo relacionado com as emoções.

V.1.4.3. Exemplo de Investigação da Neurofilosofia Não-Redutiva: o caso das Emoções

Um último exemplo pode ser avançado para compreender o poder explicativo da Abordagem da Neurofilosofia Não-Redutiva. Ademais, este exemplo mostrar-nos-á de que forma esta abordagem se distingue das propostas corpóreas (*embodied*) mais tradicionais tal como as que foram analisadas no capítulo anterior.

A Teoria das Emoções de James–Lange, avançada historicamente por William James e, numa versão similar, por Carl Lange, defende que as emoções se devem fundamentalmente a alterações corporais, quer a nível fisiológico, hormonal ou motor (cf. James, 1884: 188-205). Nas palavras de James: “(...) sentimos pena porque choramos, raiva porque batemos, medo porque trememos, e não choramos, batemos ou trememos porque sentimos pena, raiva ou medo, conforme o caso” (James, 1884: 190). Ora, esta teoria reverte a forma comum de interpretar as emoções: aqui, o indivíduo reage primeiro (e.g. a fugir de algo) e só depois irá interpretar essa reação (e.g. de medo) como uma resposta emocional determinada que funciona como explicação para a acção prévia.

Esta teoria tem ressurgido de uma perspectiva mais empírica com variados estudos (e.g. Bechara, 2004; Niedenthal, 2007) que procuram demonstrar o papel fundamental do corpo nas emoções e, por consequência, da própria subjetividade (cf. Colombetti & Thompson, 2005; 2007). Qual é, nesta explicação, o papel do ambiente (i.e. do mundo)? Há duas formas de interpretar o papel do mesmo nas emoções:

- (i) Indiretamente, o corpo representa o ambiente através de representações corpóreas sensoriomotoras que irão modular a resposta emocional (teorias corpóreas das emoções);

- (ii) Diretamente, sendo que o ambiente é constitutivo da própria resposta emocional, que é diretamente relacionada com o organismo sem necessitar de um intermediário indireto como seja o corpo (teoria relacional das emoções).

Em (i), a relação entre o corpo e o ambiente é, no máximo, instrumental, sendo que há uma contribuição do segundo no primeiro, mas não há uma constituição fundamental. Já em (ii), estamos perante uma relação não-instrumental: mesmo que o *input* sensorial se mantenha, há a possibilidade de o conteúdo da percepção variar com a variação do *output* motor, havendo uma independência e uma relação direta que não se encontra em (i) (cf. Northoff, 2008: 506). Como afirma o neurofilósofo alemão:

Se a relação entre o sentimento emocional e o ambiente for direta e, portanto, constitucional, ou seja, não instrumental, as mudanças no ambiente devem poder impactar e constituir os sentimentos emocionais independentemente da representação sensoriomotora do corpo. O próprio ambiente pode então estar diretamente envolvido na constituição das emoções e sentimentos (Northoff, 2008: 506).

Mas haverá alguma evidência empírica que aponte mais em direcção a (ii) do que a (i)? Um estudo avançado por Fliessbach *et al.* (2007) procurou oferecer precisamente evidências a favor desta tese, criando a seguinte experiência científica: dois sujeitos, A e B, receberam uma recompensa em dinheiro enquanto os seus cérebros estavam a ser analisados por fMRI. O sujeito A iria receber a mesma quantia fixa, enquanto o sujeito B veria a sua recompensa variar. A ideia é perceber o impacto de algo externo ao sujeito (i.e. ambiente) como o dinheiro nos sujeitos humanos. Igualmente importante é o facto de o sujeito A saber que quantia específica o sujeito B iria receber, podendo assim comparar constantemente os valores recebidos por cada um.

Ora, de forma interessante, as reacções emocionais no circuito neuronal da recompensa não variaram consoante a diferença na quantia do dinheiro recebido pelo sujeito A, mas, antes, variou quando este comparava a quantia de dinheiro recebido pelo sujeito B. Por exemplo: se o sujeito A recebesse 60 euros e o sujeito B recebesse apenas 30 euros, então o sujeito A aumentaria a atividade no circuito neuronal da recompensa e também o seu estado de felicidade. O oposto também acontecia: caso o sujeito A recebesse os mesmos 60 euros mas o sujeito B recebesse 120 euros, então o aumento de ativação dos circuitos de recompensa não ocorria nem o sujeito se sentia mais feliz (cf. Fliessbach *et al.*, 2007: 1306).

Assim, a partir destas evidências, Northoff conclui que, apesar do *input* sensório-motor ser exatamente o mesmo (i.e. 60 euros), a reação emocional do sujeito A dependia fundamentalmente da recompensa monetária que o sujeito B iria receber quando comparada com a sua. Tal aponta para a tese argumentada em (ii): as emoções não são apenas instrumentalmente dependentes do ambiente, mas, pelo contrário, o ambiente é essencialmente constitutivo da resposta emocional:

[...] é a relação entre a pessoa e o ambiente a real diferença ou balanço entre as recompensas dos sujeitos A e B que parecem determinar os sentimentos emocionais e atividade neuronal. É uma dependência constitucional, isto é, não-instrumental dos sentimentos emocionais sobre o ambiente e sua relação com a pessoa [...] (Northoff, 2008: 506).

Mas de que forma se distingue, assim, a abordagem das concepções corpóreas de Northoff analisadas anteriormente? Esta abordagem “relacional” das emoções – como o neurofilósofo lhe chama – altera o foco de atenção da investigação do corpo para o papel ativo do ambiente nas emoções: em vez da modulação destas passar por representações corpóreas indiretas, o ambiente passa a ser uma componente diretamente constituinte das emoções.

Mas será que basear a defesa da tese a favor de (ii) em vez de (i) num único exemplo de experiência científica é suficiente? Será epistemicamente sensato perceber se existe um conjunto de evidências que aponta para (ii), e não basear uma conclusão neurofilosófica apenas de um exemplo particular?

Há uma distinção fundamental que poderá ajudar-nos a compreender por que razão devemos entender que as emoções devem ser conceptualizadas como relacionais em vez de somente corpóreas. Deve-se entender que, de certa forma, a teoria relacional das emoções apresentada por Northoff consegue incorporar, de certa forma, a teoria corpórea: a grande diferença, como iremos ver, é o papel atribuído ao corpo numa e noutra – em (i), o corpo é o início e o fim; todavia, em (ii) o corpo é também o início da explicação, mas não se reduz apenas a ele, sendo antes uma hipótese baseada neste.

Para entendermos melhor a defesa, então, da teoria relacional das emoções, devemos atender a uma distinção fundamental: entre estímulos exeroceptivos (EE) e estímulos interoceptivos (EI) nas emoções. Os (EE) são estímulos que provêm do ambiente, de fora do sujeito. Em contrapartida, (EI) são estímulos que provêm do próprio corpo, dentro do sujeito (cf. Capítulo 8).

Ora, se (i) está correta, favorecendo uma teoria corpórea das emoções, podemos prever empiricamente que haja uma maior prominência de (EI) e das regiões do cérebro associadas – e, portanto, dos estados fisiológicos do corpo – ao processamento dos mesmos durante as emoções, enquanto os (EE) devem ter um papel diminuto ou não-constitutivo. Mas será realmente isso que podemos confirmar empiricamente (cf. Northoff, 2008: 509)?

Northoff, baseado em dados da Neuroimagemologia, reúne os seguintes indícios a favor de (ii) e contra (i): se a consciência introspectiva – ligada às sensações internas corporais – estiver relacionada somente com (EI), sendo independente de (EE), então tal favorecerá (i) uma teoria corpórea das emoções. Se o contrário for o caso, isto é, se a consciência introspectiva envolver uma relação entre (EI) e (EE), então irá beneficiar (ii) uma teoria relacional das emoções.

Assim, precisamos de investigar qual das teorias, (i) ou (ii), tem uma plausibilidade empírica maior. Vejamos, assim, o que os estudos empíricos indicam e que consequências conceptuais irão acarretar.

O grupo em redor de Critchley *et al.* (2004) pediu aos sujeitos da experiência para avaliarem se o seu ritmo cardíaco era sincrónico ou assincrónico com um som em *feedback* auditivo que permitiu comparar a atenção dada a (EE) – o som externo – em relação à atenção dada a (EI) – o ritmo cardíaco, interno.

Os resultados obtidos foram os seguintes: em relação à atenção dada ao ritmo cardíaco, houve um aumento da atividade neuronal na ínsula direita, no córtex cingulado anterior dorsal (dACC), no córtex cingulado anterior subgenual (sACC) e no córtex somatomotor. Em contrapartida, em relação à atenção dada ao som auditivo, a atividade neuronal nessas regiões foi suprimida. Por fim, verificou-se também que a atividade neuronal na ínsula direita era correlacionada com sintomas de ansiedade subjetiva e com a performance da deteção do ritmo cardíaco que se correlacionavam igualmente um com o outro (cf. Critchley *et al.*, 2004: 192-193).

Para os autores do estudo, estes resultados⁵² favoreceriam uma relação próxima entre as emoções e a consciência introspectiva, sendo que as regiões cerebrais indicadas funcionaram como intermediárias entre os dois fenómenos, dando a entender que os (EE) seriam somente instrumentais ou modulatórios dessa relação, favorecendo, assim, um suporte empírico a (i) e contradizendo (ii).

⁵² Northoff analisa outras investigações semelhantes (cf. Pollatos, Herbert *et al.*, 2007; Pollatos, Traut-Mattausch *et al.*, 2007; Critchley, 2005) mas que ficaram de fora por falta de espaço e por irrelevância epistémica, evitando-se desse modo repetições de resultados semelhantes que já são instanciados por este exemplo particular.

Ora, Northoff não concorda de todo com esta forma de interpretar os dados por três razões principais:

- Os variados estudos analisados não investigaram os (EI) em isolamento dos (EE) mas relacionados entre si, sendo uma falha metodológica importante dado que se assumiu que a atividade neuronal associada à consciência introspetiva estava somente associada ao ritmo cardíaco (i.e. EI), mas pode estar igualmente associada com o som percebido (i.e. EE), refletindo, assim, ao contrário da interpretação de (i) que assume um isolamento independente entre (EI) e (EE), um balanço dinâmico entre o processamento exteroceptivo e o processamento interoceptivo favorecendo (ii);
- Os variados estudos analisados não avaliaram a questão da valência emocional, isto é, a avaliação positiva ou negativa do caráter das emoções pelos sujeitos (cf. Colombetti, 2005: 103) dado que, por exemplo, o estudo apresentado por Pollatos *et al.* (2005) não apresentou uma diferença considerável entre os indivíduos que tinham uma elevada performance em identificar o ritmo cardíaco daqueles que tinham baixa performance, apesar dos dois grupos divergirem substancialmente na excitação emocional; ora, para Northoff, tal não faz sentido dado que as regiões cerebrais associadas à consciência introspetiva estão altamente interligadas com as regiões do cérebro associadas à valência emocional, pelo que “[...] o padrão de conectividade favorece, portanto, fortemente o conceito relacional ítero-exteroceptivo do sentimento emocional que parece tornar o processamento interoceptivo isolado e, portanto, um conceito interoceptivo do sentimento emocional, bastante improvável” (Northoff, 2008: 512);
- Por fim, Pollatos *et al.* (2005) investigaram a duração temporal da atividade neuronal durante a tarefa de identificar o ritmo cardíaco com eletroencefalografia (EEG), observando que os indivíduos com maior performance tinham maior força dipolar em zonas corticais como (dACC), (sACC), a ínsula direita e o córtex somatossensorial secundário do que os indivíduos com baixa performance; ora, além disso, foi descoberto que a fonte de dipolo⁵³ ocorria no (dACC), (sACC) e no córtex pré-frontal dorsomedial (CPFDm) mais cedo – por volta de 280 milissegundos – do que na ínsula e no córtex somatossensorial – que ocorria por volta de 370 milissegundos; uma distribuição temporal semelhante pode ser encontrada no estudo de Tsuchiya e Adolphs (2007);

⁵³ Os potenciais elétricos nos neurónios podem ser modelados por dipolos, que são compostos de duas cargas opostas e de mesma intensidade separados por uma curta distância, servindo para medir a atividade químico-elétrica do neurónio (ou do conjunto de neurónios).

ora, se (i) fosse plausível, seria de esperar que os padrões temporais das duas regiões cerebrais fossem opostos, dado que é a ínsula e o córtex somatossensorial que estão mais ligados ao processamento interoceptivo; além disso, as regiões corticais associadas a (dACC), (sACC) e (CPFDM) têm sido associadas ao processamento dos (EE), principalmente estímulos ligados ou relacionados com o organismo (Northoff *et al.*, 2006; Northoff e Bermppohl, 2004); ora, para Northoff, o facto dessas regiões estarem associadas desde cedo – em termos temporais – à consciência introspectiva indica que os (EE) possam estar envolvidos no processamento dos (EI) e, além disso, dado que esse envolvimento não acontece somente numa fase tardia – o que favorecia a tese (i) de que os (EE) funcionam somente como processos modulatórios ou instrumentais de (EI) – mas acontece numa fase inicial, podemos igualmente concluir que os (EE) são de facto constitutivos de (EI), sendo mais uma evidência a favor de (ii) e contra (i) (cf. Northoff, 2008: 512-513).

A partir destas evidências, Northoff conclui que devemos reconceptualizar as nossas teorias das emoções a favor de uma teoria relacional e contra uma teoria corpórea das emoções. Especificamente, os sentimentos emocionais não podem ser teorizados como baseados somente nos (EI) ou como perceções das mudanças fisiológicas do corpo como assume (i) e a teoria das emoções James-Lange. Pelo contrário, se definirmos a estrutura teórica como uma relação entre (EE) e (EI) – e não uma dicotomia – poderemos focar com mais rigor empírico e conceptual a relação entre o corpo e o ambiente em vez do papel independente de cada parte da equação (i.e. do corpo, ou do ambiente, mas tidos como isolados), sendo assim uma teoria relacional das emoções mais plausível quer em termos de plausibilidade empírica como em termos de plausibilidade conceptual (Northoff, 2008: 518).⁵⁴

Este é mais um exemplo da diversidade que a Abordagem da Neurofilosofia Não-Redutiva permite nas suas investigações. De seguida, iremos levantar as principais objeções a esta abordagem.

⁵⁴ Poderá ser argumentado que o aspeto relacional e ambiental é irrelevante no âmbito de uma teoria das emoções: identificar zonas ativadas neuronalmente não esclarece nenhum aspecto particular. Ora, esta objeção só se dá se uma leitura reducionista do que foi exposto ocorrer: o cerne da experiência é exatamente demonstrar como é que uma alteração ambiental afeta a consciência do sujeito. A análise da atividade neuronal é apenas um dos aspetos relevantes, mas não o único: é necessário enquadrar esse elemento nos restantes, i.e., no elemento externo, ambiental, e no elemento interno, a própria experiência consciente das emoções sentidas.

V.2. Reflexão Crítica:

Depois de termos refletido sobre as várias dimensões da Neurofilosofia Não-Redutiva, focaremos agora a nossa atenção para as principais críticas e problemas desta abordagem metodológica.

V.2.1. Objeção do “Erro de Categoria e da Modalidade”

A primeira objeção que pode ser levantada à Abordagem da Neurofilosofia Não-Redutiva é que esta comete, de certa forma, o mesmo erro falacioso⁵⁵ apontado à abordagem reducionista, atribuindo ao cérebro variadas capacidades mentais que apenas podem, logicamente, ser atribuídas à pessoa, isto é, à totalidade do organismo e não a uma parte do mesmo.

Dado que, como vimos, esta abordagem caracteriza o cérebro como ponto de partida, sendo um objeto físico dentro do mundo natural, em contrapartida com a mente que pode ser considerado um objeto subjetivo dentro do mundo lógico, estamos perante duas categorias absolutamente diferentes que requerem investigações opostas.

No segundo capítulo, analisámos com detalhe as consequências metafisológicas desta posição: a abordagem isolacionista caracteriza-se exatamente pela defesa de que diferentes disciplinas devem estudar diferentes fenómenos. A Filosofia deveria focar-se no estudo do carácter e relações lógicas das proposições através de métodos *a priori*. No sentido oposto, a Neurociência deveria focar-se no estudo dos objetos físicos através de métodos *a posteriori*.

Ora, seguindo a objeção, não se poderá inferir qualquer reflexão relevante investigando o cérebro do ponto de vista da Filosofia, dado que o cérebro não é um objeto que possa ser estudado, metodologicamente, pela Filosofia – tal seria, então, cometer um erro de categoria, confundindo e misturando categorias que não podem ser cruzadas por definição (cf. Keil & Schnaedelbach, 2000).

Assim, esta objeção pressupõe que há uma distinção exclusiva entre o cérebro e a mente, por um lado, e o mundo natural e lógico, por outro. Porém, tal seria um problema se, de facto, não existisse uma região do mundo natural que se cruzasse efectivamente com o mundo lógico. Como afirma Northoff:

⁵⁵ Esta objeção é inspirada por Gilbert Ryle que, como vimos anteriormente, investigou exatamente este tipo de falácia que comete erros de categorização entre níveis que devem ser claramente distinguidos (cf. Capítulo 2).

Somente se não houver sobreposição entre condições naturais e lógicas, a caracterização do cérebro e da mente como categorias diferentes pode ser mantida. No entanto, existe alguma sobreposição entre condições naturais e lógicas em relação ao cérebro e à mente. O cérebro como um "cérebro físico", subjacente às condições naturais por si só, deve ser considerado como uma condição natural necessária para a possibilidade de gerar condições lógicas. Continuamos incapazes de executar o raciocínio filosófico e, portanto, incapazes de explicar as condições lógicas sem o cérebro. Embora as condições lógicas possam permanecer independentes do cérebro em si mesmas, o cérebro deve pelo menos ser considerado uma condição natural necessária para a possibilidade da sua criação. O cérebro pode então servir como uma "ponte entre condições naturais e lógicas" e, assim, como uma "janela para a mente". A distinção entre condições naturais e lógicas em relação ao cérebro e à mente não pode mais ser considerada mutuamente exclusiva (Northoff, 2004: 122).

Assim, dado que a ideia de haver uma exclusividade mútua entre os dois mundos se encontra equivocada, por consequência, a tese de que a mente e o cérebro são também de diferentes categorias estará também equivocada.

Mas de que forma este cruzamento categorial não coloca em causa a modalidade das hipóteses neurofilosóficas? Dado que o mundo natural é contingente por definição e o mundo lógico é necessário por definição, será que o estatuto modal das hipóteses neurofilosóficas se encontra em causa?

Como detalhámos, a Abordagem da Neurofilosofia Não-Redutiva pressupõe, em variados passos iterativos, a introdução de elementos *a posteriori* com elementos *a priori* e vice-versa. Ora, tal poderá colocar em causa qualquer tentativa de criar hipóteses que sejam necessárias (isto é, verdadeiras em todos os mundos possíveis), algo que costuma caracterizar teorias filosóficas robustas e sólidas, dado que há cruzamentos de elementos contingentes com elementos necessários em termos epistémicos.

Ora, tal poderá ser uma imprecisão, novamente. Esta crítica assume que a introdução de elementos contingentes – através da introdução de elementos *a posteriori* – nas hipóteses neurofilosóficas implica a total eliminação da necessidade modal, interpretando o binómio contingente versus necessário como mutuamente exclusivo. Todavia, é precisamente esta tese que um defensor da Abordagem da Neurofilosofia Não-Redutiva irá recusar: como vimos, a natureza destas hipóteses não é somente caracterizada pela "plausibilidade empírica" (contingente), mas também pela "plausibilidade lógica" (necessária):

Como a “hipótese neurofilosófica” requer “plausibilidade empírica e lógica”, a contingência e necessidade podem coexistir para que não sejam mais mutuamente exclusivas. A introdução de traços de contingência, portanto, não elimina necessariamente a necessidade por completo – a hipótese neurofilosófica não pode ser considerada como necessariamente inválida em termos ontológicos. Dado que o mesmo é igualmente verdadeiro no caso do *a priori* e do *a posteriori*, a hipótese neurofilosófica não pode também ser considerada necessariamente inválida em termos epistêmicos (Northoff, 2004: 122).

Fica, assim, claro que esta primeira objeção falha dado que pressupõe teses de assimetria total entre o *a priori* versus o *a posteriori* e a contingência versus a necessidade que um defensor da Abordagem da Neurofilosofia Não-Redutiva não aceita, preferindo considerar estes como extremos de cada posição e não como dois extremos mutuamente exclusivos e independentes um do outro, mas, antes, como fazendo parte de um espectro que poderá tendencialmente variar em termos de graus de aprioridade e aposterioridade dependendo da idiosincrasia de determinada hipótese neurofilosófica.

V.2.2. Objeção da “Irrelevância Metodológica”

Uma segunda objeção que pode ser avançada contra a Abordagem da Neurofilosofia Não-Redutiva é a ideia de que há uma certa irrelevância metodológica presente nesta abordagem. Ao colocar-se no meio ou no limite tanto da investigação filosófica tradicional como da investigação neurocientífica tradicional, esta objeção pretende avançar com a tese de que nenhuma hipótese neurofilosófica – como apresentada anteriormente – pode ser relevante tanto do lado da Filosofia como do lado da Neurociência, exatamente porque há uma defesa de um certo carácter de excecionalidade da versão da Neurofilosofia de Northoff, dado que não se pode reduzir nem à disciplina filosófica nem neurocientífica.

Assim, argumenta-se que a Neurofilosofia concebida deste modo peculiar é tão irrelevante para a Filosofia como para a Neurociência, dado que não pode, por definição, contribuir para nenhuma das disciplinas devido às suas diferenças metodológicas e metafilosóficas (que analisaremos de seguida). Tal acontece porque não se compreende, especificamente, de que forma o cruzamento é possível entre as duas disciplinas sem ser algo considerado arbitrário.

Portanto, esta objeção procura reclamar da implausibilidade metodológica da estratégia não-redutiva: por um lado, defende-se que nem a Filosofia nem a Neurociência, por si só, são estratégias eficazes para lidar com o problema do cérebro consciente; depois, defende-se que devemos dar um salto

de fé metodológico criando uma nova metodologia que cruza elementos das duas disciplinas anteriores sem, contudo, se reduzir a nenhuma delas; finalmente, argumenta-se que tal estratégia metodológica, que difere tanto da metodologia filosófica como da metodologia neurocientífica, possa causar algum progresso nas duas disciplinas.

É efetivamente este último passo que parece ser dúbio, filosoficamente: se x é diferente de y , não se percebe como é que y pode causar progresso em x sem assumir-se que tal se dá por mero acaso ou seguindo um princípio arbitrário implausível ou, ainda, sendo um mero enviesamento casuístico do defensor da Abordagem da Neurofilosofia Não-Redutiva:

O desenvolvimento de “hipóteses neurofilosóficas” pode ser considerado como irrelevante e desnecessário para a filosofia e a neurociência. Dado que as hipóteses neurofilosóficas são “misturas brutas e arbitrárias” entre hipóteses empíricas e pressupostos teóricos, as primeiras permanecem incapazes de fazer contribuições significativas tanto à filosofia como à neurociência. E se, portanto, a hipótese neurofilosófica não pode contribuir para nenhuma disciplina, então deve ser considerada irrelevante em geral (Northoff, 2004: 122).

Ora, será que há alguma forma de escapar a esta objeção? O defensor desta metodologia pode argumentar contra a ideia de que a mesma representa uma metodologia arbitrária entre o cruzamento de duas disciplinas tradicionalmente opostas quanto ao seu método e objecto de estudo.

Assim, esta tese pode ser rejeitada defendendo que o cruzamento iterativo da metodologia de Northoff é não-arbitrário, isto é, sistemático. Se conseguirmos demonstrar que há uma sistematização rigorosa da dança iterativa entre as duas disciplinas (cf. Figura 5.3) estaremos em condições de apresentar de que forma a Abordagem da Neurofilosofia Não-Redutiva, apesar de ser metodologicamente diferente tanto da Filosofia como da Neurociência pode, ainda, ser relevante e prolífera para ambas sem, no entanto, reduzir-se a uma das disciplinas.

Primeiramente, é necessário fazer notar que caracterizar o processo iterativo de cruzamento entre o conceptual e o empírico como uma mera “mistura brusca” é um simplismo que não pode ser permitido. De facto, a metodologia desta abordagem é deveras complexa: o conceito de “mistura” implicaria que o elemento empírico e o elemento conceptual possam ser cruzados bastando colocar-se arbitrariamente em conjunção. Ora, esta forma de interpretar o cruzamento iterativo descrito anteriormente (cf. Figura 5.3) está longe do que um defensor desta abordagem defende.

Pelo contrário, defende-se que o cruzamento acontece num sistema articulado de acoplamento ou ligação dos dois elementos que constituem a hipótese neurofilosófica: o empírico com o conceptual. Ora, é impossível, nesta estratégia, compilar o lado empírico com pressupostos ontológicos e epistémicos antagónicos com o esquema conceptual utilizado: caso haja uma inconsistência a esse nível deveremos, como detalhado anteriormente (cf. V.1.3.), proceder à correção quer do nível empírico, quer do nível conceptual, até que o cruzamento iterativo permita confluir numa hipótese neurofilosófica plausível e, assim, numa nova tese sobre determinado tema.

Ademais, considerar a ligação entre os dois níveis como uma “mistura brusca” implicaria que não haveria qualquer distinção entre os mundos natural e lógico, nem se consideraria o seu intercâmbio ou justaposição. Todavia, a ligação na metodologia da Abordagem da Neurofilosofia Não-Redutiva considera essas diferenças entre o natural e o lógico (cf. princípio da assimetria), assim como a dança constante entre os dois mundos (cf. princípio da bidirecionalidade).

Finalmente, a ligação ou cruzamento, nesta estratégia metodológica, não pode, justamente, ser considerada como arbitrária dado que há regras e princípios metodológicos que devem ser seguidos passo a passo de modo a que o cruzamento aconteça de forma sistematizada e rigorosa. Como afirma Northoff:

“Arbitrário” implica que não haja regras e estratégias que sirvam como orientação para a geração do “vínculo”. No entanto, existem vários princípios, isto é, os princípios da metodologia transdisciplinar que estabelecem estratégias concretas para “ligação” entre hipóteses empíricas, isto é, condições naturais e pressupostos teóricos e as condições lógicas. Tal reflete-se melhor no “princípio da circularidade transdisciplinar”, que define etapas específicas para a sua “ligação”. Adequadamente, a “ligação” entre hipóteses empíricas e suposições teóricas pode ser considerada como “sistemática” em vez de “arbitrária” (Northoff, 2004: 123).

A objeção da irrelevância metodológica aponta, assim, para que o defensor da Abordagem da Neurofilosofia Não-Redutiva esclareça com mais detalhe de que forma acontecerá a estratégia de cruzar iterativamente os dois níveis em análise. O próprio representante máximo desta abordagem, Georg Northoff, indica que as “estratégias metodológicas para a ligação entre os dados da primeira [filosofia] e terceira [neurociência] pessoas sobre estados mentais e neuronais precisam de ser mais refinados e desenvolvidos no futuro” (Northoff & Heinzl, 2006: 6).

Assim, deve ser afirmado que se espera que a investigação no futuro demonstre com maior precisão de que forma a metodologia neurofilosófica não-reduziva (i) se distingue tanto de uma mera conjunção entre (a) a metodologia filosófica e (b) a metodologia neurocientífica; (ii) se distingue de abordagens que utilizam a mesma intuição de respeitar tanto (a) quanto (b) tentando cruzá-las, como a Abordagem da Neurofenomenologia apresentada anteriormente e (iii) apresenta um progresso real capaz de causar impacto tanto em (a) como em (b).

Mas, mais importante que estas questões futuras, importa perceber com exatidão que metafilosofia temos nesta Abordagem da Neurofilosofia Não-Reduziva, algo que consideraremos na próxima objeção.

V.2.3. Objeção da “Metafilosofia Imprecisa”

Uma objeção final pode ser levantada a esta abordagem: a sua metafilosofia poderá ser problemática ou, pelo menos, imprecisa.

Fizemos notar na secção sobre o “princípio da circularidade transdisciplinar” (cf. V.1.2.3) que Northoff faz uso do conceito de “circularidade disciplinada” para corroborar a sua estratégia metodológica. Ora, esse conceito foi avançado por Francisco Varela (cf. Varela, 1998: 342) que, como já detalhámos no anterior capítulo, é um defensor da Abordagem da Neurofenomenologia. Assim sendo, é necessário compreender de forma explícita com que tipo de metafilosofia estamos a lidar com esta Abordagem da Neurofilosofia Não-Reduziva.

Mas também fizemos notar na secção dedicada à introdução da estratégia metodológica desta abordagem que Northoff segue Quine na sua rejeição da distinção entre o *a priori* e o *a posteriori* e, por consequência, do trabalho filosófico e do trabalho neurocientífico. Nas palavras de Quine, citado por Northoff:

Vejo a filosofia não como uma proposta propedêutica *a priori* ou uma disciplina-base da ciência, mas em continuidade com a ciência. Vejo a filosofia e a ciência como estando no mesmo barco – um barco que, para reverter à figura de Neurath, como tantas vezes faço, só podemos reconstruir no mar enquanto estivermos a flutuar nele [...]. Todas as descobertas científicas, todas as conjeturas científicas actualmente plausíveis são, portanto, a meu ver, tão bem-vindas para uso em filosofia quanto em outras disciplinas (Quine, 1969: 126).

Como fizemos observar anteriormente, a Filosofia de Quine era a base dos ideais defendidos pela abordagem reducionista, que defendia a redução – e em alguns casos a eliminação – gradual do trabalho filosófico pelo trabalho neurocientífico.

A posição de Quine vai dar origem ao naturalismo e às suas diversas variantes (cf. Capítulo 1, I.2.): o naturalismo metafilosófico fraco e o naturalismo metafilosófico forte. É necessário, assim, compreender claramente três questões relevantes: (1) de que forma esta abordagem se distingue da Abordagem do Reduccionismo em termos metafilosóficos; (2) de que forma esta abordagem se distingue da Abordagem da Neurofenomenologia em termos metafilosóficos; e (3) em que posição metafilosófica poderemos encaixar esta Abordagem da Neurofilosofia Não-Redutiva.

Em relação à Abordagem do Isolacionismo e à metafilosofia aí pressuposta, o defensor desta metodologia rejeita explicitamente a tese da exclusividade do trabalho filosófico, assim como a dicotomia metafilosófica extrema do total isolamento entre o trabalho filosófico e o trabalho neurocientífico: um não afeta o outro e vice-versa, havendo uma total rejeição de qualquer intercâmbio. Ora, o defensor da Abordagem da Neurofilosofia Não-Redutiva, defendendo uma posição de continuidade em relação à relação entre a Filosofia e a Ciência, rejeitará por consequência qualquer intuição metafilosófica que não apoia uma possível interação entre os dois empreendimentos. Mas será que a rejeição do isolacionismo e a defesa do naturalismo de Quine torna imediatamente a Abordagem da Neurofilosofia Não-Redutiva numa abordagem reducionista como a abordagem reducionista analisada no capítulo terceiro?

De uma perspectiva mais geral, se utilizarmos uma definição do conceito de “reduccionismo” alargado e não-técnico, um defensor da abordagem isolacionista poderá de facto acusar a posição de Northoff de “reducionista” dado que se assume a importância tanto do elemento empírico como do elemento conceptual para abordar os problemas da Filosofia da Mente. Mas o sentido relevante de uma acusação do tipo reducionista é dado pela abordagem reducionista.

Portanto, é necessário compreender (1): de que forma se distinguem as duas abordagens da neurofilosofia (a redutiva em oposição à não-redutiva)? Vejamos as principais diferenças com uma contraposição detalhada:

- Naturalismo metodológico: a abordagem reducionista faz uso de um tipo de naturalismo metodológico particular, o “naturalismo de substituição”, onde se defende a tese de que os métodos lógico-conceptuais da Filosofia devem ser tendencialmente substituídos pelos métodos da Ciência, dado que se assume que os primeiros são absolutamente desprovidos

de utilidade epistêmica, sendo vistos como ferramentas arcaicas incapazes de causar qualquer progresso na área de estudo; pelo contrário, a Abordagem da Neurofilosofia Não-Redutiva, apesar de também fazer uso do naturalismo metodológico, faz uso de uma versão essencialmente dissemelhante do naturalismo de substituição: em vez deste, um neurofilósofo não-redutivo aplica um “naturalismo de cooperação” que defende uma interação concreta e específica (cf. I.1.2.) entre os métodos da Filosofia e os métodos da Ciência; antes, considera que ambos os métodos são importantes e relevantes estando num contínuo que não defende a redução de uma disciplina à outra; assim, percebemos perfeitamente a agenda anti-reducionista da Abordagem da Neurofilosofia Não-Redutiva e por que razão a acusação de “reducionista”, no seu sentido técnico, não pode ser aplicada à mesma;

- Em consequência de aplicar um naturalismo de substituição, o mental é substituído pelo neuronal na Abordagem do Reducionismo: todos os estados mentais devem ser reduzidos a estados cerebrais e, em alguns casos, devem ser eliminados, dando origem a uma metodologia “reduzida ao cérebro” (*brain-reductive*); no sentido oposto, a Abordagem da Neurofilosofia Não-Redutiva, fazendo uso de um naturalismo de cooperação, não reduz o mental ao cerebral/neuronal: antes, o cérebro é considerado central e a base da investigação, mas não se reduz somente a este; é necessário considerar a interação do cérebro com o corpo e com o ambiente: apenas tendo presente esta tríade é que poderemos dar conta da natureza e complexidade do mental; assim, o cérebro e a sua atividade neuronal são metodologicamente relevantes – ao contrário da Abordagem do Isolacionismo, e no mesmo sentido que a Abordagem do Reducionismo – mas esta relevância é tomada como uma condição somente necessária do mental, em oposição a uma condição suficiente – como na abordagem reducionista (cf. Northoff, 2014a: 100-102).

Todavia, dado que acabámos de afirmar que a abordagem de Northoff se distingue das restantes por considerar o corpo e o ambiente como relevantes na equação do mental, de que forma é que a mesma se distingue da Abordagem da Neurofenomenologia, uma metodologia que tem, na sua base, precisamente a crítica anti-cartesiana à obsessão do cerebral pela Neurociência tradicional e a importância do corpo na investigação?

Analisemos, assim, a questão levantada em (2): de que forma se distingue a Abordagem da Neurofilosofia Não-Redutiva da Abordagem da Neurofenomenologia?

De facto, esta objeção tem bastante poder dado que Northoff, em variadas ocasiões, se afirma como um investigador que tenta colocar em prática uma investigação “neurofenomenal”, como afirma no prefácio do seu *Unlocking The Brain – Consciousness*:

O meu objetivo é desenvolver hipóteses neurofenomenais específicas que mostrem como as características intrínsecas do cérebro, ou seja, a sua atividade no estado de repouso e o código neuronal predis põem as características intrínsecas da consciência e as suas características fenomenais de uma maneira necessária e inevitável (Northoff, 2014d: xii).

Ademais, reparámos durante a introdução a esta abordagem a preocupação do autor em conectar os vários níveis – cerebral ou neuronal, corporal, ambiental – na sua explicação do fenómeno mental, algo que não podemos negar que aparece com a tradição fenomenológica e atualizada pela Neurofenomenologia, além de citar, em várias ocasiões, Francisco Varela, o fundador da Abordagem da Neurofenomenologia.

Assim, é necessário compreender explicitamente de que forma as duas abordagens em (2) diferem. Northoff avança com uma distinção subtil, mas fundamental, para compreender a diferença entre as duas abordagens: a distinção entre “hipóteses neurofenomenais” (Hn) e “hipóteses neurofenomenológicas” (Hnf).

A investigação da Abordagem de Neurofilosofia Não-Redutiva avançará, então, com (Hn): estamos perante uma abordagem que se foca nas características intrínsecas do cérebro, codificação da atividade neuronal e na estrutura espaçotemporal dessa atividade intrínseca em relação ao corpo e ao ambiente, defendendo uma ligação direta entre o nível fenomenal ou mental e o nível neuronal ou cerebral (cf. Northoff, 2014d: lxxi).

Ora, é aqui que encontramos uma diferença importante entre (Hn) e (Hnf): o foco nas características intrínsecas do cérebro distingue (Hn) de (Hnf) dado que em (Hnf) o foco principal dá-se nas funções sensoriomotoras – corpo – e na forma como estas estão relacionadas com a consciência. Todavia, para Northoff, as (Hn) são mais básicas e fundamentais que (Hnf). De forma a ficar claro: é o foco particular nas estruturas sensoriomotoras, em contraposição à relevância da atividade intrínseca do cérebro que difere entre as duas abordagens, exatamente porque, para Northoff, a segunda pressupõe ou antecede as primeiras.

Nas palavras do neurofilósofo:

[Embora] os termos “neurofenomenal” e “neurofenomonológico” pareçam estar muito próximos, as duas abordagens devem, no entanto, ser distinguidas uma da outra. Ambas fazem uso de pressupostos teórico-conceptuais diferentes assim como de abordagens empíricas diferentes ao cérebro [...] Empiricamente, a abordagem neurofenomenológica reivindica que as funções sensoriomotor são centrais na ligação do cérebro ao meio ambiente e, assim, constituírem a subjetividade e a consciência [...]. No sentido oposto, a minha abordagem neurofenomenal não considera nenhuma função específica, seja afetiva, sensoriomotora ou cognitiva como condição necessária e predisponente da consciência [...] Tal acontece porque a minha abordagem neurofenomenal concentra-se naqueles processos e mecanismos neuronais que precedem a constituição e diferenciação dessas diferentes funções. Metaforicamente falando, a abordagem neurofenomenal concentra-se no próprio terreno, a atividade intrínseca do cérebro sobre as quais colunas diferentes, isto é, as diferentes funções, são assentes. Em contrapartida, a abordagem neurofenomenológica concentra-se numa dessas colunas, as funções sensoriomotoras (Northoff, 2014d: 538).

Fica, assim, claro de que forma a Abordagem da Neurofilosofia Não-Redutiva difere da Abordagem da Neurofenomenologia: a sua consideração-base de que a função sensoriomotora é a base da consciência subjetiva implica uma concepção corpórea da mesma. Ora, embora o papel relevante do corpo não seja negativo em (Hn), este não assume o papel de relevância ou prioridade ontológica sobre o cérebro ou o ambiente: em (Hn), os estímulos sensoriomotores têm também de passar pelas características intrínsecas do cérebro, sendo, como as restantes funções – cognitivas ou afetivas – estímulos pós-fenomenais em vez de pré-fenomenais. Assim, há uma incorporação da relevância do corpóreo para o mental, mas não há redução do segundo ao primeiro, dado que o corpóreo é influenciado pelas condições pré-sensoriais da atividade intrínseca do cérebro (cf. Northoff, 2014d: 538).

Ademais, há também uma diferença que vale a pena apontar: a Abordagem da Neurofenomenologia é uma abordagem que, metodologicamente, tem na sua base uma estratégia que começa pela mente e a caracterização dos estados mentais “como aparecem à consciência” (cf. Capítulo 4). Pelo contrário, a Abordagem da Neurofilosofia Não-Redutiva rejeita essa estratégia, como vimos: o primeiro passo é uma estratégia baseada no cérebro e na sua atividade intrínseca que estão subjacentes à mente e aos estados mentais. Na conceptualização de Northoff, procuraremos inicialmente compreender os elementos pré-fenomenais da consciência que serão fundamentais para caracterizar a mesma sem focar exclusivamente na própria consciência para, só depois, conseguirmos caracterizar os elementos pós-fenomenais da consciência.

Finalmente, vimos também no exemplo das emoções (cf. V.1.4.3.) de que forma as duas abordagens diferem substancialmente na linha de raciocínio agora apresentada: ambas enaltecem a importância do corpo, mas a Abordagem da Neurofilosofia Não-Redutiva equipara o cérebro e o ambiente ao mesmo nível de prioridade ontológica, enquanto a Abordagem da Neurofenomenologia prioriza o corpo em relação tanto ao cérebro quanto ao ambiente.

Surge, todavia, uma dificuldade interessante para a estratégia de Northoff: ao dar relevância ao ambiente, e ao transformar o problema mente-corpo num problema mundo-corpo-cérebro (cf. Northoff, 2018: 269-314), o neurofilósofo alemão resolve um problema, mas cria um segundo: qual é a caracterização ontológica deste mundo? Seria necessário apresentar uma defesa ontológica da estrutura da realidade cuja ontologia estivesse enquadrada tanto na ontologia do corpo como na do mental.

Northoff chega a avançar com a defesa de um Realismo Estrutural Moderado (cf. Northoff, 2018: 242-255), mas tal poderá ser insuficiente: há variados problemas com essa ontologia relacional entre estruturas mais básicas. Por exemplo: Esfeld e Lam notam que, pressupondo uma ontologia deste tipo, pode surgir a objeção de que é impossível individualizar os objetos particulares do mundo (cf. Esfeld & Lam, 2008; 2010).

É também necessário compreender de que forma o tempo e o espaço – o que Northoff chama de “tempo interno” para a atividade temporal neuronal e o “tempo externo” para a atividade temporal do mundo – podem interagir de forma coerente. Assim sendo, será necessário demonstrar, futuramente e numa investigação subsequente focada no carácter ontológico da Abordagem da Neurofilosofia Não-Redutiva, de que forma o mundo pode ser definido para acomodar o problema do mental.

Faltará, então, responder a (3) e compreender claramente que diferenças metafisólicas existem perante as abordagens analisadas até este momento. Atenemos à seguinte esquematização que resume a investigação metodológica já realizada:

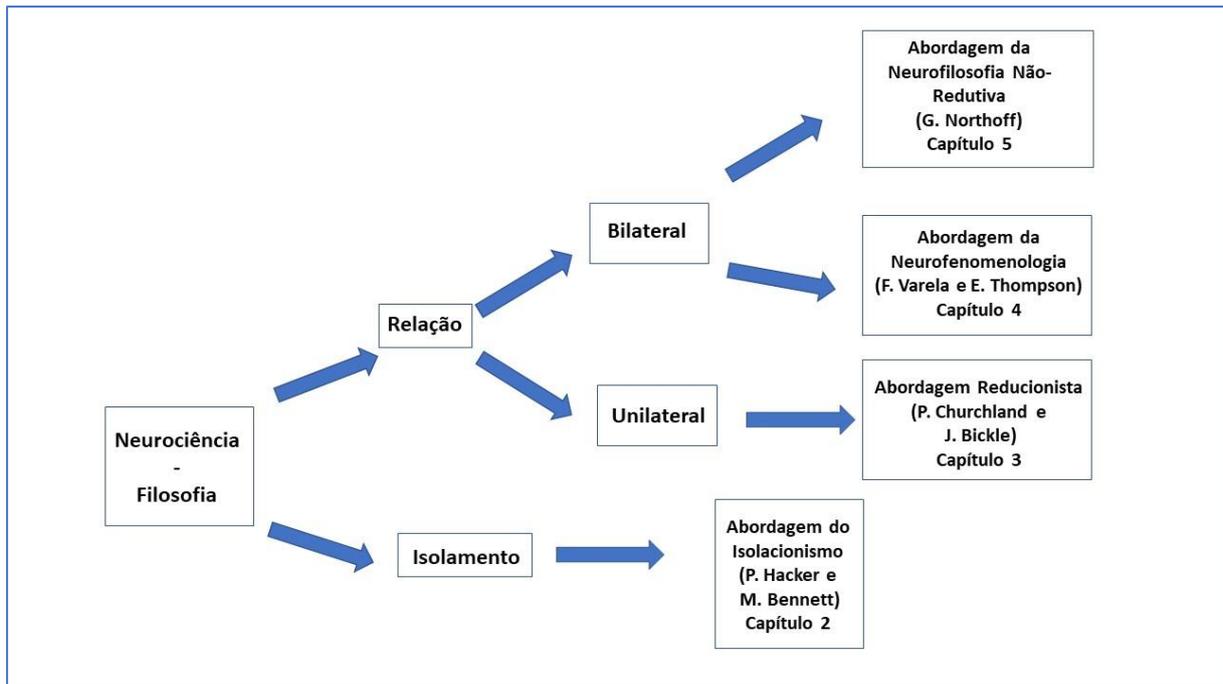


Figura 5.4.: As várias posições que procuram responder ao problema da relação (ou não) entre a investigação filosófica e a investigação neurocientífica.

Esta esquematização apresenta claramente de que forma as diferentes abordagens se distinguem ou se cruzam entre si. O primeiro grande critério de separação entre as várias abordagens é se aceitam ou recusam uma relação entre as duas disciplinas.

Como se analisou no segundo capítulo, a Abordagem do Isolacionismo recusa qualquer relação entre as duas disciplinas dado que são metafilosoficamente distintas uma da outra, sendo a única abordagem que defende explicitamente essa separação. Todavia, se recusarmos o isolamento disciplinar, poderemos defender que há, pelo contrário, uma relação entre as duas disciplinas.

Dentro dessa relação, levanta-se o segundo critério: a mesma pode ser unilateral, como defendido pela Abordagem do Reducionismo, que procura unilateralmente reduzir o trabalho filosófico ao trabalho neurocientífico. Ou poderá ser bilateral, isto é, poderá defender-se que a relação entre as duas disciplinas deve ser fundamentada numa interação constante, como advogam a Abordagem da Neurofenomenologia e a Abordagem da Neurofilosofia Não-Redutiva.

Mas como é que poderemos diferenciar metafilosófica e metodologicamente as diversas abordagens apresentadas na Figura 5.4.? A seguinte Figura 5.5. ajudará a clarificar visualmente esses dois critérios relevantes:

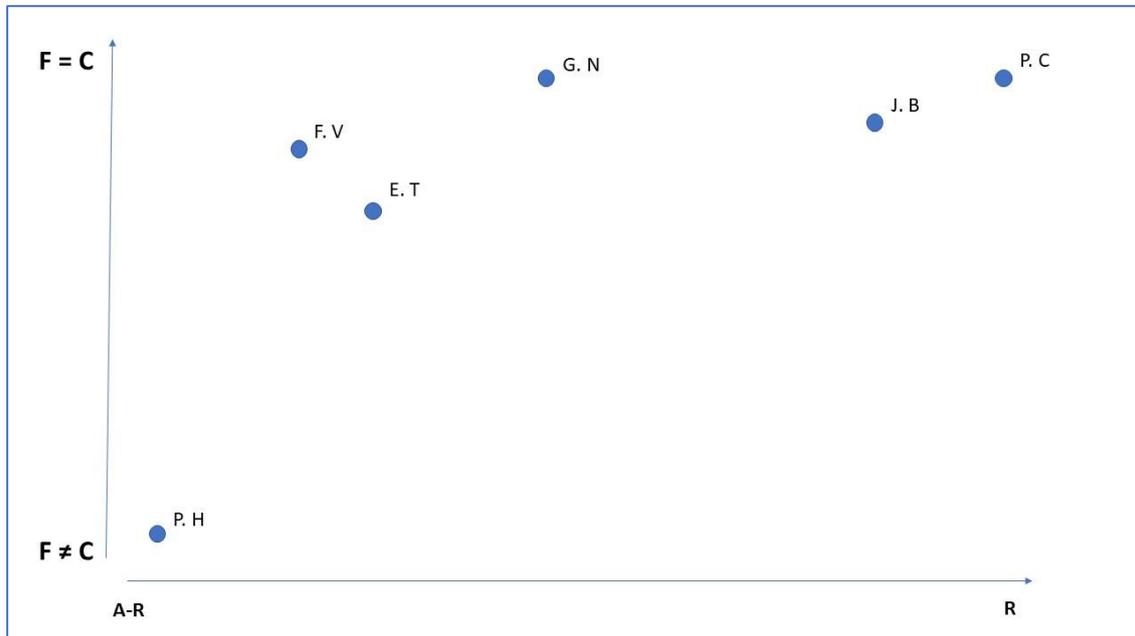


Figura 5.5.: As posições metafilosóficas das abordagens analisadas. No eixo vertical, poderemos observar dois extremos metafilosóficos entre a posição que defende que a Filosofia (F) e Ciência (C) são dois empreendimentos metafilosoficamente semelhantes e a posição que defende que (F) e (C) são metafilosoficamente dissemelhantes. No eixo horizontal, podemos observar dois extremos metodológicos entre a posição que defende que (F) deve ser reduzida a (C) – a posição reducionista (R) – e a posição que defende que (F) e (C) possam ser intercaladas – a posição anti-reducionista (A-R). Indicado com pontos no gráfico encontrar-se-ão os vários autores analisados nesta tese doutoral e a sua posição em relação aos dois eixos: (P.H) = Peter Hacker; (F.V) = Francisco Varela; (E.T) = Evan Thompson; (G.N) = Georg Northoff; (J.B) = John Bickle e (P.C) = Patrícia Churchland.

Percebemos, com a Figura 5.5., de que forma as variadas abordagens se distinguem, tanto em termos metodológicos como em termos metafilosóficos, umas das outras. Como se afirmou no início da investigação, as variadas abordagens metodológicas fazem uso de diferentes concepções metafilosóficas entre a tese (i) “filosofia-como-ciência” e (ii) filosofia-diferente-da-ciência” (cf. 1.2.) para lidar com os diversos problemas que procuram resolver.

Mas a grande novidade “filosófica” desta análise foi compreender de que forma os pressupostos metafilosóficos de cada autor influenciam a resposta dos mesmos ao problema metodológico das

relações entre a Filosofia e a Neurociência. A habitual “desatenção” desses mesmos pressupostos metafilosóficos – algo que já fizemos notar ser comum nos diversos investigadores das variadas áreas – poderá realmente ser um dos obstáculos ao próprio progresso dos complexos problemas da Filosofia da Mente em geral, e da Neurofilosofia em particular.

Finalmente, iremos, no próximo capítulo, analisar de que forma cada estratégia metodológica – a Abordagem do Isolacionismo, a Abordagem do Reduccionismo, a Abordagem da Neurofenomenologia e a Abordagem da Neurofilosofia Não-Redutiva – definem dois conceitos centrais: o conceito filosófico de *qualia* e o conceito neurocientífico de informação. Encerraremos com dois capítulos que deixarão de se focar no aspeto metodológico da relação entre a Filosofia e a Neurociência: dado que a última abordagem analisada nos parece ser a mais promissora, devido ao seu equilíbrio metodológico, por um lado, mas também pela aceitação de que apenas coligando a Filosofia e a Neurociência – sem ignorar uma área em detrimento da outra – se poderá alcançar algum progresso, mesmo que esse progresso seja diminuto ou controverso.

Esta conclusão advém da relevância dos vários argumentos avançados pelas abordagens anteriores: o rigor conceptual é necessário (cf. Capítulo 2); o rigor empírico é necessário (cf. Capítulo 3); uma abordagem corpórea é necessária (cf. Capítulo 4); mas apenas o conjunto equilibrado destas condições necessárias poderá oferecer uma mais-valia metodológica – e uma condição suficiente – para investigar os variados fenómenos mentais. Portanto, a Abordagem da Neurofilosofia Não-Redutiva é a consequência de uma dialética entre as várias teses e antíteses, sendo a síntese metodológica exatamente a proposta desta abordagem.

Assim, terminaremos a nossa investigação com a aplicação de Abordagem da Neurofilosofia Não-Redutiva a dois temas actuais: a natureza da percepção e a validade da teoria do processamento preditivo para o estudo da consciência, compreendendo de que forma as intuições metodológicas desta abordagem poderão auxiliar a investigação desses fenómenos em questão.

V.3. Conclusão:

Neste capítulo examinámos com detalhe os pressupostos metodológicos e metafilosóficos da Abordagem da Neurofilosofia Não-Redutiva. Primeiramente, demos conta da relevância da cooperação intermetodológica que esta abordagem propõe entre a investigação filosófica e a investigação neurocientífica. De seguida, detalhámos os princípios metodológicos que fundam a estratégia desta abordagem, nomeadamente o princípio da assimetria, o princípio da bidirecionalidade e o princípio da circularidade transdisciplinar. Avançámos com a análise da crítica às abordagens neurocientíficas que se reduzem ao cérebro ao invés de se basearem neste, optando por uma defesa de uma Neurociência da Primeira Pessoa. Por fim, observámos a diversidade explicativa desta abordagem, exemplificando os seus princípios pela análise da Abordagem da Neurofilosofia Não-Redutiva ao Eu, às emoções e a relação do Eu com fatores ambientais (e.g. a Covid-19).

Na segunda parte, investigámos as principais objeções a este projeto metodológico: a ideia de que esta abordagem poderá estar a cometer um erro de categoria ao dar tanta importância ao cérebro e à sua atividade neuronal; a ideia de que estamos perante uma metafilosofia irrelevante dado que, defendendo uma excepcionalidade do trabalho neurofilosófico, poderá ficar em causa a relação ou implicação para o trabalho filosófico e o trabalho neurocientífico convencionais; finalmente, analisámos a objeção da metafilosofia imprecisa, defendendo que esta abordagem poderá não se apresentar verdadeiramente como uma alternativa às abordagens anteriores dada a sua ambiguidade metodológica.

VI. QUALIA E INFORMAÇÃO EM NEUROCIÊNCIA E FILOSOFIA

Depois de termos analisado quatro abordagens que procuram responder ao problema metodológico entre a Filosofia e a Neurociência, iremos analisar de que forma cada uma dessas abordagens define um conceito central da Filosofia – *qualia* – e um conceito central da Neurociência – informação. Como iremos notar, a forma como cada metodologia define o seu próprio modo de investigar os problemas da mente irá impactar na definição desses dois conceitos: as metodologias não são, assim, neutras, contribuindo, diferentemente, para uma delimitação significativa dos conceitos.

Mas antes de avançarmos com a nossa tarefa, façamos uma breve introdução histórica aos dois conceitos.

VI.1. Conceito de *Qualia*

Qualia é a palavra latina para o plural de *quale* para descrever a propriedade subjetiva ou qualitativa da experiência. Na literatura filosófica, esta experiência particular costuma ser descrita como a experiência de “como é ser um morcego” – em que o ponto de vista do morcego pode ser substituído por qualquer outro. Convencionalmente, os *qualia* são tidos como a qualidade da experiência que pode ser relevada ou identificada através da introspeção. A existência dos *qualia* é, muitas vezes, usada contra as teorias reducionistas e materialistas como um obstáculo intransponível dos métodos da terceira pessoa.

Tradicionalmente, o conceito foi utilizado por C. I. Lewis (1929)⁵⁶ no âmbito da discussão sobre a teoria dos dados dos sentidos (*sense-data theory*), que identificava os *qualia* com esses mesmos dados dos sentidos (*sense datum*) (cf. Capítulo 7), afirmando que

[...] existem aspetos qualitativos reconhecíveis do dado [*the given*], que podem ser repetidos em experiências distintas e, portanto, são uma espécie de universais. Chamo a esses *qualia*. Mas, embora esses *qualia* sejam universais, no sentido de serem reconhecidos de uma para outra experiência, devem ser distinguidos das propriedades dos objetos. A confusão desses

⁵⁶ Embora a utilização do termo não tenha sido original. O singular do conceito, *quale*, é usado durante o período medieval em latim. Segundo o Dicionário de Inglês Oxford, o termo é utilizado pela primeira vez na língua inglesa em 1834, por Abraham Tucker (cf. Tucker, 1834: 462). Um excelente estudo sobre a origem do termo pode ser encontrado em Keeley, 2009: 71-89.

dois sentidos é característica de muitas concepções históricas, bem como das teorias actuais sobre a essência dos objetos. O *quale* é diretamente intuído, dado e não é objecto de nenhum erro possível, porque é puramente subjetivo (Lewis, 1929: 121).

Contudo, no sentido mais contemporâneo, o conceito tem sido alargado para referir, então, as propriedades da experiência: os objetos da percepção externa (verídica e ilusória), as sensações corporais (dor, fome, sede), as emoções (medo, ansiedade, felicidade), entre outros aspetos da vida mental, são comumente associados a possuírem uma propriedade qualitativa respetiva, um *quale* correspondente.

O conceito de *qualia* é tido como central em Filosofia da Mente e um dos maiores obstáculos à Neurociência contemporânea. Todavia, de facto, podemos encontrar-nos na situação de Agostinho de Hipona em relação ao tempo: se ninguém perguntar o que os *qualia* realmente são, todos saberão dar uma resposta; contudo, quando começamos a tentar perceber ao detalhe o que são os *qualia*, veremos que estamos perante um conceito no mínimo ambíguo e no máximo obscuro. Nas palavras de Tim Crane:

[...] parece-me que não existe um consenso claro sobre como o termo *qualia* deve ser entendido e, nessa medida, o problema contemporâneo da consciência não é bem colocado. A dificuldade aqui pode ser vividamente trazida à tona considerando-se o facto de que parece haver uma disputa real sobre a existência dos *qualia* (...) (Crane, 2000: 2).

Assim, podemos definir o uso do conceito de *qualia* de quatro formas principais:

- (1) *Qualia* podem ser identificados com o carácter fenomenal da experiência. O carácter fenomenal de perceber uma flor vermelha é diferente do carácter fenomenal de perceber uma flor azul: embora sejam duas experiências perceptiva de flores, o carácter subjetivo e único de cada experiência gera um carácter particular e irrepetível. Ora, se se focar no carácter fenomenal dessa experiência através da introspeção irá tornar consciente determinadas qualidades – são essas qualidades a que, neste sentido de (1), podemos chamar de *qualia* (cf. Peirce, 1982: 223);
- (2) *Qualia* podem ser identificados, como já fizemos notar, com o seu sentido original: como sendo uma propriedade dos dados dos sentidos. A apreciação de uma obra de arte pode ajudar neste segundo sentido: veja-se, por exemplo, a pintura *La Gioconda*. Por um lado, um espectador pode apreciar o conteúdo temático da obra mas pode, também, apreciar as

relações formais da obra – o jogo de cores e sombras, os equilíbrios e desequilíbrios espaciais, etc. Estar consciente de uma experiência visual funciona como uma imagem interna, não-física, da pintura. A visualização da mesma criou uma representação mental que, através da introspeção, releva tanto o conteúdo da obra como as suas relações formais – mas estas são tidas como características não-representacionais da obra. Assim, neste sentido (2), *qualia* são intrínsecos, características conscientemente acessíveis e não-representacionais dos dados dos sentidos que são responsáveis pelo seu carácter fenomenal (cf. Lewis, 1929);

- (3) *Qualia* podem também ser identificados como uma propriedade intrínseca não-representacional sem ter de estar conectada aos dados dos sentidos (isto é, à teoria de *sense-data*). Neste sentido, independentemente destas propriedades poderem ser físicas ou não-físicas, são consideradas intrínsecas no sentido de (2): são propriedades irreduzíveis, acessíveis conscientemente e não-representacionais responsáveis pelo carácter fenomenal. Este sentido mais abrangente (não reduzido apenas aos dados dos sentidos) é o mais comum dentro da Filosofia da Mente contemporânea (cf. Nagel, 1974; Peacocke, 1983; Block, 1990);
- (4) Finalmente, *qualia* podem ser definidos como (3) mas em conjunção com a ideia de que são propriedades inefáveis da experiência. Em versões ainda mais restritas de (4), além de inefáveis, os *qualia*, neste sentido, são “dados” aos sujeitos de forma incorrigível, isto é, é impossível para os sujeitos estarem errados sobre um determinado *quale* que experienciam (cf. Dennett, 1991).

Ora, ligado ao conceito de *qualia*, está também uma diversidade de argumentos e experiências mentais que procuram mostrar, em geral, que os *qualia* são reais e que são, usualmente, irreduzíveis ao físico.

Uma das mais famosas experiências mentais a favor dos *qualia* (cf. (5)) foi avançada por Frank Jackson que definia *qualia* como “[...] determinadas características das sensações corporais [...] e de determinadas experiências perceptivas [...] que nenhuma quantidade de informação puramente física inclui” (Jackson, 1982: 127). Neste artigo, o filósofo australiano relata a história de Mary, uma brilhante neurocientista que conhece todos os factos empíricos associados tanto à neuroanatomia e neurofisiologia quanto ao sistema neurovisual da perceção das cores. Contudo, Mary encontra-se numa situação estranha: desde a sua nascença apenas viveu num quarto fechado. Pior que isso, apenas viveu num

mundo monocromático: nunca viu cores, e todas as investigações e estudos que fez foram a preto e branco.

Um dia, Mary é libertada do quarto onde se encontrava prisioneira e vê, pela primeira vez, a cor vermelha de uma flor. Mary fica maravilhada com a cor que está a ver pela primeira vez e percebe que, apesar de saber tudo sobre a Neurociência da percepção e a Física das cores, havia mais uma informação que lhe faltava – o dado de ver, de facto, a cor vermelha, a experiência consciente da cor vermelha ou, usando a terminologia técnica exposta acima, o *quale* associado à experiência de ver vermelho.

Esta experiência mental possui duas finalidades principais: provar que os *qualia* existem, isto se aceitarmos – e é aqui que está o problema central do argumento – que Mary aprende realmente algo particular que não conhecia antes, ou seja, o conhecimento do *quale* que corresponde à experiência de ver vermelho, devendo ser admitido que os *qualia* são propriedades reais, uma vez que existe uma clara diferença entre uma pessoa que tem um acesso direto a um *quale* particular e outra que não o tem; e, por fim, refutar a propriedade física da mente: dado que Mary estava na posse de todas as informações físicas e empíricas possíveis sobre as cores e, mesmo assim, aprendeu algo novo, então o fisicalismo é implausível, dado que há verdades sobre as cores que não são físicas.

Mas de que forma se segue, logicamente, a tese de que os *qualia* refutam o fisicalismo – a tese metafísica de que todos os objetos que existem terem de ser descritos pelas leis da natureza? Vejamos uma possível formulação do argumento (5) que torna clara essa refutação:

- (5.1) Antes de ser libertada, Mary possuía *todas* as informações sobre a experiência de ver cores;
- (5.2) Após ser libertada, Mary aprende algo novo – o *quale* de ver cores – sobre a experiência de ver cor;
- (5.3) Logo, antes de ser libertada, Mary não possuía, afinal *todas* as informações sobre a experiência de ver cor, mesmo que tivesse na posse de *todas* as informações físicas sobre a experiência de ver cor;
- (5.4) Logo, existem verdades sobre a experiência de ver cores que não são físicas;
- (5.5) Portanto, o fisicalismo é falso.

Ora, esta formulação torna clara de que forma a ideia de que Mary aprende algo novo – a premissa (5.2) será uma das mais disputadas – refuta claramente o fisicalismo dado que esta teoria metafísica deve explicar todos os fenómenos – literalmente – e bastará encontrar um contraexemplo que se possa admitir

existir mas fora do âmbito de explicação naturalista. O argumento da Mary procura ser esse contraexemplo, provando que os *qualia* são essa exceção.

De seguida, iremos analisar de que forma cada uma das abordagens examinadas anteriormente se relacionam com a existência (ou não) dos *qualia* e de que forma respondem ao Argumento da Mary.

VI.1.1. A Abordagem Isolacionista e Qualia

Peter Hacker e Bennett Maxwell começam por fazer notar de que forma o sentido de *qualia* apresentado em (3) tem sido mais influente entre os filósofos: os *qualia* são experienciados com um carácter qualitativo único. Por exemplo, Searle defende que

Qualquer estado consciente tem um certo “sabor” [*feel*] qualitativo, e podemos ver isto se examinarmos exemplos. A experiência de saborear cerveja é muito diferente da de ouvir a Nona Sinfonia de Beethoven, e ambas têm um carácter qualitativo diferente de cheirar uma rosa ou ver um pôr-do-sol. Estes exemplos ilustram as diferentes características qualitativas das experiências conscientes (Searle, 2000: 560).

No entanto, e de forma curiosa, parte dos investigadores em Neurociência decidiram seguir os filósofos e defender uma concepção de *qualia* muito parecida. Vejamos três exemplos neste sentido:

- António Damásio: “Os *qualia* são as simples qualidades sensitivas que se podem encontrar no azul do céu ou na tonalidade do som produzido por um violoncelo, e os componentes fundamentais das imagens [*sc.* causadas pela percepção] na metáfora do filme são desta forma feitos de *qualia*” (Damásio, 1999: 9);
- Gerald Edelman e Giulio Tononi: “[*sc.* o problema dos *qualia*] é provavelmente o mais desencorajador problema da consciência [...] cada experiência consciente diferenciada representa um *quale* diferente, seja ele essencialmente uma sensação, uma imagem, um pensamento, ou mesmo uma disposição” (Edelman & Tononi, 2000: 157);
- Ian Glynn: “Embora os *qualia* se associem mais obviamente às sensações e percepções, também podem ser encontrados noutros estados mentais, tais como crenças, desejos, esperanças e medos, durante episódios conscientes destes estados” (Glynn, 1999: 392).

Quer seja para filósofos quer para neurocientistas, os *qualia* são qualitativos no sentido em que “para qualquer estado consciente [...] há alguma coisa que sabe [*to feel*] qualitativamente como estar nesse estado [*sc.* consciente]” (Searle, 1997: xiv) ou, como já indicámos, este sentido ficou popularizado nas palavras de “há uma coisa que é ser como *x*”, há algo que é ser como esse algo (cf. Nagel, 1979: 166), e esse algo corresponde ao carácter ou qualidade da experiência subjetiva a que chamamos de *qualia*.

Deste sentido, parece haver duas formas relevantes de ligar esta forma de definir *qualia* com a consciência:

(Q.1) Uma criatura é consciente ou tem experiência consciente se e só se houver uma coisa que é como para a criatura ser a criatura que é;

(Q.2) Uma experiência é uma experiência consciente se e só se houver uma coisa que é como para o sujeito da experiência ter essa experiência (cf. Bennett & Hacker, 2003: 272).

Ora, para os defensores da abordagem isolacionista, toda esta abordagem e descrição dos *qualia* e do seu carácter qualitativo único está impregnada numa teia de confusões conceptuais que necessita de um esclarecimento lógico e não demonstrações empíricas.

Será que ter uma experiência sabe (*to feel*) sempre a algo que é ter essa experiência e que, portanto, cada experiência é unicamente diferente? Para os autores, esta questão pode ter variados sentidos. Para um indivíduo normal, ter a experiência de ver uma mesa não tem nada de especialmente característico. No máximo, um indivíduo que era cego e que voltou a conseguir ver poderá achar a experiência de ver a mesma “maravilhosa”. Mas tal não indica, para os autores, que a experiência tenha um “sabor” qualitativamente diferente de ver, por exemplo, uma cadeira. A existir uma diferença, essa diferença não tem que ver com “a que sabe ver uma mesa”, mas, antes, que experiências diferentes podem diferir no sentido em que o objeto da perceção difere (cf. Bennett & Hacker, 2003: 273).

Deste modo, perceber usando diferentes modalidades – visual, auditiva, sensorial, etc. – não implica qualquer qualidade (e.g. agradável, apazível, desagradável) e muito menos essa qualidade iria diferir consoante um objeto diferente. Portanto, para os autores, a questão “A que é que sabe *x*” não tem sentido lógico, sendo apenas uma confusão conceptual que necessita de ser eliminada da investigação atual.

Seguindo a citação de Searle, num certo sentido, o filósofo de Berkeley está certo de que há algo de diferente entre “o sabor de uma cerveja” e “ouvir uma sinfonia”. Todavia, essa diferença não tem que

ver com a natureza qualitativa da experiência, mas, antes, somente com a diferença de modalidade a que essas experiências apontam.

O “sabor” qualitativo de beber uma cerveja ou de ouvir uma sinfonia pode até ser igual: ambas as experiências podem, por exemplo, ser aprazíveis da mesma forma. E é plausível afirmar que o “sabor” qualitativo seja diferente: que beber uma cerveja seja desagradável e ver o pôr-do-sol seja agradável. Mas desta afirmação truista não se segue que *todas* as experiências tenham um carácter qualitativo diferente e, por isso, único (cf. Bennett & Hacker, 2003: 250).

Para os isolacionistas, a leitura falaciosa dos defensores dos *qualia* sobre a natureza dos mesmos, quer do lado da Filosofia, quer da Neurociência, assenta em várias confusões conceptuais/lógicas que devem ser evitadas:

(6.1) As experiências podem ser sujeitos lógicos de predicados de atitude (como algo ser agradável, ou maravilhoso, ou aborrecido); estes atributos é que devem ser considerados “o carácter qualitativo da experiência”, e não as experiências em si; os *qualia*, assim, não são “o carácter qualitativo da experiência” mas, antes, são as qualidades dos objetos da experiência;

(6.2) As experiências em geral podem ser sujeitos lógicos possíveis (e não reais) de predicados de atitude (positivos ou negativos); mas tal não significa que todas as experiências são sujeitos reais ou factuais desses predicados de atitude;

(6.3) Nem as diferenças entre experiências, nem o seu carácter único se reduzem ao modo a que sabe ter essas experiências, dado que duas experiências podem “saber” [*to feel*] o mesmo sendo diferentes (e.g. cheirar uma rosa em oposição a cheirar uma tulipa: ambas as experiências podem possuir o carácter qualitativo de ser agradáveis, mas isso não as distingue);

(6.4) As experiências são individualizadas ou especificadas não pelo carácter qualitativo da experiência, mas, antes, pelos objetos dessa experiência (o que faz diferente a experiência em cheirar uma rosa versus uma tulipa não é o carácter da experiência em si, mas os objetos dessa experiência, isto é, a rosa e a tulipa, serem distintos) (cf. Bennett & Hacker, 2003: 251-252).

Assim, para os autores, o único interesse epistémico em pensar a pergunta de Nagel “Como é [a que sabe] ser *x*” é a tese de que apenas organismos conscientes podem ser sujeitos de experiências e que podem ser predicados por atributos de qualidade (e.g. agradável, desagradável, interessante, aborrecido, etc.). Nas palavras dos autores:

Mas para achar a experiência E como sendo F (agradável, desagradável, etc.) não é uma característica das experiências enquanto tais, ou das experiências tidas enquanto conscientes, ou de experiências das quais estamos conscientes, e muito menos da consciência transitiva em geral (que envolve muito mais do que experiências) (Hacker & Bennett, 2003: 78).

No fundo, a plausibilidade da questão de Nagel aparece devido a duas formas muito semelhantes, mas profundamente distintas, de interpretar a pergunta “como é ser um morcego”:

(V.1) Como é que é para um x ser um y ,

(V.2) Como é que é para um x ser um x ,

Ora, em (V.1), encontramos uma formulação totalmente plausível: no seu âmago, estamos perante um princípio de comparação perfeitamente aceitável, dado que faz todo o sentido perguntar, e.g., como é que é para um professor universitário ser pai. Todavia, o problema é que as interrogações em relação à natureza dos *qualia* nada têm que ver com (V.1), mas com (V.2): existe uma “reiteração ilegítima” de x que, logicamente, deve ser rejeitada, dado que o primeiro x em (V.2) (e.g. como é *para um morcego ser um morcego*) não desempenha a função legítima que deveria desempenhar nessa construção frásica.

Um outro aspeto problemático dos *qualia* é a tese apresentada em (4) de que estes são, por definição, indescritíveis e inefáveis: só cada sujeito pode experienciar o seu *quale*, mas esse *quale* não poderá ser experienciado por outra pessoa. Ora, num certo sentido, é razoável afirmar-se que, por exemplo, o sujeito A e B percecionam a mesma rosa vermelha: ambos, neste caso, vêem o mesmo objeto dado que ambos podem relatar o que estão a ver e compreender, por essa descrição, que estão de facto a ver o mesmo objeto (cf. Hacker, 2002: 168).

Todavia, a noção de *qualia* procura afirmar, nesta situação, o seguinte: “É certo que vemos a mesma flor mas A não pode ter a visão de B !”. Mas, para os isolacionistas, isto é uma mera confusão linguística, dado que (i) nem A pode ter a visão de B , (ii) nem B pode ter a visão de B , porque a visão não é algo que se “tenha”.

Wittgenstein avançou com um argumento interessante para mostrar de que forma a conceção de experiência consciente como privada e inefável assenta em várias confusões que não necessitam de esclarecimentos empíricos – como alguns neurocientistas e filósofos fazem notar – mas, antes, de esclarecimentos conceptuais.

O argumento – ou a série de argumentos que tipicamente é chamado de “argumento da linguagem privada” – tem sido interpretado de variadas formas. Podemos ler Wittgenstein como defendendo a tese que não há estados fenomenais que são privados, mesmo ontologicamente; ou que os estados fenomenais não são suficientes para determinar a dor (ou outra sensação) – uma interpretação que parece ser suportada pela primeira parte do parágrafo do “Besouro na caixa”: “[...] seria bem possível que todos tivessem algo diferente na sua caixa” (1953: §293, 100e); ou ainda pode ser lido como uma defesa de que os estados fenomenais não são sequer necessários para termos dores (ou sensações) – como parece ser afirmado na segunda afirmação do parágrafo “Besouro na caixa”: “A caixa pode até estar vazia” (cf. Hacker, 1992: 68).

Independentemente da interpretação mais favorecida, o foco da investigação de Wittgenstein é a ideia sedutora de que a consciência é essencialmente privada no sentido já indicado de que apenas o sujeito que possui esse estado pode saber se se encontra nesse estado ou não – relembremos que a ideia de Nagel é justamente mostrar que há algo de inacessível para um sujeito de fora compreender na sua totalidade o que é experienciar o mundo como um morcego.

Como indicado, uma das consequências desta ideia é a tese forte de que existe uma assimetria entre a primeira e a terceira pessoas: enquanto o uso da terceira pessoa é baseado na observação, a primeira pessoa não é baseada na observação. Como vimos, a natureza desta tese obriga ou afirma que a consciência garante não só a privacidade, mas também a imediatez incorrigível e a infabilidade do que é apresentado à consciência.

Assim, seja qual for a interpretação que façamos do argumento da linguagem privada, o objetivo de Wittgenstein é claro: mostrar que, seguindo o seu exemplo, existe uma falsa representação do conceito de dor – esta tese poderá ser generalizada a qualquer experiência subjetiva ou *qualia*. Em primeiro lugar, é falso que, no sentido ordinário ou comum da expressão, uma pessoa não possa saber que outra pessoa está com dores: pelo contrário, muitas vezes sabemos facilmente que outra pessoa está de facto a sentir dor, bastando olhar para ela ou ouvir os seus pedidos de socorro ou auxílio.

Em segundo lugar, a ideia de que, pela natureza da dor, um sujeito sabe com toda a certeza que ele está com dores é, para os autores, uma apreensão incorreta do sentido da palavra “dor”. É verdade que não parece fazer sentido para o sujeito da dor, mas não para outra pessoa, expressar dúvidas sobre se está a sofrer. Mas tal não é uma questão de o sujeito ser dotado, pela natureza da dor, de uma autoridade epistémica excepcionalmente privilegiada sobre o estado atual da sua consciência, dado que as regras gramaticais que regem a palavra “dor” não são mais deriváveis da natureza da dor do que qualquer outra regra gramatical derivável da realidade (cf. Hacker, 1992: 369).

Assim, quando afirmamos “estar com dores” não estamos a oferecer uma descrição intimista, mas somente uma expressão cujo sentido é baseado no jogo de linguagem do conceito de “dor” partilhado pela maior parte dos seres humanos. Esta conclusão segue-se ao facto de que, se assumirmos que as dores são realmente privadas, várias consequências implausíveis se seguem: apenas cada pessoa poderia saber a que chama de “dor” e não o que os outros chamam de dor. Mas, se tal fosse o caso, aquilo que cada sujeito nomeia de dor passaria a ser irrelevante para o uso da palavra “dor” na nossa linguagem comum: o uso da palavra numa linguagem pública não pode ser constrangido por considerações sobre o que é essencialmente privado para cada sujeito. Assim, saber se alguém compreende a palavra “dor” e a usa corretamente é determinado pelo critério acessivelmente público, e não privado, do uso desse conceito (cf. Bennett & Hacker, 2003: 297).

Portanto, se a dor for privada, a definição ostensiva de dor também o seria, e faltaria aprender ou conhecer a regra para o uso da palavra “dor”. Assim, chegamos à conclusão de que a alegada privacidade essencial dos estados da consciência implica não só que uma pessoa não pode transmitir o seu “conhecimento” privado para outros, mas também que não o pode expressar para ele próprio, uma conclusão contraditória que, segundo os autores, deve ser rejeitada (cf. Hacker, 1992: 368).

Mas será que duas pessoas podem realmente ter “a mesma dor”? Alguém afirmar “O sujeito *A* não pode ter a minha dor (do sujeito *B*)” parece querer significar que (i) dois sujeitos não podem ter a mesma dor (i.e. uma dor numericamente idêntica) mas apenas dores semelhantes (i.e. uma dor qualitativamente idêntica): seguir-se-ia, de (i) e (ii), que ninguém pode saber verdadeiramente se um sujeito está com dores dado que o *quale* de dor é privado e inacessível do ponto de vista da outra pessoa.

Este cenário descreve mais uma confusão conceptual inaceitável para os isolacionistas: a distinção entre identidade numérica e identidade qualitativa que se aplica a substâncias não tem qualquer aplicação a sensações ou experiências. O facto de o sujeito *A* ter uma dor na “sua” cabeça (i.e. cabeça de *A*) e o sujeito *B* ter uma dor de cabeça na “sua” cabeça (i.e. cabeça de *B*) aponta para uma diferença de localização espacial (i.e. cabeça de *A* não é a cabeça de *B*) que, pela aplicação da lei de Leibniz, implicaria uma diferença numérica. Ora, para os autores, tal é somente uma confusão teórica obscura: dois sujeitos terem a mesma dor no mesmo lugar (e.g. no lado esquerdo frontal da cabeça) é justamente o que significa possuir uma dor específica numa zona particular do corpo (cf. Hacker, 1992: 370).

Portanto, os critérios de identidade de uma dor incluem características fenomenais (de como se sente, e.g. desagradável), de intensidade (do grau de dor e.g. elevada) e de localização (da parte do corpo e.g. cabeça). Se essas características corresponderem de igual forma no sujeito *A* e no sujeito *B*, então é perfeitamente razoável afirmar que os dois sujeitos possuem a mesma dor. Como afirmou

Wittgenstein, a expressão “estou com dores” é somente uma afirmação que reitera o facto de alguém estar com dores, ou um “contrassenso” de alguns filósofos – e, poderíamos afirmar igualmente, de alguns neurocientistas – dado que se está a contrariar a lógica gramatical da aplicabilidade do conceito de dor (cf. Hacker, 2018: 131-137).

Restará fazer notar que estes esclarecimentos conceptuais a que chamamos de “argumentos da linguagem privada” de Wittgenstein não são, para os autores, uma forma de behaviorismo em que reduzimos a expressão “o sujeito *x* tem dor porque ele se comporta como se estivesse a ter dor”: antes, o próprio Wittgenstein sempre frisou a distinção entre “dor”, por um lado, e “comportamento de dor”, por outro. Ademais, nunca rejeitou totalmente a ideia de que o mental é uma ficção, mas antes que uma certa imagem filosófica do mental é uma ficção gramatical (cf. Hacker, 2018: 123-126).

Para concluir esta análise, faltará apontar de que forma esta abordagem interpreta e responde à experiência mental da Mary. Muito simplesmente, aos ombros de Wittgenstein, nega-se que Mary tenha aprendido algo de novo dado que, de facto, não poderia aprender por ter em falta o próprio conceito de “vermelho” no seu esquema conceptual. Mesmo que Mary afirme: “a flor é vermelha”, ela não está a expressar uma experiência ou um *quale* específico do vermelho, mas, antes, está apenas a expressar uma crença metalinguística que associa a flor à cor vermelha dado que ela conhece todas as informações factuais sobre as cores e as flores e sabe que ambas andam em conjunção uma com a outra. Contudo, o importante é compreender que, neste sentido, Mary não possui de facto o conhecimento experiencial do *quale* vermelho (cf. Wittgenstein, 1977: I, 13).

Para que fosse possível Mary ter um *quale* da cor vermelha, seria necessário que ela soubesse identificar essa cor: dado que ela nunca esteve perante a cor vermelha, e dado que ninguém da comunidade linguística a que pertence lhe ensinou ou apontou para uma instanciação do vermelho, quando confrontada pela primeira vez com essa cor quando é libertada, a neurocientista brilhante não estará em condições de conciliar o conceito de cor com a experiência que está a ter.

Assim, estamos novamente perante o processo de elucidação lógico-conceptual em relação aos *qualia* que é a marca da abordagem isolacionista. De forma interessante, compreendemos perfeitamente de que forma a metafilosofia desta abordagem influencia a sua definição do conceito de *qualia*: é necessária uma cautela conceptual a interpretar este conceito que deve ser visto com mais ceticismo na investigação contemporânea sobre a mente consciente.

VI.1.2. A Abordagem Reducionista e Qualia

De seguida, iremos analisar de que forma a Abordagem do Reducionismo e os seus proponentes lidam com o conceito de *qualia*. Curiosamente, e como analisámos nos capítulos 2 e 3, estaremos perante duas metafisologias fundamentalmente opostas entre si. Ora, devemos concluir que, por esse motivo, a abordagem reducionista considera os *qualia* com mais relevância do que a abordagem anterior? Como iremos notar, apesar de possuírem metafisologias diferentes, a visão cética ou, pelo menos, crítica, das duas abordagens em relação aos *qualia* é semelhante, embora por motivos diferentes.

Como notámos anteriormente, a existência dos *qualia* – que parece ser comprovada por vários argumentos – levanta uma série de constrangimentos às teorias fisicalistas ou materialistas do mental dado que parece implicar que há propriedades fundamentais da consciência que não se podem reduzir ao puramente físico. Ora, uma forma de evitar o problema que os *qualia* causam à ciência contemporânea é simplesmente afirmar a sua não-existência.

Essa será a estratégia da abordagem reducionista: como vimos no terceiro capítulo, esta abordagem defende que, à medida que a Neurociência se irá desenvolver em conjunção com outras disciplinas científicas como a Psicologia Evolutiva, os conceitos tradicionais da Filosofia ligados ao mental – percepção, sensação, emoção, *qualia* – serão eliminados do esquema linguístico e conceptual dessas mesmas disciplinas dado que perderão a sua utilidade explicativa e preditiva.

Esta posição em relação aos *qualia* é apelidada na literatura científica como “eliminativismo” exatamente porque, na sua essência, defende que os *qualia* devem ser eliminados das nossas teorias sobre o mental. Como fizemos observar, há variados exemplos da história da Ciência que parecem demonstrar que a estratégia eliminativista de conceitos incoerentes faz sentido e é, até, útil para fazer progredir determinado objeto que estaria a causar algum embaraço teórico (cf. Churchland, 1985: 43-45).

Relembrando o exemplo da teoria ptolemaica: uma melhor teoria não tem de explicar o papel da esfera celestial – antes, bastará negar a sua existência e perceber que estávamos perante um pseudoproblema. De forma analógica, Churchland argumenta que o mesmo deve acontecer com o conceito de *qualia*: o avanço científico do estudo do mental irá demonstrar que o problema levantado pelos *qualia* é apenas um pseudoproblema (cf. Churchland, 1986: 292-293).⁵⁷

⁵⁷ Como vimos anteriormente, o sucesso das mudanças conceptuais em determinadas áreas científica (e.g. Física) não obriga a que o mesmo aconteça em outras (e.g. Neurociência). Além disso, o sucesso passado não garante o sucesso futuro.

De forma interessante, é curiosa a forma como a abordagem reducionista lida com o problema levantado pelos *qualia*: apresenta uma refutação meramente *a priori* do fenômeno, dado que não pode oferecer nada mais substantivo que tal refutação. Tal ponto já foi frisado no “desafio de ‘demonstra ou silencia-te’” (cf. III.1.6), uma objeção que, no seu âmago, desafiava os reducionistas a apresentar uma redução real ou a eliminação de qualquer aspecto da mente consciente. Como fizemos notar, Bickle apresentou algumas reduções do fenômeno da memória – mas deixa de fora qualquer *quale* associado – e Churchland apenas afirma que uma teoria futura haverá de livrar-se desses conceitos mentalistas.

Por causa disso, a única fonte das ideias acerca dos *qualia* apresentada pela abordagem reducionista tem um teor meramente negativo: aparece nas críticas aos argumentos a favor da existência dos *qualia*, como seja o Argumento da Mary (ou do Conhecimento). Apresentemos, assim, as várias reflexões dos reducionistas sobre o argumento e, por consequência, sobre a natureza dos *qualia*.

A análise do argumento por Paul Churchland começa com a ideia de que os esquemas conceptuais ao longo da história da Ciência que têm mostrado ser profundamente contingentes são substituídos por mais e melhores esquemas ao longo do progresso e desenvolvimento de áreas científicas particulares, afirmando que devemos “[...] considerar [...] a possibilidade de aprender a descrever, conceber e apreender introspectivamente os meandros da nossa vida interior, dentro da estrutura conceptual da neurociência amadurecida” (Churchland, 1985: 16).

Churchland reconstrói o argumento da Mary da seguinte forma:

(7.1) Mary sabe tudo o que há para saber sobre os estados cerebrais e as suas propriedades;

(7.2) Não é o caso que Mary saiba tudo o que há para saber sobre sensações e as suas propriedades;

Logo, pela lei de Leibniz,

(7.3) As sensações e as suas propriedades são diferentes dos estados cerebrais e das suas propriedades.

Ora, como já analisámos anteriormente, o argumento de Jackson sofre de uma falácia intensional: a instanciação que o filósofo australiano faz do conhecimento é transparente e extensional em (7.1) e (7.2).

Mas para Churchland, tal não é de todo o caso, por duas razões principais:

- (i) Mesmo que a referência ao “conhecimento” seja transparente tanto em (7.1) como em (7.2), o significado não é totalmente unívoco em ambas as premissas e, se tal for o caso,

então o argumento perde a sua validade; segundo Churchland, a diferença é que em (7.1) a referência ao conhecimento é equivalente a um conjunto de proposições; já a referência ao conhecimento na premissa (7.2) está ligada à representação de propriedades anteriores à representação sublinguística da experiência (cf. Churchland, 1985: 23); se tal for o caso, então, fazendo uso da distinção de Frege dos modos de apresentação, a forma ou o modo como as coisas podem ser conhecidas é diferente da diversidade das próprias coisas;

- (ii) Jackson assume que mesmo um cenário de perfeição epistémica em relação ao conhecimento do cérebro iria ainda deixar os *qualia* de fora: “(...) mesmo um conhecimento utópico da neurociência deve deixar Mary irremediavelmente no escuro da natureza qualitativa subjetiva das sensações ainda não desfrutadas” (Churchland, 1985: 25); ora, para Churchland, a plausibilidade desta afirmação deriva não de um problema com a abordagem reducionista em geral mas, antes, do erro de Jackson em considerar o que significa “saber tudo o que há para saber sobre o cérebro físico”; portanto, Mary poderia muito bem, apesar de nunca ter experienciado o *quale* de determinada cor, imaginar o que seria experienciar esse *quale* dado que não há nenhuma razão *a priori* para descartar essa hipótese (cf. Churchland, 1985: 26).

Além desta distinção relacionada com os modos de apresentação, Churchland faz uso também da distinção de Russell entre “conhecimento por contacto” e “conhecimento acerca de”: entre as coisas de que temos representações e as coisas que apenas alcançamos por meio de frases denotativas. Assim, Mary aprendeu todo o conhecimento neurocientífico por descrição, e não por contacto; a neurocientista não aprendeu nada de novo – antes, apreendeu algo através de um novo modo de representação que já conhecia antes (cf. Churchland, 1985; Tye, 1986).

Além desta resposta mais filosófica, Churchland oferece também um contra-argumento pragmático ao argumento da Mary comparando-a a uma criança selvagem. Usualmente, uma criança selvagem cresce e vive afastada dos seres humanos em isolamento extremo. Se levarmos a sério o facto de Mary tenha vivido sempre fechada num quarto monocromático, e se atendermos ao conhecimento empírico resultante das investigações da Psicologia do Desenvolvimento em crianças selvagens, então perceberemos que, levando literalmente a experiência mental a sério, o cérebro e o sistema visual de Mary não teriam aprendido a habilidade de ver ou de conhecer a cor vermelha, dado que o cérebro necessita de aprender e desenvolver mecanismos específicos em variados níveis celulares – e.g. é

necessário formar padrões na secção V4 do córtex visual, e tal apenas se dá pela exposição da pessoa a variados comprimentos de onda de luz, principalmente no início do desenvolvimento cerebral.

Ora, no máximo, o que poderia acontecer pragmaticamente é que Mary iria apenas ter uma experiência de cor a preto e branco, e isso contraria a conclusão de que Mary aprenderia algo novo, nomeadamente o *quale* da cor vermelha. Embora esta refutação seja interessante, ela falha em considerar a própria natureza das experiências mentais: estas não procuram ser adequadas às considerações contingentes do mundo natural. Antes, procuram problematizar o carácter modal de certas teses assumidas.⁵⁸ Assim sendo, haverá outra forma de problematizar o argumento de Mary e a concepção de *qualia*?

Patrícia Churchland avançou igualmente com variadas abordagens críticas ao argumento de Mary e, por consequência, ao conceito de *qualia*. A filósofa começa por questionar a pressuposição de perfeição epistémica que a experiência mental assume:

Como posso avaliar o que Mary vai saber e perceber se ela sabe tudo o que há para saber sobre o cérebro? Tudo é muito, e significa, com toda a probabilidade, que Mary tem uma compreensão radicalmente diferente e mais profunda do cérebro do que qualquer coisa que mal se possa imaginar nos nossos voos mais extravagantes (Churchland, 1986: 332; cf. Dennett, 1991: 399-400).

Para a filósofa americana, não parece fazer sentido assumir que haja sequer um “problema difícil da consciência”: por que razão devemos supor que continuará a haver um mistério ininteligível? Supondo que a Neurociência consiga explicar todos os processos neurobiológicos do mental, parece ser uma pressuposição aleatória afirmar que a consciência ou os *qualia* ficarão de fora dessa explicação neurocientífica.

Churchland apelida a este tipo de argumentos de hipóteses “left-out”: são, no fundo, argumentos que já assumem à partida aquilo que procuram, supostamente, demonstrar de forma imparcial. São versões mais sofisticadas de petições de princípio não muito óbvias à primeira análise (cf. Churchland, 1997: 39). Não há qualquer impedimento lógico ou de princípio ou qualquer razão *a priori* para acreditar que a consciência esteja sempre fora da esfera científica: o que há somente é, neste momento particular

⁵⁸ As experiências mentais são usadas em diversas disciplinas para investigar diversos assuntos, como por exemplo na Economia (cf. Thoma, 2016: 129-146), na Filosofia Pré-Socrática (cf. Rescher, 1991: 31-42) ou na Ciência e Ética (cf. Shepard, 2008: 3-35). Duas obras fundamentais sobre a natureza e utilidade das experiências mentais são Sorensen (1992) e Stuart *et al.* (2018).

da história da humanidade, falta de conhecimento neurocientífico que explique exatamente de que forma o cérebro cria e dá origem à mente consciente.

A filósofa americana dá uma resposta muito clara ao argumento de Mary, levantando uma objeção analógica: “Se um aluno realmente compreender a natureza da gravidez ao aprender tudo o que há para conhecer sobre a natureza causal da gravidez, ninguém esperaria que a estudante engravidasse [sc. devido a isso]” (Churchland, 2005b: 289). Ora, com esta analogia, Churchland pretende pôr em causa a ideia de que o mero conhecimento proposicional que Mary aprendeu a preto e branco a impede de ter os mecanismos da percepção da cor desenvolvidos para conseguir apreender a cor da flor e, por consequência, para ter o *quale* da cor vermelha.

Para pôr em causa esta experiência mental e mostrar, igualmente, que o facto de uma experiência mental demonstrar que algo é concebível e depois possível nada nos ajuda a investigar empiricamente a mente consciente, Churchland apresenta uma outra experiência mental para dar a entender que tudo parece ser permitido neste tipo de exercício lógico que não toma em consideração o mundo natural:

“Deadbies” são coisas compostas de células com membranas, núcleos com DNA, as organelas usuais e assim por diante. “Deadbies” reproduzem-se, digerem, respiram, metabolizam, fabricam proteínas, crescem e assim por diante, assim como os restantes organismos na Terra. Ao contrário de nós, no entanto, os “deadbies” não estão realmente vivos. Esta é uma possibilidade lógica. Então a vida é explicitamente independente da biologia; (...) imagine um planeta onde a velocidade das moléculas num gás aumenta, mas a sua temperatura não. Isso diz-nos que a temperatura é explicitamente independente da energia cinética molecular média? Certamente que não. O que isso nos informa sobre a relação real entre a energia cinética média da molécula e temperatura de um gás? Efectivamente nada (Churchland, 2002c: 178).

O que Churchland procura problematizar com esta experiência mental é a seguinte pressuposição anti-reducionista: por que razão havemos de concluir que, depois da ciência explicar um fenómeno particular, haverá algo de misterioso que fique de fora dessa imagem científica particular? Para a filósofa americana, a maior parte das experiências mentais tem, no seu fundamento, um *argumentum ad ignorantiam*: dado que não compreendemos totalmente um fenómeno *F*, e dado que a ciência não consegue explicar totalmente *F*, segue-se que:

(A.1) F nunca poderá ser explicado;

Ou

(A.2) Por muito que a ciência se desenvolva, nunca poderá descobrir algo relevante para aprofundar a compreensão de F ;

Ou

(A.3) F nunca poderá ser explicado em termos de propriedades S (cf. Churchland, 1997: 41).

Assim, para Churchland, do facto de atualmente desconhecermos algo sobre a natureza de F não implica que F seja misterioso, fora do âmbito científico ou inteligível eternamente. A filósofa reconhece que é absolutamente verdadeiro que, actualmente, a Neurociência tenha dificuldade em lidar com o fenómeno da mente consciente. Mas tal não se deve a qualquer propriedade ou caracterização ontológica da mesma, mas somente a um problema epistémico ligado ao conhecimento atual:

O mistério do problema não é um facto sobre o problema, não é uma característica metafísica do universo, é um facto epistemológico sobre nós. É sobre o conhecimento que temos na ciência atual, é sobre o que podemos e não podemos compreender [...]. Não é uma propriedade do próprio problema (Churchland, 1996b: 406).

Assim, a aposta da filósofa está no progresso científico para aclarar a escuridão da mente consciente, e não assumir uma posição de obscuridade em relação a um fenómeno que não se compreende totalmente na atualidade, mas que, tal como todos os fenómenos anteriormente não compreendidos pela ciência, poderá ser melhor entendido conforme o desenvolvimento da Neurociência e da tecnologia associado ao estudo do cérebro. Claro, a autora ignora, no entanto, o subconjunto dos insucessos da história da Ciência que são muito mais numerosos que o subconjunto dos casos de sucesso, podendo estar a cometer um viés cognitivo da confirmação que poderá colocar em causa a ambição pressuposta da abordagem reducionista.

Será que há alguma forma de problematizar a resposta dos Churchlands ao argumento da Mary? Torin Alter (2013) pensa que sim, dado que os defensores reducionistas assumem algumas teses sobre o argumento que podem ser implausíveis. É importante fazer notar que a inferência que Jackson faz da experiência de Mary para a refutação do fisicalismo é indireta, dado que a experiência apenas conclui o que Mary conhece ou não conhece sobre o mundo. Assim, há uma inferência direta daquilo que Mary conhece para a tese fisicalista sobre a natureza do mundo, isto é, entre (5.3) e (5.4), dado que o próprio

Jackson qualifica o fisicalismo como a doutrina que afirma que toda a informação verdadeira sobre o mundo é informação física.

Ora, Churchland ataca exactamente a base desta inferência direta que, segundo o filósofo, se baseia num princípio que é implausível, nomeadamente o princípio constitutivo que afirma que o fisicalismo é verdadeiro apenas se o conhecimento científico (físico) constituir – em oposição a descrever ou explicar – o conhecimento subjetivo, i.e., os *qualia* (cf. Torin, 2013: 61).

Todavia, tal ideia é implausível: o conhecimento fenomenal pode ser conhecimento não-proposicional de algum tipo – não requer que seja conhecimento proposicional *per se*. Nas palavras do reducionista:

Jackson espera que o conhecimento explícito-discursivo-científico constitua de alguma forma um conhecimento subjetivo das experiências visuais. Mas essa expectativa não é razoável à partida. Também seria de esperar que o seu conhecimento discursivo exaustivo da micro-organização do córtex motor de um golfista profissional constitua um “conhecimento” prático real da sua parte, de como acertar numa bola de golfe a 200 jardas no meio do *fairway*, mesmo se nunca tiver utilizado um taco de golfe na sua vida (Churchland, 2011: 30).

Embora esta forma de criticar o argumento de Mary possa levantar vários problemas (cf. Torin, 2013: 62-65), o mais relevante desta análise é compreender de que forma os *qualia* são analisados pela Abordagem do Reduccionismo: a sua natureza não é um obstáculo a uma futura teoria neurocientífica que poderá reduzir ou até eliminar esse conceito pelo seu correspondente neurocientífico. Ademais, será através do foco na investigação do cérebro e das suas funções que encontraremos a chave para desvendar o mistério da consciência.⁵⁹

Vejamos, de seguida, a Abordagem da Neurofenomenologia aos *qualia*.

⁵⁹Para quem defende uma abordagem baseada no conhecimento empírico e rigoroso, estamos perante uma tese bastante mística e até profundamente anticientífica neste ponto particular. Os *qualia* e a consciência deixam de ser um problema para esta abordagem porque o enquadramento metodológico deixará de postular essas mesmas entidades. Ora, tal é algo que terá de ser provado e não pode ser assumido à partida, como parece ser o caso. Estamos assim perante uma petição de princípio metodológica que devemos colocar em causa.

VI.1.3. A Abordagem da Neurofenomenologia e Qualia

No quarto capítulo, analisámos ao detalhe a Abordagem da Neurofenomenologia avançada por Francisco Varela e os seus defensores. Como observámos, Varela acreditava que nem o método da terceira pessoa da Neurociência, nem o método da primeira pessoa da fenomenologia seriam suficientes por si só para compreender a natureza única e subjetiva dos *qualia*.

Antes, era necessário prestar atenção ao fundamento da mente e da experiência consciente: a estrutura da experiência humana e o acesso que, através desta, temos do mundo e de nós mesmos. Assim, não é necessário procurar um “ingrediente extra” para resolver o problema difícil da consciência: o que é necessário é encontrar pontes que conectem os dois domínios.

Observámos que tal se daria a partir das estruturas oferecidas pela redução fenomenológica e pela atenção dada à experiência quando esta é posta em “suspensão”: apenas aplicando este método poderíamos estar perante a natureza real da experiência consciente. Mas este método permitia também ser intersubjetivamente acessível à análise externa e, igualmente importante, poderia servir de guia para a própria investigação neurocientífica da terceira pessoa (cf. Varela, 1996: 341).

Ora, a Abordagem da Neurofenomenologia tem uma perspetiva muito particular sobre os *qualia* dado que não aceita o problema que estes causam à partida. A versão clássica do problema difícil da consciência assume uma lacuna absoluta entre o físico e o mental, sendo que uma unidade conceptual entre os dois níveis, entre o cérebro e a consciência, parece ser uma mera miragem. Os *qualia* são tidos como propriedades fenomenais que não são redutíveis ou explicáveis a partir de qualquer análise funcional ou neurocientífica.

Dado que o corpo é visto, tradicionalmente, de forma secundária, este apenas pode funcionar como um mecanismo do processo consciente. Nesta formulação dos *qualia*, há apenas dois caminhos a seguir: (i) ou reduzir a consciência ou os *qualia* a mecanismos ou estados cerebrais; (ii) defender um dualismo de propriedades, assumindo que os *qualia* são irredutíveis ao físico.

Para um defensor da Abordagem da Neurofenomenologia, esta forma de pôr o problema é já problemática em si mesma: o problema dos *qualia* é posto porque a forma como o problema é posto é puramente cartesiana. Na mesma linha da Abordagem do Isolacionismo do segundo capítulo, a Abordagem da Neurofenomenologia identifica o mesmo erro em grande parte da investigação científica atual da consciência: a rejeição do dualismo cartesiano de substâncias mantendo, todavia, a estrutura conceptual subjacente a esse cartesianismo.

Todavia, para a Neurofenomenologia, tal é um erro que deve ser ultrapassado: o problema de ligar um conceito de consciência subjetiva a um conceito de corpo objetivo é um falso problema. Antes, é necessário compreender de que forma o “corpo vivo” influencia diretamente a experiência consciente e vice-versa.

Portanto, se a Abordagem do Reduccionismo aposta na ideia de que uma explicação neuronal ou cerebral dos *qualia* irá emergir da investigação futura em Neurociência ou irá simplesmente eliminar o seu conceito, a Abordagem da Neurofenomenologia critica essa previsão. Os *qualia* e a experiência consciente em geral não poderão ser explicados apenas tendo como referência o cérebro: antes, os processos mentais devem ser caracterizados em termos de processos dinâmicos que conectam um intercâmbio constante entre o mental, o corporal, o comportamental e o ambiental (cf. Noë, 2004, 2009).

Ora, estes processos não se reduzem simplesmente ao cérebro: a consciência “não está dentro da nossa cabeça”, como famosamente argumentou Alva Noë – há uma dança persistente entre as várias partes da equação que compõe uma explicação total do fenómeno do mental. Ademais, a própria consciência ou os *qualia* estão diretamente ligados a um organismo que age, isto é, que se comporta no mundo através de ações intencionais: este processo não é passivo, como assumem as abordagens mais neurocêtricas da abordagem reducionista, mas, antes, o corporal em interação com o ambiente é condição necessária, em conjunção com o cérebro, para explicar adequadamente a consciência.

Em relação aos *qualia* ligados às cores, a posição da Abordagem da Neurofenomenologia é clara:

A afirmação não é que a cor (tendo em conta o exemplo padrão), como experienciada exista puramente na estrutura do mundo (que devemos aprender a “enativar”, isto é, a comportarmo-nos adequadamente perante ela). Tampouco é que a cor exista puramente nas nossas mentes e não realmente no mundo. Em vez disso, [...] a cor existe na nossa interação com o mundo (Beaton, 2013: 310).

Deste modo, a posição desta abordagem aos *qualia* é absolutamente clara e crítica da abordagem reducionista:

A afirmação geral é, então, que os termos como “mente fenomenal” e “experiência subjetiva” (e, de facto, termos como “crença”, “desejo”, “conhecimento”, “emoção” – todos os termos mentais de nível pessoal) não são sobre dinâmica cerebral; pelo contrário, são sobre (...) a relação activa do sujeito com o mundo. Se mantivermos essa dimensão normativa em mente, então (...) poderemos encontrar os *qualia* (...) (Beaton, 2013: 310).

De que forma lida esta abordagem, em concreto, com o fenómeno da consciência subjetiva? Uma forma muito comum de abordar a investigação dos *qualia* é procurar aquilo a que chamamos de “correlatos neuronais da consciência” (CNC) (cf. Metzinger, 2000; Rees, Kreiman & Koch, 2002). Embora o termo possua o conceito de correlato, o objetivo é descobrir não meramente um correlato, mas que essa estrutura neuronal particular seja suficiente para dar conta dos conteúdos conscientes (cf. Kanwisher, 2001).

Adaptamos de David Chalmers a seguinte fórmula que formaliza os (CNC):

(CNC): x é um CNC (conteúdo) se x é um sistema representacional neuronal mínimo N , de modo que a representação de um conteúdo em N é suficiente, nas condições C , para representação desse conteúdo na experiência consciente E (cf. Chalmers, 2000: 31).

Ora, a partir da definição de (CNC), fica claro que estamos perante uma pressuposição de que o conteúdo da consciência x é equivalente ao conteúdo neuronal do sistema representacional N . Este pressuposto é apelidado por Noë e Thompson (2004) da “Doutrina de Conteúdo Correspondente” (dCC) que defende o seguinte:

(dCC): Para qualquer experiência consciente E , existe um sistema representacional N , tal que (i) N é o sistema representacional neuronal mínimo cuja ativação é suficiente para E ; e (ii) existe uma correspondência entre o conteúdo de E e o conteúdo de N (cf. Noë e Thompson, 2004: 4);

Desta formulação, segue-se que (dCC) procura isolar os (CNC) para formas específicas de experiências conscientes E 's que, embora inicialmente possa ser vista como neutra em termos causais – dado o conceito de mera correlação – o objetivo é determinar não somente a correlação, mas uma teoria neurofisiológica com poderes causais necessários e suficientes do fenómeno em questão.

Uma implicação relevante que se segue desta estratégia metodológica é assumir que, determinando-se os (CNC) de um E particular, poderemos generalizar para outras instâncias de E : por exemplo, descobrir os (CNC) da visão poderá ser generalizado a todas as outras modalidades (cf. Crick & Koch, 1998).

Contudo, para a Abordagem da Neurofenomenologia, há vários erros com a forma de interpretar os (CNC) e conduzir a investigação empírica segundo esta estratégia:

- (a) Embora a investigação empírica tenha descoberto alguns (CNC) de variados estados conscientes, não se provou qualquer correspondência entre o conteúdo de N e o conteúdo de E ;
- (b) É impossível que haja, sequer, essa correspondência entre os conteúdos de N e E dado que resultam de diferenças conceptuais e de categoria; tal estratégia poderá ser um mero erro de categoria (cf. Tononi & Edelman, 1998; Searle, 2004b);
- (c) A tese da “neutralidade causal” inicial é uma limitação que pode afetar a própria investigação acerca da natureza da experiência consciente dado que (i) cada E (em uma escala de fração de segundo) envolve a seleção transitória de um conjunto neuronal dinâmico em larga escala e (ii) que o comportamento global de (i) modula e limita a atividade neuronal local influenciando causalmente os processos mentais e a própria ação intencional (cf. Thompson, Lutz & Cosmelli, 2005: 47).

Assim, os proponentes da Abordagem da Neurofenomenologia vão recusar a estratégia dos (CNC) e vão defender uma outra abordagem a que Searle chama de “Modelo de Campo Unificado” (mCU). Neste modelo, um E particular não é constitutivo de um estado consciente agregado, mas da alteração do campo unificado consciente: “Experiências conscientes vêm em campos unificados. Para se ter uma experiência visual, um sujeito já deve estar consciente, e a experiência é uma modificação do campo” (Searle, 2000: 572)

Com este novo modelo (mCU), estamos perante um modelo que se foca na totalidade da atividade neuronal em larga escala em vez de se focar somente nos (CNC) de E particulares, além de se considerar que os processos neurofisiológicos são dinâmicos, influenciando-se uns aos outros e os próprios conteúdos da experiência consciente e, por fim, considera-se a atividade cerebral não somente em estados conscientes transitivos, mas também intransitivos (cf. Thompson, Lutz & Cosmelli, 2005: 48). Nas palavras dos autores:

A neurofenomenologia é consistente com o “modelo de campo unificado”. Como observa Searle, de acordo com a abordagem de campo unificado, “o que é necessário procurar é uma atividade maciça do cérebro capaz de produzir uma experiência consciente holística unificada” (p. 574). Uma das hipóteses de trabalho da neurofenomenologia é que essa atividade maciça é a integração em larga escala via sincronia de fase (ou formas mais

complexas de sincronia generalizada). Essa auto-organização global é metaestável (novos padrões surgem constantemente sem que o sistema se estabeleça em nenhum deles), rápida (ocorrendo em uma escala de tempo de 100 a 300 milissegundos) e constitui um pólo fundamental de integração para a seleção e modulação da atividade neuronal local (Thompson, Lutz & Cosmelli, 2005: 48)

Fica, portanto, clara a tese da Abordagem da Neurofenomenologia em relação aos *qualia* ou experiência consciente: estes poderão ser investigados desde que se tenha uma abordagem holística, unificada e corpórea da experiência consciente, não reduzindo a mesma a sistemas neuronais particularmente isolados e estáticos.

Para concluir, qual é, então, a posição desta abordagem em relação ao argumento da Mary? Em primeiro lugar, a Abordagem da Neurofenomenologia não aceita, como vimos, a lacuna que os *qualia* supostamente fazem nascer entre a perspectiva da primeira e da terceira pessoas: em vez de procurar “fechar” essa lacuna, esta abordagem procura uma “ponte” (cf. Capítulo 4).

Uma resposta direta ao argumento é avançada por Thompson (1992): a experiência mental, como descrita por Jackson, é impossível de ser considerada seriamente devido a falhas metodológicas ou descritivas. O facto de Mary ter sido confinada a um quarto monocromático não a impediria de não ver o *quale* de uma cor particular, dado que ela poderia ainda ver cores a sonhar, ou esfregando os seus olhos, ou ainda pela percepção de imagens posteriores a partir da percepção da luz (cf. Thompson, 1992: 322).

Para Thompson, a forma como os *qualia* são definidos por Jackson na experiência de Mary entra em colisão com as suas ideias em relação à mente consciente. Em primeiro lugar, se os *qualia* são não-relacionais e se são qualidades fenomenais sem qualquer intencionalidade para o mundo, então nada têm que ver com a nossa experiência consciente do dia-a-dia e, portanto, neste sentido, a Abordagem da Neurofenomenologia negaria a relevância dos *qualia*. Todavia, vejamos de que forma a metafilosofia desta abordagem influencia a própria definição de *qualia*: se os *qualia* não forem definidos como neste primeiro sentido, mas forem simplesmente identificados como as características qualitativas da experiência, então fará sentido usar este conceito, desde que seja entendido como um conceito corpóreo, relacional e ligado ao mundo pela intencionalidade da sua experiência (cf. Thompson, 1992: 344).

Thompson acaba por criticar intensamente a experiência mental de Mary, afirmando que

[...] a moral metodológica a ser traçada é que não devemos prosseguir as nossas investigações em filosofia da mente avançando argumentos *a priori* sobre o que a ciência poderá ou não dizer-nos sobre a nossa experiência e as experiências de outras criaturas sencientes. A única maneira de sabermos o que a ciência nos pode dizer é fazendo ciência. Fazer ciência não significa abandonar a filosofia, mas significa que as nossas explorações filosóficas não devem envolver experiências mentais ou de pensamento irrestritos e apelos sem argumentos a intuições sobre propriedades misteriosas. O conhecimento das características qualitativas da experiência consciente, científico ou não, nunca pode ser obtido a menos que comecemos com um retrato mais preciso e fiel do que aquele que encontramos no argumento do conhecimento de Jackson e em muitas discussões actuais sobre *qualia* (Thompson, 1992: 345).

E esta crítica faz sentido se enquadrarmos a mesma na metafilosofia particular – que, como vimos, se encontra equivocada – da Abordagem da Neurofenomenologia, principalmente o seu cunho naturalista de que o problema da consciência necessita da intervenção tanto da investigação científica como de investigação filosófica para ser solucionado.⁶⁰

Analisemos, por fim, a última abordagem investigada a que prestaremos atenção sistemática, percebendo de que forma os *qualia* se enquadram na estratégia multidisciplinar da Abordagem da Neurofilosofia Não-Redutiva.

VI.1.4. A Abordagem da Neurofilosofia Não-Redutiva e os *Qualia*

Como analisámos com detalhe no quinto capítulo, a Abordagem da Neurofilosofia Não-Redutiva avançada por Georg Northoff procura cruzar iterativamente a investigação filosófica com a investigação neurocientífica, propondo hipóteses de trabalho verdadeiramente neurofilosóficas.

Desta forma, esta abordagem fará o estudo mais preciso e focado nos *qualia* das quatro abordagens aqui analisadas. A estratégia passará por investigar os *qualia* – e a consciência em geral – quando esta não está presente: analisando os mecanismos neuronais em pacientes que se encontram

⁶⁰ Há, no entanto, inúmeros casos de sucesso em que experiências mentais ajudaram ou guiaram o progresso científico de determinados problemas. Alguns exemplos famosos incluem o de Lucrécio a atirar uma lança na borda do universo, o demónio de Maxwell, o elevador de Einstein, o gato de Schrödinger, o balde e a bala de canhão de Newton, o microscópio de Heisenberg, os corpos em queda de Galileu ou o Dilema do Prisioneiro (cf. Stuart et al., 2018: 1; Ylikoski, 2003: 35-59; Van Dyck, 2003: 1-21), pelo que Thompson precisaria de indicar uma refutação particular desta experiência mental, em vez de avançar com ceticismo generalizado a todas as experiências mentais.

em Estado Vegetativo (EV). Por que razão escolher pacientes em (EV) para estudar os *qualia*? Esses pacientes perdem a sua habilidade de ter qualquer experiência subjetiva pelo que, se investigarmos as mudanças e alterações neuronais, poderemos ter uma forma indireta de compreender que mecanismos estão envolvidos no cérebro que impliquem a manifestação dos *qualia*.

Surge, neste ponto, uma distinção relevante que é necessário explicitar: entre (8.1) os conteúdos da consciência, (8.2) o nível da consciência e, por fim, (8.3) a forma da consciência. Em relação a (8.1), estamos perante aquilo que se encontra dentro do domínio fenomenal da experiência subjetiva particularizada; em (8.2) estamos perante os níveis ou graus alterados da consciência, em que (8.1) pode ser tido como mais claro ou distinto dependendo do nível de consciência efetivo; por fim, em (8.3) estamos perante a forma da consciência, a sua estrutura ou organização espaço-temporal onde os estímulos externos e internos interagem, estando particularmente ligada, como veremos, à atividade intrínseca do cérebro ou ao estado de repouso do cérebro (cf. Northoff, 2014a: 415).

Um estudo feito pela equipa de Steven Laureys (cf. Vanhaudenhuyse *et al.* 2010) investigou a conectividade funcional (CF) no estado de repouso (ligada à Rede de Modo Padrão – *Default Mode Network*) fazendo uso da ressonância magnética funcional (fMRI). O estudo envolvia pacientes perfeitamente saudáveis e pacientes em estados problemáticos de consciência: um paciente em estado de encarceramento (*locked-in syndrome*); quatro pacientes em estado minimamente consciente; quatro pacientes em estado vegetativo e, finalmente, cinco pacientes em coma (inconscientes).

O primeiro passo do estudo foi perceber exatamente a conectividade funcional (CF) da (RMP) em geral – foi observada elevada (CF) nas regiões da:

- Linha média anterior (anterior córtex cingulado, córtex pré-frontal ventromedial, córtex cingulado anterior subgenual e dorsomedial);
- Linha média posterior (córtex cingulado posterior, precuneus, córtex retrosplenial);
- Medial temporal (hipocampo e para-hipocampo);
- Na junção temporoparietal bilateral.

Foi na linha média posterior que a (CF) se mostrou mais elevada em comparação com as restantes regiões (cf. Northoff, 2014a: 415).

O segundo passo foi dividir os pacientes pelo nível de consciência de cada um para analisar de que forma a (CF) iria variar segundo diferentes níveis de consciência (e.g. pacientes saudáveis, pacientes em estado vegetativo, etc.). Os resultados foram interessantes: por ordem decrescente, a (CF) estava

mais elevada na (RMP) (em conjunção com o tálamo e o tronco cerebral) nos (i) pacientes saudáveis; de seguida, (ii) os pacientes em estado de encarceramento tinham níveis de (CF) muito similares a (i); pelo contrário, (iii) os pacientes em estado minimamente consciente exibiram níveis de (CF) mais baixos; (iv) os pacientes em estado vegetativo apresentaram níveis de (CF) ainda mais baixos que (iii) e, finalmente, (iv) os pacientes em coma não apresentavam uma (CF) relevante (cf. Vanhaudenhuyse *et al.*, 2010: 167).

Estes resultados apontam para uma interdependência direta entre o grau de (CF) na (RMP) e o grau de consciência. Nos pacientes em (EV), ficou clara a diminuição da (CF) em todo o cérebro, incluindo no tronco cerebral, no tálamo e particularmente nas estruturas da linha média como o núcleo da (RMP) (cf. Boly *et al.*, 2009). Estes resultados estão em linha com outros estudos semelhantes que apontam para os mesmos resultados em pacientes em (EV) (cf. Silva *et al.*, 2010; Huang *et al.*, 2013; Laureys & Schiff, 2012).

Além disso, um paciente em morte cerebral não mostrou qualquer (CF) de longa distância, i.e., no cérebro em geral, tendo sido apenas identificadas correlações a nível local. Finalmente, a relevância do tálamo em pacientes em (EV) foi também confirmada pela recuperação de um dos pacientes em (EV) que, tendo sido estimulado eletricamente nessa região, recuperou a consciência (cf. Schiff, 2009, 2010).

Quer em termos espaciais, quer em termos temporais, a (CF) está mais localizada em regiões específicas do cérebro e está menos generalizada pelo mesmo, não se estendendo tanto a outras regiões cerebrais nos pacientes em (EV) comparando com os pacientes saudáveis. Ademais, oscilações de alta frequência também se encontram reduzidas em pacientes (EV), embora as frequências de baixa frequência se mantenham de algum modo conservadas (cf. Northoff, 2014d: 422). Nas palavras do neurofilósofo:

Metaforicamente falando, o nível de ruído temporal e espacial no estado de repouso significando interação repouso-repouso [*rest-rest interaction*] parece ser significativamente reduzida em EV. O estado de repouso é, metaforicamente falando, mais silencioso em EV. Há apenas uma criança solitária a brincar silenciosamente na sala do estado de repouso do cérebro em EV em comparação com o elevado barulho das 20 crianças a brincar na sala do estado de repouso do cérebro saudável (Northoff, 2014d: 422).

Ora, a partir destes dados (e outros), Northoff propõe um espectro biofísico-computacional do cérebro:

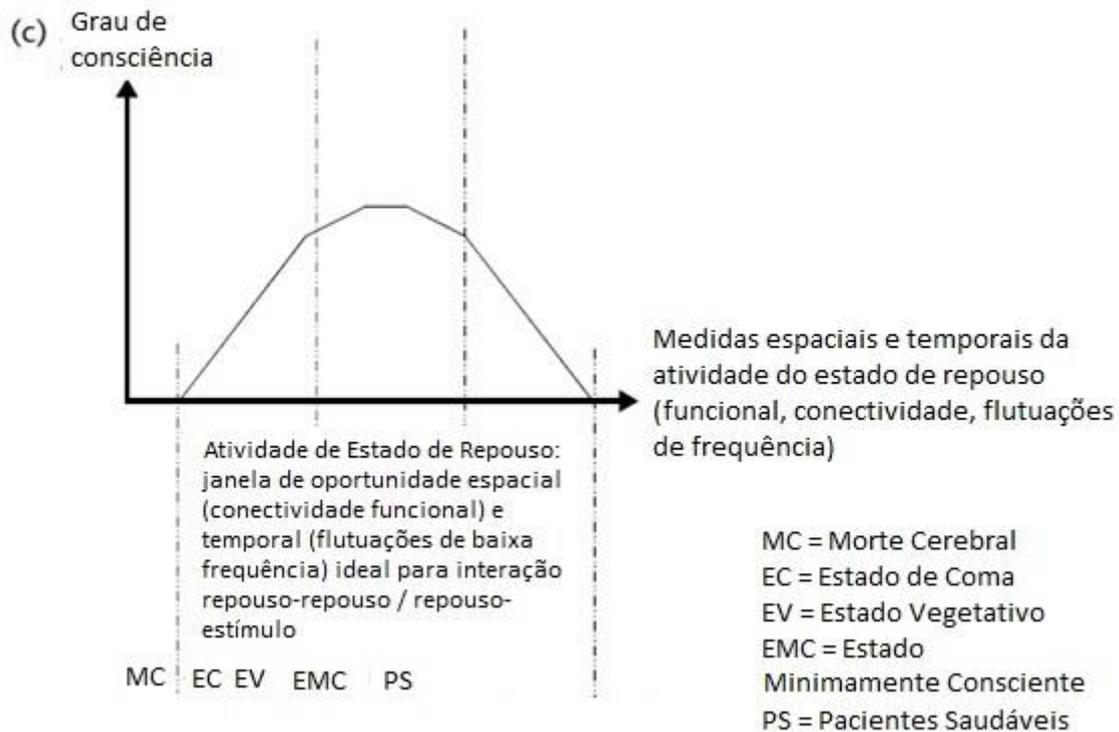


Figura 6.1.: Espectro biofísico-computacional do cérebro: constituído pelos limites mínimos e máximos do nível de atividade do estado de repouso. MC = Morte Cerebral; EC= Estado de Coma; EV = Estado Vegetativo; EMC = Estado minimamente consciente; PS = Pacientes Saudáveis (Northoff, 2014d: 426).

Assim, há uma dependência direta do nível de consciência em relação à atividade do estado de repouso do cérebro, o que poderá, por sua vez, influenciar a capacidade de alteração desse estado de repouso. O nível da dependência dependerá da posição específica no espectro biofísico-computacional do cérebro.

Como indicado da Figura 6.1., o nível de consciência mais elevado está relacionado não com os extremos do espectro – o limiar mais alto ou baixo de atividade de repouso do cérebro – mas, antes, encontra-se no meio desses dois extremos (na parte central do gráfico da Figura 6.1.). Tal levará o neurofilósofo a avançar com a “hipótese do espectro biofísico de consciência”. O que afirma tal hipótese? A hipótese propõe que a posição atual do estado de repouso em relação ao espectro biofísico-computacional do cérebro esteja diretamente ligada ao grau possível de consciência (cf. Northoff, 2014d: 427)

Northoff liga esta hipótese a uma segunda: a “hipótese de codificação baseada na diferença da consciência”. Aqui propõe-se que o nível ou estado da consciência esteja diretamente ligado ao grau de diferença da codificação durante o estado de repouso do cérebro ou durante a atividade induzida por estímulo. A tese do neurofilósofo prevê que diferenças maiores de graus de codificação tornarão possível

uma maior codificação das diferenças espaciais e temporais na atividade neuronal. Quanto maior forem essas diferenças, maior será o nível de consciência associado (cf. Northoff, 2014d: 432).

De que forma se podem cruzar as duas hipóteses agora mencionadas? Vejamos as palavras do autor:

Como é que a “hipótese de codificação baseada na diferença da consciência” está relacionada com a “hipótese do espectro biofísico de consciência”? Propus que o grau de metabolismo global e, portanto, o fornecimento de energia seja central na determinação da “posição” do estado de repouso em relação à posição do “espectro biofísico-computacional do cérebro” [...] Quanto maior for o metabolismo, maior será a probabilidade de o estado de repouso estar “localizado” no meio do espectro biofísico-computacional. Ao mesmo tempo, o metabolismo e o fornecimento de energia são relevantes em predispor o grau de possível codificação baseada em diferenças: graus mais altos de energia devem acompanhar graus mais altos de codificação baseada em diferenças e a consequente codificação de maiores diferenças espaciais e temporais na atividade neural. E, como já disse antes, quanto maiores as diferenças codificadas, maior o nível de consciência que pode possivelmente ser atribuído à atividade neural daí resultante (Northoff, 2014d: 434).

Assim, em suma: a reatividade ou propensão neuronal de mudança da atividade do estado de repouso está diretamente relacionada com o grau de consciência e, indiretamente, relacionado com os *qualia*. De que forma todos estes detalhes neurocientíficos se ligam aos *qualia*, o foco desta secção?

Northoff sugere que os *qualia* estejam ligados ao grau das diferenças espaciais e temporais codificada na atividade neuronal: quanto mais elevadas as diferenças espaciais e temporais codificadas, maior grau de subjetividade terá a experiência consciente particular. Além disso, as diferenças espaciais e temporais dependem das distribuições estatísticas de frequência dos diversos estímulos ou da atividade de repouso através de pontos discretos no tempo e no espaço. Portanto, os *qualia*, nesta Abordagem da Neurofilosofia Não-Redutiva são vistos como tendo uma natureza (i) espaçotemporal e (ii) estatística, sendo que os mesmos são dependentes da estratégia de codificação particular do cérebro (cf. Northoff, 2014d: 436).

O que significa caracterizar os *qualia* como espaçotemporais? Significa que, intrinsecamente, os *qualia* têm uma natureza espacial e temporal homogênea e unificada, não sendo discretos e distintos. Essa caracterização espacial e temporal estará relacionada, como vimos, com a atividade do estado de repouso e da sua interação com estímulos externos. Uma implicação desta definição de *qualia* é que, se

essa estrutura espaçotemporal estiver em falta, então a associação de um caráter qualitativo subjetivo a determinado estímulo estará igualmente em falta.

Considere-se, por exemplo, os pacientes em (EV): devido à falta do fornecimento de energia, a atividade do estado de repouso e as estruturas espaçotemporais ligadas à mesma estão “congeladas”, não permitindo que se possam adaptar aos estímulos exteroceptivos e interoceptivos. Assim, é importante notar que o estado de repouso, por si só, não chega para compreender a natureza intrínseca dos *qualia*: é necessário, antes, cruzar o estado de repouso do cérebro com as estruturas espaçotemporais e as suas diversas camadas e interações para que tal seja caracterizado como uma condição necessária dos *qualia* (cf. Northoff, 2014d: 480).

Uma segunda lição importante em relação aos *qualia* é a crítica que Northoff faz às concepções convencionais de *qualia* utilizadas tanto por filósofos como neurocientistas: estas costumam ser conectadas meramente aos estímulos externos, como no caso do Argumento da Mary. Ora, Northoff não concorda com este argumento, porque pressupõe uma abordagem aos *qualia* baseada meramente no estímulo extrínseco e na sua atividade induzida por estímulo, negligenciando a importância do estado de repouso, da interação do estímulo com o mesmo e, por fim, da relevância da estrutura espaçotemporal entre o mundo (estímulo) e o cérebro (repouso).

Embora essa forma convencional de investigar os *qualia* possa levar a hipóteses neurocientíficas empiricamente plausíveis, e caracterizando como condição suficiente para os *qualia* o mero estímulo externo, estas hipóteses não conseguem explicar por que razão esses mecanismos neuronais estão associados a uma experiência subjetiva consciente particular – um *quale* – em vez de uma experiência não-consciente. Assim, Northoff critica o pressuposto – que se manifesta em maior grau na Abordagem do Reduccionismo e em menor grau na Abordagem da Neurofenomenologia – de assumir-se à partida que o estímulo está relacionado (seja externo no primeiro caso ou mais interno no segundo) necessariamente com um *quale* específico. É exatamente esta pressuposição que leva grande parte da Filosofia da Mente e da Neurociência atual a postular a “lacuna explicativa” (cf. Northoff, 2014d: 481).

Ora, Northoff, ao focar-se menos no estímulo em si, mas mais no estado de repouso e da sua correspondente atividade, pensa ter evitado essa lacuna, afirmando, então, uma espaçotemporal e estatística da interação entre o repouso e o estímulo que permitirá explicar que, se essa interação estiver no equilíbrio certo (em diversas frequências e ondas cerebrais), permitirá que haja condições para que possa haver a associação de um *quale* a um estímulo e, portanto, tornar essa relação necessária.

Isto é algo que a concepção anterior teria dificuldade em explicar segundo Northoff: “Devido à escolha do ponto de partida ‘correto’, da estratégia de codificação do cérebro e da estrutura

espaçotemporal da sua atividade intrínseca, a questão da ‘lacuna explicativa’ não pode mais ser levantada [*sc. na minha concepção*]” (Northoff, 2014d: 483).

Finalmente, os *qualia* são tidos por esta Abordagem da Neurofilosofia Não-Redutiva como relacionais. Essa conclusão vem da análise das interações entre os estímulos do corpo (interoceptivos) e os estímulos do ambiente (exteroceptivos): Northoff argumenta que não existem estímulos puramente interoceptivos ou puramente exteroceptivos. Antes, estamos perante um espectro interativo em que um estímulo mais corporal é incorporado na própria atividade neuronal afetando as estruturas espaçotemporais que, por sua vez, já foram afetadas e alteradas pelos estímulos externos.

Northoff conclui esta caracterização dos *qualia* de investigações em pacientes (EV) que mostraram graus elevados de codificação de estímulos externos, mas graus baixos de codificação das diferenças entre a atividade induzida por esse estímulo e a atividade do estado de repouso. Ora, essa diferença explica por que razão esses pacientes não teriam quaisquer *qualia* associados ao processamento dos estímulos e demonstra, assim, que uma caracterização “isolada” dos *qualia* é implausível sendo que, pelo contrário, uma natureza relacional dos mesmos é confirmada pela investigação empírica em questão (cf. Northoff, 2014d: 527).

Assim, o defensor da Abordagem da Neurofilosofia Não-Redutiva conclui que:

Tal significa que [...] os *qualia* e a consciência devem resultar desta base estatisticamente espaçotemporal baseada numa ligação “virtual” entre cérebro, corpo e ambiente. *Qualia* são, portanto, intrinsecamente relacionais; tal implica que os *qualia* seriam impossíveis sem essa base estatística e espaçotemporal e [*sc. portanto*] essa ligação “virtual” baseada no cérebro, corpo e meio ambiente (Northoff, 2014d: 528).

Tal ajuda-nos também a compreender por que razão Northoff é cético a respeito da experiência mental de Mary: estamos perante um organismo cuja estrutura espaçotemporal se encontra perfeitamente desalinhada, onde os *qualia* são tidos como meras associações pressupostas aos estímulos externos – por exemplo, ver a flor vermelha – e onde a interação com a atividade do estado de repouso é tida como irrelevante.

Para concluir, faltará fazer notar que Northoff poderá estar muito próximo de cometer um erro que ele próprio aponta às abordagens anteriores: o de confundir o conceito de “correlação” com o conceito de “causalidade”. Dado o seu foco profundamente focalizado nas investigações da atividade cerebral, a forma como, por vezes, o neurofilósofo se refere a diversas descobertas, associando-as

diretamente à experiência consciente, deve levantar alguma suspeita. O erro está em, por um lado, criticar o foco das explicações reducionistas no neuronal, assumindo que essa explicação tem uma força causal – e não meramente correlacional – mas, por outro, fazer uso de estudos neurocientíficos semelhantes que poderão estar a apontar para meras correlações entre dados, e não uma causalidade direta entre os mesmos.

Contudo, ficou também claro que o defensor da Abordagem da Neurofilosofia Não-Redutiva, de certa forma, consegue conciliar as três abordagens anteriores de forma particularmente criativa: em relação à abordagem isolacionista, Northoff concorda que é um erro confundir ou até reduzir o trabalho filosófico com o trabalho neurocientífico; em relação à abordagem reducionista, concorda que o cérebro e a investigação neurofisiológica possui uma importância fundamental para o estudo da mente consciente; contudo, essa importância deve ser distribuída com a relevância do papel do corpo e do ambiente na construção dessa mente consciente, lição que Northoff vai buscar à Abordagem da Neurofenomenologia.

Finalmente, deve ser notado que esta abordagem procura enquadrar-se tanto ao nível conceptual como ao nível empírico sendo que, por isso, a caracterização avançada de *qualia* representa, em relação às abordagens analisadas anteriormente, maior plausibilidade conceptual e empírica e que, por isso, merece a nossa aposta.

Restará analisar, sucintamente, de que forma as quatro abordagens definem o conceito neurocientífico de “informação” para depois concluir esta reflexão com a exemplificação de uma investigação neurofilosófica não-redutiva ao fenómeno da percepção.

VI.2. Conceito de “Informação” e as Diversas Abordagens

Dada a natureza multidisciplinar do conceito de Informação, será uma tarefa difícil especificá-lo em cada uma das abordagens anteriormente analisadas, porque não se trata de um conceito primariamente central, embora possa ter um papel relevante em termos secundários.

A primeira referência que encontramos ao conceito de informação aparece em Platão, sendo um conceito ligada à sua teoria do conhecimento e teoria ontológica das Formas. Quando Cícero e Agostinho de Hipona discutiram a teoria platónica fizeram uso do termo *informare* e *informatio* para descrever uma variedade de termos gregos como *eidos*, *idea*, *typos*, *morphe* e *prolepsis*, sendo que a palavra atual contém uma referência direta a *morphe* (forma): in-forma-ção (cf. Capurro & Hjørland, 2003). Capurro

(1978) explora com enorme detalhe as origens gregas e latinas do conceito de informação e o seu desenvolvimento pelo que evitamos recriar aqui essa mesma história.

Embora esta seja longa e complexa, o que o conceito de informação levanta é um travão conceptual devido à sua natureza profundamente multidisciplinar e, por isso, ambígua. Estamos, portanto, ao lado de Bogdan (1994) que afirmou o seguinte sobre este intrincado conceito:

O meu ceticismo em relação a uma análise definitiva da informação reconhece a infame versatilidade este conceito. A noção de informação foi usada para caracterizar uma medida de organização física (ou diminuição da entropia), um padrão de comunicação entre fonte e recetor, uma forma de controlo e *feedback*, a probabilidade de uma mensagem ser transmitida através de um canal de comunicação, o conteúdo de um estado cognitivo, o significado de uma forma linguística ou a redução de uma incerteza. Esses conceitos de informação são definidos em várias teorias, como em física, termodinâmica, teoria da comunicação, cibernética, teoria da informação estatística, psicologia, lógica indutiva e assim por diante. Parece não haver uma ideia única de informação sobre a qual esses vários conceitos convergem e, portanto, nenhuma teoria proprietária da informação (Bogdan, 1994: 53).

Embora fosse interessante detalhar uma história do conceito de informação, remetemos a mesma para Adriaans (2018) e Capurro (1978; 2003) dado que o mais relevante é compreender sucintamente de que forma as diferentes abordagens analisadas até ao momento fazem uso do conceito de informação nas suas teorias.

Desde a revolução cognitiva de 1956, fundou-se um paradigma de investigação em Psicologia chamado de “paradigma do processamento de informação” (cf. Gärdenfors, 1999; Franklin, 1995; Gardner, 1983; 1985). Neste sentido, o conceito de informação é tido com um conceito científico que emerge de qualquer objeto cuja natureza possa ser medida qualitativamente ou possa ser quantificável.

Assim, podemos avançar com a forma como a abordagem reducionista encara este conceito: dando-lhe um cunho reducionista, esta abordagem define o conceito de informação como a capacidade do ser humano – e principalmente do sistema nervoso central ou o cérebro – extrair dados das propriedades físicas e químicas dos estímulos sensoriais.

Observemos, nas palavras da Churchland, esta atitude reducionista em relação ao conceito de informação dentro da sua abordagem reducionista:

[...] se queremos entender a natureza do processamento de informação subjacente a funções como o pensamento e o controlo sensoriomotor, as nossas teorias devem ser restringidas pela forma como os neurónios são de facto orquestrados, e não podemos entender isso sem conhecer os próprios neurónios, conhecer as suas conexões com outros neurónios e como é que formam essas conexões (Churchland, 1986: 36).

O conceito de “informação” é usado por Churchland mais de 166 vezes na sua obra principal. Todos estes usos têm em comum o sentido reducionista em que informação remete para um conceito puramente neuronal, sendo que o sujeito da predicação desse conceito é, na grande maioria das vezes, o neurónio e, em menor número, o cérebro.

Esta forma reducionista de definir este conceito será criticada de duas formas:

- (i) A Abordagem do Isolacionismo, seguindo o seu projeto metafilosófico central, defenderá que o conceito de “processamento de informação” tem uma origem indireta no cartesianismo, como analisámos em detalhe, mas que foi adaptada pela revolução computacional. Neste caso, faz-se a aplicação dos predicados psicológicos, que estão mal aplicados ao cérebro, às máquinas, e fala-se, por consequência, em “processamento de informação” (cf. Bennett & Hacker, 2003: 386), algo que, para os autores, se trata não mais do que de uma confusão conceptual que resulta do legado da revolução computacional e cognitiva. Ademais, Hacker observa que a conceção do cérebro como um órgão cuja função é processar informação – como afirmada pela Abordagem do Reduccionismo – é uma mera variante incoerente da “antiga concepção empirista” da forma como o ser humano adquire informações (cf. Smit & Hacker, 2014: 1098) dado que (a) é contraditório afirmar-se que o cérebro recebe “informação não-processada” que depois seria processada de algum modo como o conceito de “processamento de informação” supõe; e (b) nem o cérebro nem nenhuma parte do mesmo – e.g. o neurónio – processa informação dado que não vemos com os nossos cérebros, mas com os nossos olhos, nem caminhamos com o processamento de informação do nosso cérebro, mas com as nossas pernas (cf. Smit & Hacker, 2014: 1099);
- (ii) Já a Abordagem da Neurofenomenologia recusa o reducionismo do conceito de informação defendendo que este deve ser compreendido de uma forma mais histórica,

cultural e holística (cf. Karpatschhof, 2000). Na sua base está a crítica de que o sistema nervoso, ao contrário do que é afirmado pela abordagem reducionista, não processa informação mas, antes, cria o sentido da mesma: a cognição em geral deve ser entendida como corpórea e relacional com o ambiente; uma abordagem reducionista do conceito de informação irá focar-se somente no processamento que o sistema nervoso central faz dessa informação, deixando de fora informação que pode ser, no mínimo, relevante e, no máximo, absolutamente fundamental para compreender a experiência consciente (cf. Gordon, 2013). Alva Noë identifica a ideia de que o cérebro é uma máquina de processamento de informação com uma teoria computacional da mente que se encontra equivocada pelos motivos analisados no capítulo quarto. Nas suas palavras: “[...] o cérebro [...] nada sabe [...] e não tem capacidade de fazer quaisquer inferências sobre qualquer coisa, muito menos sobre o ambiente [...]” dado que “[...] somos nós – adultos e outros animais – que pensamos; que vemos, que sentimos, que inferimos” (Noë, 2009: 161).

Ora, fica claro que, no seu fundamento e embora partindo de uma metafilosofia dissemelhante, a crítica das abordagens isolacionista e neurofenomenológica ao conceito reducionista de informação é bastante similar. Faltará analisar, pois, a posição da abordagem de Northoff em relação ao conceito de informação que será colocada a par de uma teoria influente na Neurociência actual: a Teoria da Informação Integrada (TII).

Uma das teorias mais influentes da Neurociência contemporânea foi fundada por Tononi (2004; 2008) defendendo a ideia de que uma quantidade específica de informação integrada está relacionada diretamente com a consciência. Para esta teoria, uma experiência consciente acontece quando há uma integração muito elevada de informação em determinadas regiões do cérebro: “O nível de consciência de um sistema físico está relacionado com o repertório de estados causais (informações) disponíveis para o sistema como um todo (integração)” (Tononi & Koch, 2008: 253).

De que forma estes neurocientistas chegaram a esta conclusão? Baseado no seu trabalho empírico sobre a relação entre o sono e a atividade cerebral, Tononi pensa ter encontrado várias evidências que demonstram que a integração de informação nos circuitos talamocorticais parece ter uma implicação fundamental no surgimento da consciência dado que, quando estes são interrompidos – como no sono não REM, na anestesia ou em pacientes em (EV) – o nível de consciência é também interrompido.

Assim, para os defensores da (TII) a integração de informação é uma condição suficiente para induzir a consciência e os seus estados particulares, sendo a teoria vista como uma teoria sobre os correlatos neuronais da consciência (CNC) com poderes causais. Contudo, a Abordagem da Neurofilosofia Não-Redutiva foca-se, antes, nas condições necessárias da consciência, isto é, nas condições neuronais e corpóreas que necessariamente estão presentes para que haja uma experiência consciente num organismo.

Ora, para Northoff, a sua abordagem não é necessariamente contrária ou cética em relação ao projeto da (TII): antes, o neurofilósofo prefere defender que existe uma complementaridade entre as duas teorias. Uma diferença importante, contudo, é o ponto de partida das duas abordagens: enquanto que a (TII) infere da atividade fenomenal – quantificada matematicamente⁶¹ – para as características neuronais, a Abordagem da Neurofilosofia Não-Redutiva inverte o processo, inferindo das características intrínsecas da atividade neuronal a própria atividade fenomenal ou a consciência (cf. Northoff, 2014a: 545).

Portanto, Northoff aparenta ser mais otimista em relação à utilidade do conceito de informação na sua investigação. Embora não o veja como meramente reducionista ao modo do que é defendido pela abordagem reducionista, o neurofilósofo de Otava toma em consideração este conceito como tendo uma utilidade conceptual, descritiva e preditiva muito relevante. Mas urge fazer notar de que forma a sua abordagem faz uso do conceito de informação: ao contrário tanto da abordagem reducionista quanto da Teoria da Integração da Informação, Northoff não olha para este conceito por “si mesmo”. Antes, para Northoff, o processamento de informação é resultado de diversos pré-processamentos que ocorrem na atividade cerebral em repouso e que influenciarão diretamente, num segundo nível, o próprio processamento de informação.

Deste modo, o autor não defende a ideia de que o processamento de uma quantidade elevada de informação em determinada zona do cérebro seja suficiente para induzir o processo consciente – algo que a (TII) infere diretamente e a abordagem reducionista indiretamente – mas é necessário, antes, ter em conta (i) o que acontece antes (os pré-processos) do processamento de informação de determinado estímulo acontecer e (ii) o que acontece depois (os pós-processos) do processamento do estímulo.

Veja-se o seguinte exemplo (cf. Figura 6.2) da atividade cerebral numa zona escolhida aleatoriamente do cérebro:

⁶¹ A afirmação central da teoria é que um sistema é consciente se possui uma propriedade denominada de Φ , ou *phi*, que é uma medida que pode ser quantificada matematicamente de "informação integrada" do sistema. *Phi* corresponde ao feedback e à interdependência de diferentes partes de um sistema (que pode ser biológico ou não-biológico) (cf. Tononi, 2004: 42; 2008: 216-240).

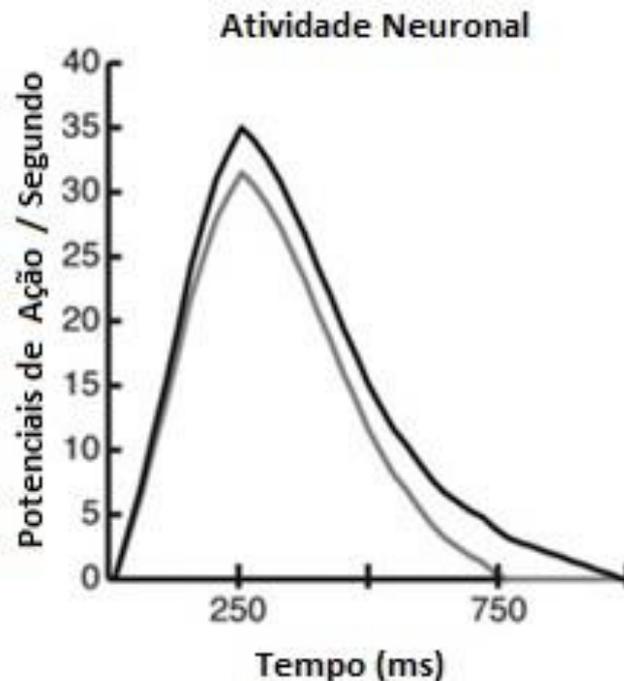


Figura 6.2.: Exemplo de atividade neuronal (cf. adaptado de James & Gauthier, 2006)

Ora, a diferença relevante entre as duas abordagens é a atenção ou relevância empírica que diferentes partes da curva da atividade neuronal do gráfico têm para cada abordagem: a Abordagem do Reduccionismo e a (TII) focam-se somente na alteração direta do sinal ou do processamento de informação (entre as linhas tracejadas da Figura 6.3), sem dar grande relevância ao que acontece antes desse pico nem o que acontece depois. Já a Abordagem da Neurofilosofia Não-Redutiva, além de considerar também como relevante a parte da atividade neuronal entre linhas a tracejado, considera igualmente fundamental perceber tanto o que acontece antes do pico do sinal como o que acontece depois, dado que se pressupõe que o que acontece entre as linhas a tracejado é profundamente influenciado tanto pelo que acontece anteriormente ao processamento de informação quanto pelo que acontece posteriormente (cf. Figura 6.3.):

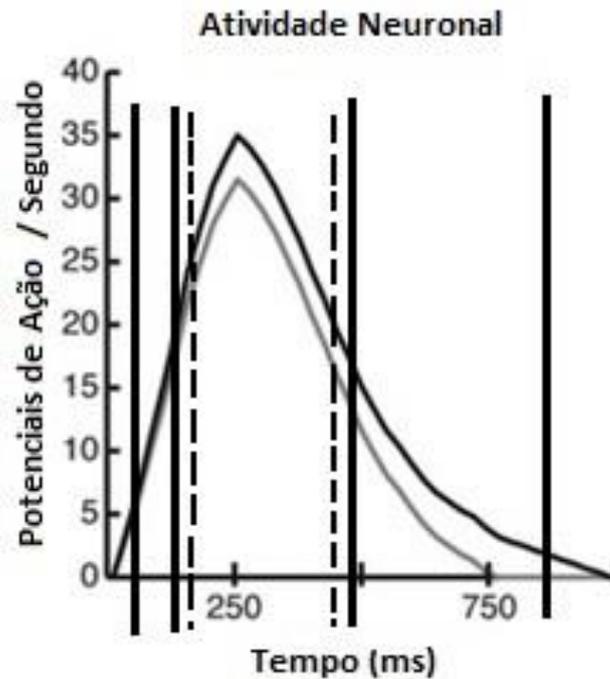


Figura 6.3.: Exemplo das diferenças de análise entre as Abordagens Reducionistas e da Neurofilosofia Não-Redutiva (adaptado de James & Gauthier, 2006).

Oferecendo um exemplo simples: para a Abordagem do Reduccionismo, interessa perceber as diferenças de atividade neuronal em diversas regiões cerebrais através de mecanismos de indução de estímulos focando-se somente nessas alterações e mudanças que o estímulo provoca na atividade neuronal local. Contudo, para Northoff, tal não chega: é preciso compreender que essa alteração é já influenciada pelos níveis de atividade neuronal do cérebro em estado de repouso – e da sua atividade (*spontaneous activity*) – e que são esses níveis que vão influenciar positiva ou negativamente o próprio processamento de informação do estímulo induzido.

É por isso que esta abordagem tem um foco específico na doença mental: percebemos exatamente esta diferença metodológica quando, por exemplo, há reações diversificadas ao mesmo estímulo por pacientes com diferentes distúrbios mentais. Por exemplo: o mesmo estímulo x pode provocar uma excitação neuronal elevada a um paciente esquizofrênico que interpreta o mesmo como um sinal de alguma alucinação auditiva. Conquanto, precisamente o mesmo estímulo x pode provocar uma inibição neuronal elevada a um paciente em estado depressivo profundo que desconsidera totalmente o valor desse estímulo.

Com este exemplo, compreendemos as principais diferenças entre as diversas abordagens e de que forma a Abordagem da Neurofilosofia Não-Redutiva pode ser uma abordagem com os instrumentos tanto conceptuais como empíricos mais afinados para lidar com os diversos problemas que a mente consciente tem causado à ciência contemporânea.

Para concluir a nossa reflexão, deixaremos a análise metodológica para demonstrar de que forma a metodologia que favorecemos em detrimento das restantes – a Abordagem da Neurofilosofia Não-Redutiva –, dado o seu equilíbrio entre o empírico e o conceptual, consegue lidar com o conceito de percepção, por um lado, e com a teoria do processamento preditivo, por outro, demonstrando assim o seu ecletismo.

Como indicámos anteriormente, esta conclusão advém da relevância dos vários argumentos avançados pelas abordagens anteriores: o rigor conceptual é necessário (cf. Capítulo 2); o rigor empírico é necessário (cf. Capítulo 3); uma abordagem corpórea é necessária (cf. Capítulo 4); mas apenas o conjunto equilibrado destas condições necessárias poderá oferecer uma mais-valia metodológica – e uma condição suficiente – para investigar os variados fenómenos mentais.

VI.3. Conclusão:

Dedicámos este capítulo à análise da segunda parte do projeto: de que forma as abordagens analisadas na primeira parte do projeto (nomeadamente a Abordagem do Isolacionismo, a Abordagem do Reduccionismo, a Abordagem da Neurofenomenologia e a Abordagem da Neurofilosofia Não-Redutiva) definem ou interpretam dois conceitos centrais: o conceito filosófico de *qualia* e o conceito neurocientífico de informação.

Em relação ao conceito de *qualia*, fizemos notar a diversidade de definições possíveis que podem ser encontradas na literatura científica sobre o tema, assim como a ligação deste conceito a uma das mais famosas experiências de pensamento da Filosofia da Mente: o Argumento da Mary ou do Conhecimento, avançado por Frank Jackson.

Em relação à Abordagem do Isolacionismo, notámos um certo ceticismo em relação à utilidade deste conceito, principalmente se for tido como algo inefável e indizível, dado que, nos ombros de Wittgenstein, os defensores desta abordagem recusam qualquer conceito que tenha na sua essência uma natureza intrinsecamente privada. Já a Abordagem do Reduccionismo procurou simplesmente negar o conceito por completo, acreditando que o mesmo será reduzido ou até eliminado do esquema conceptual que no futuro surgirá sobre as realidades mentais e cerebrais. No sentido contrário, a Abordagem da Neurofenomenologia defende que o conceito não pode ser entendido como reduzido ao cérebro, mas que deve ser visto como um conceito que abarca tanto a interação do cérebro com o corpo como a interação deste com o ambiente. Finalmente, a Abordagem da Neurofilosofia Não-Redutiva tentou apresentar uma análise mais empírica dos *qualia*, tentando perceber de que forma os mecanismos neuronais podem variar com a variação do grau de consciência de variados pacientes com lesões.

Para concluir, fizemos notar que o conceito de informação é um conceito utilizado em diversas disciplinas, sendo que o seu sentido mais útil seria a aceção de “processamento de informação”. Assim, analisámos de que forma este conceito é interpretado ou definido pelas diversas abordagens: a Abordagem do Reduccionismo reduz o conceito a uma mera troca de informação entre neurónios. Curiosamente, tanto a Abordagem do Isolacionismo quanto a Abordagem da Neurofenomenologia levantaram a mesma crítica (mas por diferentes motivos): não faz sentido aplicar o predicado de “processamento de informação” a sujeitos como “neurónios”. Finalmente, a Abordagem da Neurofilosofia Não-Redutiva tentou definir o conceito de forma holística e integradora.

De seguida, e para concluir a investigação, deixaremos a reflexão metodológica para passar à aplicação da metodologia que favorecemos na análise do fenómeno da perceção e da teoria do processamento preditivo.

VII. PERCEÇÃO: UMA ABORDAGEM NEUROFILOSÓFICA NÃO-REDUTIVA

VII.1. Introdução Geral

O fenómeno da percepção é um importante aspecto da vida mental. A grande maioria das experiências que temos advém da experiência perceptiva: o sabor de um bom café ou de uma saborosa pizza. A experiência perceptiva é, também, vital para a nossa sobrevivência. Desde os tempos antigos, o conceito de percepção tem sido assumido como parte das investigações sobre a natureza da realidade ou de como temos acesso ao mundo externo. A questão filosófica acerca da natureza da percepção é, contudo, diferente da questão empírica: isto é, como trabalham, em geral, os órgãos de percepção? Nestes termos, indaga-se acerca dos mecanismos específicos da visão ou do olfato, por exemplo. Mesmo se, num cenário de perfeição epistémica, em que conhecemos todos os factos e detalhes, por exemplo, uma teoria neuropsicológica completa da visão e do olfato, poder-se-á sempre perguntar: por que é que este processo – descrito empiricamente em todos os seus detalhes – oferece razões para justificar, por exemplo, a crença de que alguém está a ver, neste momento, uma mesa?

Neste capítulo, iremos dar atenção a uma das propostas mais relevantes para o problema da percepção: nomeadamente o Realismo Directo (RD) e o seu enquadramento numa Abordagem da Neurofilosofia Não-Redutiva. Iremos caracterizar a principal ideia desta abordagem e das suas várias subposições (e qual o modelo da percepção – sensorial, cognitivo ou outro – a que estão interligados) subjacente à principal intuição do (RD). Demonstraremos, também, por que razão é que esta abordagem é mais plausível do que as teorias da percepção do Realismo Indirecto. No entanto, (RD) padece de alguns problemas, sendo o mais notório a sua adoção de um modelo de percepção que é implausível. Pressupondo um modelo de percepção empiricamente mais plausível, elaboraremos, assim, uma alternativa neurofilosófica não-redutiva: uma abordagem espaçotemporal da percepção.

VII.1.1. Teorias da Percepção I: Realismo Direto

A experiência perceptiva é um dos tópicos mais antigos do pensamento humano. Por exemplo, a Alegoria da Caverna platônica ou o Génio Maligno cartesiano levantam questões sobre a validade do conhecimento que os nossos sentidos recebem do mundo. O problema sobre a natureza da experiência perceptiva pode ser colocado como uma luta entre duas posições claras sobre a percepção e sobre a possibilidade real de um certo tipo de erro perceptivo, a saber, a ideia de que temos acesso direto ao mundo, mas que, às vezes, esse acesso padece de precisão (cf. Crane, 2005: 241).

Por que razão é tão difícil combinar estas duas ideias? Vejamos com mais detalhe a problemática ligada à percepção. A nossa conceção ordinária a respeito da percepção está dividida entre (a) a ideia de que há objetos que existem no mundo; (b) que estes objetos não estão dependentes da nossa atividade perceptiva (i.e. que são independentes da percepção que temos deles) e, por vezes, (c) estamos perceptivamente conscientes de tais objetos materiais que existirão alegadamente de modo independente da mente.

A primeira tese (a) é uma posição ontológica, sendo uma pretensão sobre o que existe e o que é real. A segunda tese (b) é uma posição epistémica já vinculada a uma teoria específica da percepção. Por agora, iremos apenas indicar que os objetos que percecionamos são independentes da mente no sentido de que a sua existência não depende da nossa atividade perceptiva em relação a eles. A ideia inversa seria afirmar que (alguns) objetos no mundo são dependentes da mente, isto é, para um objeto x ser dependente da mente, tem de ser um objeto da experiência perceptiva – tem de ser percecionado.

A ideia está, de certa forma, ligada a uma famosa posição filosófica de George Berkeley: “esse est percipi” – ser é ser percecionado. Uma consequência importante, por agora, é que estes x 's são considerados entidades mentais com uma ontologia particular. A última pretensão (c) é uma afirmação trivial de que durante a maior parte da nossa vida – quando não estamos a dormir ou sonhar – vemos objetos no mundo que existem independentemente de serem percecionados.

Contra estas teses sobre percepção, temos vários casos que podem contradizer algumas das nossas afirmações iniciais. Por enquanto, peço ao leitor que observe a imagem seguinte:

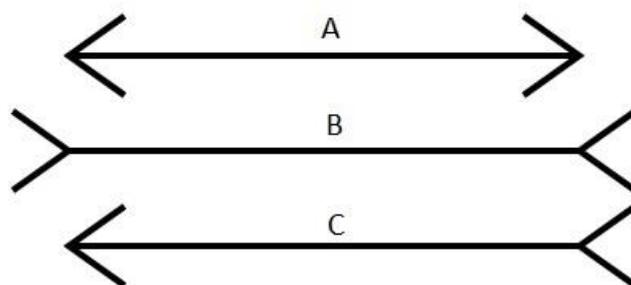


Figura 7.1.: Ilusão de Müller-Lyer (original)

A experiência perceptiva da imagem (cf. Figura 7.1.) mostra, claramente, que as linhas A, B e C são diferentes: A é menor que C e B é maior que C. Não estamos embriagados, não consumimos qualquer alucinogênio, não estamos a sonhar e o nosso sistema visual está a funcionar perfeitamente. Temos, então, que concluir que a experiência visual direta que estamos a ter deve ser coerente com as nossas três afirmações [(a), (b) e (c)] sobre a percepção. Portanto, as linhas A, B e C são independentes da mente e realmente existem com tais e tais características no mundo.

Porém, se essas linhas não fossem tão diferentes quanto a imagem que a nossa experiência perceptiva criou, o que tal significaria para a primeira abordagem à percepção? Tal significaria que, talvez, (a), (b) ou (c) possam estar efetivamente equivocadas.

Observe-se agora a seguinte imagem:

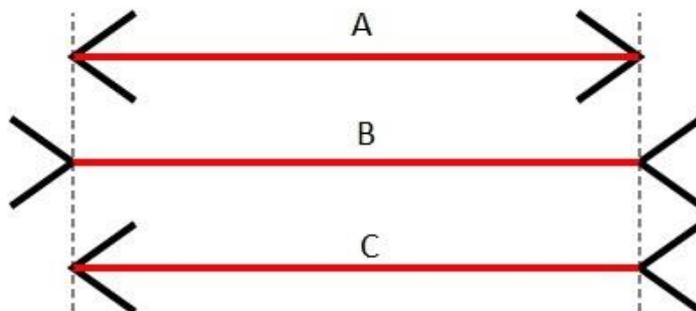


Figura 7.2.: Ilusão de Müller-Lyer (corrigida).

Como pode ser claramente observado depois de sermos confrontados com a Figura 7.2., a nossa primeira percepção sobre as linhas estava equivocada: na verdade, todas as linhas têm o mesmo comprimento ($A = B = C$). No entanto, tal vai contra a nossa primeira experiência visual da imagem. O que está errado nesta abordagem?

Note-se que, por mais artificial que este exemplo seja, casos semelhantes de ilusão e alucinação ocorrem em variadas doenças mentais. Por exemplo, na esquizofrenia, os indivíduos possuem geralmente várias alucinações – podem ouvir, ver, cheirar ou sentir coisas que mais ninguém percebe. A alucinação mais comum nesta doença específica é ouvir vozes nas suas cabeças. Estas alucinações são muito reais para a pessoa, mesmo que indivíduos ao redor da mesma não consigam ouvir qualquer voz.

No caso da alucinação auditiva, existem evidências empíricas sólidas para o aumento da atividade do estado de repouso no córtex auditivo (Northoff, 2014e; Northoff & Qin, 2011). Essa atividade pode levar à constituição de *inputs* previstos involuntariamente fortes que não são afetados por nenhum *input* real, dando ao *input* previsto e gerado internamente maior influência do que o *input* real, tendo como resultado para o sujeito ouvir as alucinações auditivas. Portanto, em geral, o cérebro de um esquizofrênico confunde o *input* previsto com o *input* real, ignorando o segundo (gerada externamente no mundo) e concentrando-se no primeiro (gerado internamente), algo que é empiricamente verificado por variados estudos (cf. Northoff, 2014e; Northoff & Qin, 2011).

Dados empíricos indicam que este tipo específico de alucinação auditiva pode ser desencadeada por um desequilíbrio entre a Rede de Modo Padrão (*Default Mode Network* – RMP)⁶² e a Rede Executiva de Controlo (*Control Executive Network* - REC) e da sua relação com o córtex auditivo: a (RMP) parece menos conectada no estado de repouso,⁶³ que está fortemente ligado ao (REC) (cf. Northoff, 2015; Ganzetti & Mantini, 2013).

Portanto, as alucinações auditivas verbais (AAV) – a experiência de perceber uma voz externa sem estímulos externos – é um caso natural que afeta mais de dois terços dos pacientes com esquizofrenia. Estes pacientes podem também sofrer vários delírios: estes delírios⁶⁴ são falsas crenças ou más interpretações de um evento específico e do seu significado. Por exemplo, um paciente pode pensar que alguém está a tentar controlar o seu cérebro ou que alguém está a persegui-lo.

⁶² A (RMP) inclui várias regiões: as estruturas da linha média cortical anterior e posterior, tal como o córtex parietal lateral e o hipocampo (cf. Qin & Northoff, 2011).

⁶³ O “estado de repouso” pode ser definido operacionalmente como a ausência de estímulos específicos, como por exemplo, durante os olhos fechados (cf. Northoff e Qin, 2011).

⁶⁴ Uma análise detalhada do conceito de delírio pode ser encontrada em Gonçalves (2018).

Ademais, as alucinações e ilusões podem também acontecer com outros sentidos: por exemplo, a sinestesia é uma experiência perceptiva em que a ativação de um sentido leva automaticamente a uma experiência involuntária de um segundo sentido – por exemplo, pessoas que veem cores poderão associar, ao mesmo tempo, um determinado som. Todos os sentidos podem ser cruzados entre si, tornando a percepção da realidade diferente da maioria das pessoas. O problema geral que estes casos levantam é que estamos diante de um fenómeno que levanta muitas dúvidas sobre a visão comum da percepção. Assim, devemos conceptualmente dividir a experiência perceptiva em duas formas principais:

(i) Casos Positivos (ou experiência verídica): quando temos uma experiência perceptual real (não ilusória) – por exemplo, alguém percebe uma mesa tendo a experiência perceptiva de esta ser castanha e, de facto, ela parecer castanha;

(ii) Casos Negativos (ou experiência ilusória ou alucinatória): quando temos uma experiência perceptiva ilusória ou alucinatória – por exemplo, alguém percebe uma mesa tendo a experiência de esta ser castanha quando, na verdade, ela é preta.

Esta divisão entre (i) e (ii) será relevante para levantar algumas objeções ao Realismo Direto (RD). Podemos agora apresentar, com base no que afirmámos até então, uma versão do que é conhecido na literatura pelo “Argumento da Ilusão ou Alucinação” (AIA):

- (1) A experiência perceptiva em (ii) consiste numa experiência dependente da mente do objeto;
- (2) A experiência perceptiva em (i) e (ii) é a mesma;
- (3) Portanto, a experiência perceptiva em (i) também consiste numa experiência dependente da mente do objeto.

Este argumento é efetivamente relevante porque nega uma das principais características da visão comum sobre a percepção: (b) e (c), ou seja, que percebemos objetos que são independentes da mente. Para garantir a validade do argumento, (2) deve ser lido no seguinte sentido:

(2*) Se duas experiências perceptivas (x e y) são iguais, se x envolve a apresentação de um objeto dependente da mente, então y também envolve a apresentação de um objeto dependente da mente.

Podemos, portanto, detalhar todas as premissas do argumento: (1) afirma que em (ii) algo deve estar presente para o sujeito. Como vimos acima, os pacientes com esquizofrenia geralmente apresentam sintomas positivos como alucinações auditivas (por exemplo, ouvir vozes) e não afirmam que não estão a ouvir nada: há algo (que é ouvido) que não tem uma causa externa e que afeta a mente do paciente. No entanto, se algo for percebido em (ii), então só pode ser algo dependente da mente – dado que, por definição, essa experiência não tem uma causa externa. Portanto, devemos concluir (1), isto é, em (ii), estamos perante a percepção de um objeto que é dependente da mente do sujeito.

E em relação a (2)? Que significa a afirmação de que a experiência perceptiva em (i) e (ii) é a mesma? Para discutir essa premissa, temos de apresentar algumas ideias que levarão à conclusão (2). A primeira é afirmar que a experiência perceptiva em (i) e (ii) é fenomenológica ou subjetivamente indistinguível para o sujeito que a tem. Ou seja, ver uma mesa (real) ou ver uma mesa ilusória é a mesma experiência subjetiva – é por isso que acreditamos que nos dois casos existe uma mesa à nossa frente.

Se as duas experiências são indistinguíveis – pelo menos para o sujeito da experiência –, podemos concluir que as mesmas têm a mesma natureza. Podemos, então, combinar as duas ideias e concluir (2): a experiência perceptiva em (i) e (ii) tem a mesma natureza. Finalmente, concluímos (3) por dedução de (1) e (2).

Podemos, na sequência desta exposição, perguntar: se o (AIA) é válido – e, vamos supor, por enquanto, cogente –, o que se segue para o problema de percepção? A consequência mais importante é que este argumento colide com a imagem comum sobre a percepção que declaramos acima em (a), (b) e (c). Essas ideias completam o que chamamos de Realismo Direto (RD): uma teoria da percepção cujas intuições principais são (a), (b) e (c). Mas, como pode resistir o (RD) ao (AIA)? E como é que, detalhadamente, (AIA) vai contra o (RD)? Finalmente: com que tipo de modelo de percepção está associado o (RD)?

VII.1.2. Teorias da Percepção II: Realismo Indireto

Depois de contextualizar a posição do Realismo Direto na percepção, podemos introduzir o principal argumento contra (RD). O argumento pode ser formulado da seguinte forma:

(4) Em (i) e (ii), existe um elemento (*quale*) na experiência perceptiva;

- (5) Como o *quale* é qualitativamente idêntico, a análise de (i)_{quale} deve ser igual à de (ii)_{quale};
- (6) Em (i) e (ii), estamos conscientes de algo *F*;
- (7) Mas em (ii), por definição, não podemos ter algo externo / material; portanto, deve ser uma entidade mental (ou representações ou dados dos sentidos);
- (8) Mas seguindo (5), o que se segue é que apenas experienciamos os dados dos sentidos;
- (9) Portanto, seguindo (8), concluímos que nunca experienciamos objetos materiais ou externos, mas apenas dados dos sentidos;
- (10) Portanto, (RD) é falso.⁶⁵

Este segundo argumento mostra claramente como (AIA) vai contra o (RD). Se o argumento for válido e cogente, teremos que concluir que o (RD) é implausível e que uma abordagem realista direta da percepção está, por consequência, condenada ao fracasso. Talvez tenhamos que adotar, como o argumento sugere, uma abordagem realista indireta da percepção. O que é, então, diferente no realismo indireto (RI)?

O realismo indireto a respeito da percepção é uma teoria que nega (b): ou seja, que os objetos que observamos do mundo são independentes do sujeito da experiência. Como mostra o argumento acima, temos que concluir que não podemos ter experiência de objetos que são independentes da mente do sujeito, mas apenas uma impressão subjetiva que depende da mente. Mas o que é “indireto” e “realista” nesta abordagem?

A parte “realista” concorda com o (RD): experienciamos objetos físicos que são independentes da mente. No entanto – e aqui entra a parte indireta – fazemo-lo via ou em virtude da percepção de dados sensoriais dependentes da mente. Esses dados sensoriais são, é claro, causados e representam objetos físicos do mundo – essa é, novamente, a parte realista desta abordagem.

Portanto, o (RI) é a teoria que afirma que experienciamos os dados sensoriais imediatamente e os objetos físicos indiretamente. Novamente, a postulação de “dados sensoriais” vem da ideia de que experienciamos algo com alguma propriedade *F* e que, quando experienciamos algo com essa propriedade *F*, há algo que possui efetivamente essa propriedade *F*, um sujeito físico no mundo. No entanto, vimos que (ii) os casos negativos da percepção mostram-nos que o objeto físico não possui a propriedade *F*. Temos que concluir que o que tem a propriedade *F* é algo mental, não-material, os dados dos sentidos.

⁶⁵ Recriação da formalização de Searle (cf. Searle, 2015: 22-23).

O conceito de dado dos sentidos (*sense-datum*) foi introduzido por G. E. Moore (cf. Moore, 1953: 30). A sua definição precisa é altamente controversa, pois é usada de maneira distinta por diferentes filósofos. Por exemplo, a principal intuição do conceito introduzido por Moore era de alguma forma neutra em relação à dependência ou independência da mente daquilo que experienciamos. Portanto, se queremos ser rigorosos, para Moore não há relação necessária entre dados dos sentidos e dependência da mente, pois o conceito também pode fazer sentido em relação a um modelo de percepção de independência da mente (cf. Snowdon, 2007). Mas o sentido em que é usado atualmente defende a noção de que os dados dos sentidos estão necessariamente conectados à dependência da mente.

Uma implicação relevante que surge da utilização deste conceito é que o mundo em si mesmo é uma representação, uma réplica do mundo externo real, como se o mundo fosse algo presente na televisão e o ser humano estivesse apenas sentado no teatro, assistindo a uma representação indireta do mundo, e não ao mundo real.

O que experienciamos neste modelo indireto? Experienciamos o conteúdo das sensações ou aparências, os dados dos sentidos, que são entidades mentais em oposição a objetos físicos. Portanto, os dados sensoriais existem apenas enquanto estão a ser experienciados, em oposição aos objetos físicos que continuam a existir mesmo que ninguém os esteja a experienciar. Finalmente, os dados sensoriais são privados sem que outros tenham acesso aos mesmos em oposição a objetos físicos que são públicos – o mesmo objeto pode ser experienciado por diferentes pessoas.

É esta abordagem uma alternativa viável ao realismo indireto? Um bom ponto contra (RD) é que parece difícil compreender o papel causal dos objetos do mundo externo neste modelo indireto. Como é que os objetos do mundo real podem afetar a percepção humana direta e imediatamente?

A ciência empírica (e.g. optometria, neurologia, etc.) que estuda o fenómeno da percepção indicam-nos que a mesma envolve uma linha causal muito complexa e longa. Se tal é realmente o caso, é impossível experienciar o mundo diretamente, mas algo mais deve estar a mediar esse processo entre a percepção e o objeto (por exemplo, dados dos sentidos, representação ou ideia) – essa é a entidade direta da percepção.

Mas será que tal implica de facto a falsidade do (RD)? Achamos que não. Quando um realista direto afirma que objetos externos são os objetos diretos da percepção, o que um realista direto está a negar é que há algo mais *a priori* a mediar o processo causal longo que termina com a percepção.

Portanto, é claro que a percepção é, num sentido restrito, indireta: envolve uma longa série de relações causais. No entanto, é implausível que a percepção seja indireta, no sentido mais amplo do

termo de envolver algum mediador entre o objeto externo e a experiência perceptiva. O argumento contra (RD) somente confunde como percebemos com o que (aquilo que) percebemos.

As causas são simplesmente determinadas condições que tornam possível a percepção de um objeto. Mas é óbvio que o objeto, por si só, não é suficiente para haver uma experiência perceptiva – é também necessário um sistema nervoso a funcionar adequadamente assim como um corpo e a interação com o ambiente.

No entanto, mesmo que os dados sensoriais fossem o meio pelo qual os objetos externos do mundo são experienciados, então os dados sensoriais seriam somente o intermediário causal desse processo e nós não experienciaríamos sequer os dados sensoriais diretamente. Mas os dados dos sentidos são, por definição, o que experienciamos. E esse é um paradoxo difícil de resolver para um realista indireto.

Além disso, o (RI) é uma resposta implausível ao problema da percepção, porque a postulação dos dados dos sentidos requer necessariamente uma conexão com a representação. Mas o representacionalismo falha claramente como pressuposto, porque é assumido que algo no cérebro deve estar a interpretar os dados perceptivos do mundo, ou seja, algo deve estar no teatro a assistir ao filme que são as experiências perceptivas.

Que tipo de entidade pode estar no cérebro? Deve existir algo: geralmente tal entidade é apelidada de homúnculo. O problema com este tipo de entidade é que leva a um segundo problema: necessariamente, requer uma regressão infinita, porque teríamos que explicar como o homúnculo₍₁₎ pode entender as informações dos sentidos e explicar que é porque um homúnculo₍₂₎ está a interpretar as informações perceptivas e assim por diante até ao infinito. Devido a tal regressão infinita, essa abordagem parece conceptualmente implausível.

Ademais, o (RI) parece deixar uma porta aberta ao ceticismo sobre a existência do mundo externo: se nunca temos uma experiência direta do mundo, como podemos ter certeza de que este realmente existe? Uma maneira de evitar este problema é aceitar essa consequência e ser um idealista sobre a percepção: a única coisa que existe são as ideias. Essa abordagem é demasiado implausível, pois pode levar à postulação de um deus (e.g. Idealismo de Berkeley) ou, alternativamente, ao solipsismo.

A postulação dos dados dos sentidos também é filosoficamente implausível: estes são geralmente caracterizados por serem privados (ninguém pode aceder aos meus dados dos sentidos), estes são exatamente como afetam a minha experiência perceptiva (não podem ser mal interpretados, não podemos estar errados em relação a eles); são dependentes da mente (existem apenas quando estão a ser

experenciados) e representam coisas no mundo externo. Todas estas características são um tanto problemáticas, como mostraremos abaixo.

Finalmente, como defenderemos na segunda parte deste capítulo, a percepção parece ser não-representacional.⁶⁶ Esboçaremos uma estrutura detalhada que explicará como a percepção realmente funciona sendo plausível tanto ao nível conceptual como empírico.

O que há de errado, então, com o argumento contra o (RD)? O (AIA) apoia-se na alegação de que, quando alguém parece experienciar algo que possui alguma propriedade F , existe efetivamente algo que tem essa propriedade F . Será essa um pressuposto sólido? Seguindo Searle (cf. Searle, 2015: 18-24), o problema com o argumento está relacionado à premissa (6):

(6) Em (i) e (ii), estamos conscientes de algo F .

O problema com (6) é que esta premissa sofre de uma falácia: a falácia da ambiguidade. Como Anscombe (1965) mostrou, há uma distinção vital a ser feita na postura intencional: entre conteúdo e objeto. Ou seja: se eu vejo uma mesa à minha frente, o conteúdo é “existe uma mesa na minha frente” e o objeto é a própria mesa.

Portanto, a percepção tem conteúdo, mas esse conteúdo não é, em si, o objeto da percepção: o oposto criaria cenários semelhantes ao “argumento da ilusão”. Nos casos negativos de percepção (ii), apenas temos conteúdo, mas nenhum objeto. O conteúdo é causado internamente por um problema no sistema nervoso da pessoa que está a ter uma alucinação ou ilusão, mas deve ficar claro que, neste caso, não existe nenhuma causa externa.⁶⁷

Não podemos, em consequência do que acabámos de afirmar, tratar a própria experiência perceptiva como objeto da experiência perceptiva em (ii), pois simplesmente não há objeto que esteja a causar essa experiência perceptiva. Além disso, seguindo Northoff (2018: 269-314), não podemos inferir do domínio fenomenal da atividade consciente a sua base ontológica. Tal equivale ao que pode ser descrito como uma “falácia fenomenológico-ontológica” (11) que deve ser claramente evitada.

Contra o que acabamos de reivindicar, alguns filósofos argumentam que (5) é uma afirmação razoável a seguir. Por exemplo, J. Foster afirma que “[...] se a percepção e a alucinação têm o mesmo

⁶⁶ Uma intuição semelhante é seguida, por exemplo, por O'Regan e Noë (2001).

⁶⁷ Na mesma linha, Johnston afirma que “(...) simplesmente temos um contexto intencional, não um objeto intencional inexistente” (Johnston, 2004: 174).

caráter subjetivo, a maneira mais simples e mais óbvia de explicar isso seria supor que também tenham, no nível fundamental da descrição, o mesmo caráter psicológico na sua totalidade” (Foster, 2000: 8).

Outros filósofos, com a mesma intuição, mas com abordagens diferentes, defendem a razoabilidade de (5), como Martin (1997) ou Crane (2005). Mas mesmo que assumamos, a benefício do argumento, que os objetos de percepção nos casos negativos (ii) são os dados dos sentidos, não se segue necessariamente que estes devam ser também os objetos de percepção nos casos positivos (i).

A falácia pode ser formalizada da seguinte forma:

(11) Se x e y são fenomenologicamente indistinguíveis, então x e y são ontologicamente indistinguíveis.

Todavia, novamente, como indicámos acima, não temos nenhuma razão plausível para supor que a experiência subjetiva deva arrastar as consequências necessárias para a ontologia. Por exemplo, um holograma de uma mesa é fenomenologicamente indistinguível de uma mesa real. Mas concluir que, por causa disso, as duas mesas são reais – ou que as duas mesas dependem da mente – seria cometer (11).

Concluimos a primeira parte do capítulo, afirmando que (RD) é a abordagem correta a seguir dada a implausibilidade de (RI). No entanto, temos de descobrir qual subposição dessa abordagem mais ampla é mais plausível. Faremos isso analisando diferentes modelos de percepção e suas consequências para o problema de percepção.

Recordamos que o objetivo é perceber de que forma a metodologia que nos pareceu mais favorável na reflexão anterior (cf. Capítulo 1-6) – a Abordagem da Neurofilosofia Não-Redutiva e postulação de um modelo espaçotemporal – poderá auxiliar a investigação do fenómeno da percepção.

VII.2. Realismo da Percepção e Modelos do Cérebro em Neurociência

VII.2.1. Realismo Ingênuo e Modelo Passivo do Cérebro

Iremos, de seguida, analisar que tipo de modelo do cérebro é necessário postular para explicar a experiência perceptiva. Ao fazer isso, caracterizaremos várias abordagens e mostraremos as suas consequências para o problema da percepção. Seguindo as três alegações iniciais sobre a percepção, devemos pensar sobre o papel que o cérebro desempenha nesse fenómeno.

O primeiro modelo – o modelo ingênuo de percepção – considera o cérebro como um órgão passivo, pois a sua atividade é determinada apenas pelos estímulos sensoriais externos, sem qualquer contribuição direta ou indireta do cérebro. Esse modelo extrínseco do cérebro (cf. Northoff 2012, 2016a; Raichle, 2015) concentra-se exclusivamente no papel que o mundo externo tem na construção de nossa percepção.

A ideia principal deste modelo é afirmar que o objeto independente da mente funciona como condição suficiente para explicar o processo total que é a experiência perceptiva. Nesta abordagem, um papel ativo do cérebro é negado e a sua função é receber e processar as informações recebidas pelos órgãos dos sentidos. Além disso, a única função do cérebro é associar os diferentes estímulos entre si.

Este modelo do cérebro está associado ao realismo ingênuo, uma subposição radical da (RD) que afirma que o papel do próprio cérebro, por exemplo, a sua contribuição intrínseca para a atividade neuronal subjacente à percepção é mínima. Essa posição teme, de alguma forma, que negar isso seja sucumbir à conclusão de que (ii) os casos negativos parecem inferir: a ideia de que não temos acesso ao mundo externo e que só temos acesso às ideias que o nosso cérebro constrói.

O modelo do cérebro defendido pela abordagem ingênua pode ser caracterizado pela receção e processamento passivos de estímulos sensoriais externos. Uma consequência direta desta pressuposição empírica é afirmar que a atividade neuronal do cérebro é necessária e suficientemente dependente do estímulo externo, o que implica que o cérebro processa apenas passivamente as informações do ambiente.

A participação ativa do cérebro é completamente negada neste modelo ingênuo de percepção. Mesmo se medirmos alguma atividade espontânea no cérebro, este modelo ignorará o seu impacto e papel causal na atividade perceptiva e na atividade neuronal em geral (cf. Northoff, 2016).

Podemos então afirmar que esse modelo pode ser descrito como definindo o cérebro como um mero modelo de resposta a estímulos ou como modelo neurosensorial do cérebro. Essa abordagem

implica necessariamente que a atividade induzida pelo mesmo estímulo seja a mesma em diferentes sujeitos. Mas será que os dados empíricos mostram essa semelhança?

Como iremos analisar, a resposta será negativa, o que levará a afirmar que este modelo do cérebro é implausível para explicar o papel do cérebro no processo de percepção. Com base na descoberta da atividade neuronal dentro do próprio cérebro, totalmente independente de qualquer estímulo externo, sugerimos outro tipo de modelo do cérebro: um modelo intrínseco ao invés de extrínseco do cérebro (cf. Northoff, 2012; Northoff, 2018; Raichle, 2010).

Essa atividade espontânea parece ser empírica e conceitualmente relevante para a compreensão do papel do cérebro na percepção. O ponto crucial é que a atividade espontânea permanece independente de uma tarefa ou estímulos específicos (cf. Logothetis *et al.*, 2009; Northoff, 2014e). Devemos notar que, no entanto, a diferença entre a atividade espontânea e a atividade induzida por estímulos não é absoluta, mas sim relativa, ou seja, relacionada com diferenças espaçotemporais (cf. Klein, 2014; Northoff, 2014e).

A atividade espontânea que reflete o que operacionalmente é descrito como atividade do estado de repouso manifesta-se em várias redes neuronais como (RMP) (rede de modo padrão), (RCE) (Rede Central de Executivos), (RS) (Rede de Saliência) e (RSM) (Rede Sensoriomotora) (cf. Klein, 2014). O ponto relevante é fazer notar que, mesmo em regiões do cérebro que não dependem de estímulos externos, há atividade neuronal. Tal será relevante para o modelo do cérebro que iremos sugerir de seguida. Mais importante, mostraremos como esse modelo ingênuo do cérebro não pode explicar a atividade neuronal ligada aos pré-estímulos.

Duncan *et al.* (2015) investigaram estudantes universitários saudáveis (livres de doenças mentais ou psiquiátricas) com um questionário psicológico (CTQ) que avaliou trauma na infância. Utilizando a fMRI, mediu-se a estrutura espaçotemporal da atividade espontânea, indexada pela entropia.⁶⁸ Os resultados podem ser resumidos nas seguintes ideias (cf. Northoff, 2016):

- (a) Os indivíduos que pontuaram alto no trauma na infância apresentaram graus mais altos de entropia na atividade intrínseca do córtex cingulado anterior perigenual (CCAP) no início da idade adulta – há uma correlação entre a entropia do CCAP e a pontuação do questionário;

⁶⁸ A entropia é uma medida que corresponde ao grau de complexidade (ou desordem) baseado em estatística do sinal (conforme medido na ressonância magnética funcional) através do tempo e espaço dentro do cérebro. Mais concretamente, mede quanto o sinal num ponto no tempo e no espaço é diferente do sinal obtido nos pontos subsequentes e posteriores no tempo. Portanto, pode-se dizer que a entropia pode ser considerada como uma medida estatística da estrutura espaçotemporal da atividade do estado de repouso do cérebro (cf. Northoff, 2016).

- (b) Utilizando a espectroscopia de ressonância magnética, os investigadores mediram a concentração de glutamato no estado de repouso na mesma região do cérebro, o córtex cingulado anterior perigenual (CCAP) – observou-se que havia uma associação direta entre glutamato e trauma na infância: quanto maior o primeiro (conforme medido pelo questionário), menor a concentração de glutamato no estado de repouso no CCAP;
- (c) Duncan *et al.* (2015) incluíram ainda outro componente na experiência durante a ressonância magnética funcional: uma tarefa que aplicou estímulos aversivos,⁶⁹ tendo medido a atividade induzida por estímulos. A atividade neuronal foi medida precisamente durante a antecipação de estímulos aversivos. Observou-se que a atividade neuronal na ínsula anterior direita e no córtex motor durante a antecipação aversiva se correlacionou significativamente com as três medidas: trauma na infância, concentração de glutamato e grau de entropia.

Esses dados empíricos são importantes para o nosso modelo do cérebro, porque podemos concluir as seguintes implicações dos mesmos:

- (d) Os eventos no mundo impactam e modulam a atividade intrínseca do cérebro de tal maneira, isto é, de maneira estatística e espaçotemporal, que se podem tornar eventos traumáticos para a respetiva pessoa. A estrutura espaçotemporal da atividade intrínseca do cérebro pode, portanto, servir como “memória espaçotemporal” de como o mundo impactou e modulou o cérebro e a estrutura espaçotemporal da atividade espontânea;
- (e) Tal sugere que um evento específico no mundo pode impactar e modular as características bioquímicas da atividade intrínseca do cérebro, o seu nível de glutamato, de modo que o respetivo evento possa ser considerado traumático mais tarde – os eventos no mundo também podem modular e impactar medidas bioquímicas no cérebro;
- (f) Os dados empíricos sugerem que a atividade espontânea do cérebro é afetada passivamente pelo ambiente e ao mesmo tempo modifica ativamente a atividade cerebral; isso equivale ao que foi descrito como interação estímulo-repouso (cf. Northoff *et al.*, 2010; Huang *et al.* 2016) e não pode ser explicado por um mero modelo passivo do cérebro.

⁶⁹ Foi utilizado um estímulo tátil doloroso de curta duração no dedo indicador.

Podemos, portanto, concluir que a estrutura espaço-temporal da atividade espontânea depende da experiência anterior do respetivo sujeito. É por isso que o modelo ingénuo do cérebro falha em prever a variação interindividual. Diferentes sujeitos mostram diferentes graus de trauma na infância; isto é algo que contraria a predição do modelo ingénuo que pressupõe o mesmo grau de variação para o mesmo estímulo, sendo que essa variação poderia acontecer somente se o próprio estímulo variasse igualmente.

Se este primeiro modelo fosse o correto, deveríamos ter uma semelhança interindividual entre todos esses fatores durante a atividade induzida por estímulos. Esse não é o caso, porém, como o estudo demonstrou.

Além disso, o modelo ingénuo do cérebro concebe o estímulo externo como necessário e suficiente. Portanto, a atividade espontânea (como identificada através dos níveis de glutamato e de entropia) não deve ser contabilizada. Não é isso que os dados empíricos nos mostram, novamente. Pelo contrário, os resultados sugerem que a atividade induzida por estímulo está claramente relacionada com a atividade espontânea. Como este primeiro modelo é baseado apenas na estrutura estímulo-resposta, não pode prever essas descobertas empíricas.

Portanto, devemos concluir que o modelo ingénuo de percepção não é plausível. Podemos resumir a implausibilidade do primeiro modelo com o seguinte argumento:

- (12) Um modelo plausível de percepção deve estar alinhado com os dados empíricos;⁷⁰
- (13) O modelo ingénuo de percepção não está alinhado com os dados empíricos;
- (14) Portanto, o modelo ingénuo de percepção não é um modelo plausível de percepção.

O estudo citado dá suporte à premissa (13). Além disso, Northoff (2014b, capítulo 8) demonstra que a atividade anterior do estado de repouso predispõe o conteúdo consciente durante a atividade induzida por estímulos. Novamente, se o modelo ingénuo da percepção fosse o correto, não teríamos diferenças interindividuais e, mesmo que as tivéssemos, o modelo ignoraria o papel ativo da atividade espontânea do cérebro (pré-estímulo) e as suas consequências para o nível comportamental e para o conteúdo perceptivo.

Finalmente, este modelo não pode também explicar o papel da atividade pré-estímulo do cérebro e o seu impacto na atividade induzida por estímulo. Outro estudo de ressonância magnética focou-se no

⁷⁰ Claro, estamos perante uma metafilosofia particular nesta abordagem neurofilosófica não-reduativa, como analisámos em detalhe no capítulo quinto: o empírico e o conceptual andam lado a lado numa constante dança interativa.

córtex auditivo (cf. Sadaghiani *et al.*, 2009) e demonstrou que a atividade do estado de repouso tem uma forte influência na atividade induzida por estímulos e na percepção associada de objetos e eventos no mundo.

A estrutura da experiência era a seguinte: os sujeitos tinham uma tarefa de detecção auditiva ouvindo estímulos de ruído em banda larga em intervalos imprevisíveis de 20 a 40 ms. Sempre que o sujeito ouvia o som específico teria de pressionar um botão (caso contrário, não era necessário pressionar o mesmo). A ideia era comparar a atividade neuronal que precede as vezes em que os sujeitos detetavam o ruído com a atividade neuronal que antecede a instância do sujeito quando estes não conseguiam detetar o som específico.

O estudo constatou que a detecção bem-sucedida do som foi precedida por uma atividade pré-estímulo significativamente maior no córtex auditivo em comparação com os erros. Os mesmos dados foram analisados por outra equipa de investigadores (Sepideh *et al.*, 2015) que mostraram que redes neuronais específicas como (RMP) têm maior conectividade funcional antes dos estímulos auditivos detetados.

Considerando estes estudos (e outros: ver Nothoff *et al.*, 2010; Northoff, 2014e), podemos alegar que existem vários indícios que apontam para a tese de que o estado de repouso exerce um forte impacto sobre o conteúdo da nossa percepção. O nível de atividade pré-estímulo do estado de repouso parece ser central na determinação dos conteúdos que subsequentemente experienciamos (cf. Hohwy, 2013; 2014).

O modelo ingénuo da percepção não pode prever estes factos exatamente porque a caracterização do cérebro como um objeto estático, não temporal e isolado é implausível, como veremos no capítulo seguinte. A principal intuição de influência de Hume de que os estímulos do mundo externo são necessários e suficientes para a nossa percepção dos objetos é implausível e, portanto, concluímos (14).

Devemos fazer notar que a única coisa que demonstrámos até agora é que um modelo específico de percepção – o modelo ingénuo – baseado em (RD) é implausível. Mas o (RD) ainda é o caminho certo a seguir. Precisamos de encontrar, no entanto, o modelo específico que possa explicar adequadamente a percepção.

VII.2.2. Realismo Indireto e Modelo Ativo do Cérebro

Um segundo modelo de percepção, diferente da abordagem realista ingênua, pode ser apresentado. Em vez de se considerar o cérebro um órgão passivo com a única tarefa de associar os vários estímulos do mundo, podemos argumentar que tal não é suficiente para obter conhecimento por meio da experiência perceptiva. Em vez de oferecer um modelo passivo do cérebro, podemos dar um papel ativo ao cérebro, considerando que a sua participação no processo de percepção e cognição, em geral, também é necessária para explicar a experiência perceptiva.

Neste modelo de percepção, podemos finalmente explicar a atividade espontânea do cérebro tão empiricamente relevante. Essa atividade espontânea é independente dos estímulos externos – o papel do cérebro é, neste modelo, estruturar e organizar os dados perceptivos numa unidade coerente. Há, portanto, uma participação ativa do cérebro para processar os estímulos sensoriais. Estamos, assim, numa melhor posição do modelo anterior: podemos explicar o impacto da atividade espontânea do cérebro na atividade induzida pelos estímulos.

Podemos igualmente explicar a intervariabilidade dos sujeitos do estudo acima citado, algo que não pudemos fazer com o modelo anterior. Neste novo modelo, a atividade espontânea do cérebro afeta diretamente a atividade induzida pelo estímulo: existe uma relação direta e a primeira é uma condição necessária da última.

No entanto, este modelo não pode dar conta de alguns dos dados do primeiro estudo, a saber, o impacto da codificação dos eventos ambientais pela atividade espontânea cerebral (cf. Northoff, 2016b: 25) e talvez seja necessário um modelo mais espaçotemporal para explicar esse elemento.⁷¹

Esse modelo parece explicar o impacto da atividade espontânea na atividade induzida por estímulos. Para fazer isso, é necessário fazê-lo por meio da representação: os eventos iniciais da vida são representados no cérebro e influenciam as funções cognitivas superiores, como a tomada de decisão ou a memória.

De um modo mais geral, as impressões sensoriais do mundo são então representadas pelo cérebro para formar o conteúdo perceptivo da nossa experiência. O cérebro, aqui, tem uma contribuição ativa: processa os estímulos sensoriais e constitui as características mentais.

A única forma deste novo modelo poder contabilizar as variações interindividuais é através da representação. E este é o principal erro desse modelo: como vimos acima, os modelos

⁷¹ Essa proposta será esboçada de seguida no terceiro modelo proposto.

representacionalistas da percepção são filosoficamente problemáticos e mostraremos, agora, que são também empiricamente implausíveis:

- (15) Um modelo plausível da percepção deve estar alinhado com os dados empíricos;
- (16) O modelo representacionalista da percepção não está alinhado com os dados empíricos;
- (17) Portanto, o modelo representacionalista da percepção não é um modelo plausível de percepção.

Novamente, existem várias suposições filosóficas: (15) representa a intuição modal de que [todos] os dados empíricos devem ser explicados fazendo uso de uma linguagem representacional. Portanto, é necessário encontrar somente um contraexemplo – como veremos de seguida – que possa ser explicado teoricamente sem o uso da linguagem representacional.⁷² Tal mostrará que a necessidade da premissa é implausível.

Como podemos apoiar a nossa premissa (16)? Antes de tudo, devemos esperar que o evento inicial se correlacione apenas com a atividade induzida por estímulos, mas não com a atividade espontânea – o nível de glutamato e de entropia. Mas o estudo (Duncan *et al.*, 2015) mostra uma relação direta entre esses eventos e a atividade espontânea. No entanto, esse modelo permanece insuficiente quando se trata da codificação da vida ambiental pela atividade espontânea (cf. Northoff, 2016b: 24-25). Para conseguirmos enquadrar esse aspeto, é necessário um modelo mais dinâmico e relacional que consiga dar conta desses factos.

Além disso, esse modelo só pode explicar indiretamente os objetos da percepção: isto é, o *input* do ambiente é mediado pela representação cerebral desses objetos. O problema é que, como mostraremos a seguir, essa abordagem ignora as características temporais do cérebro que são relevantes empiricamente.

Este modelo da percepção de pendor cognitivo (em vez de sensorial) e representacionalista (em vez de ingénuo) também comete uma falácia metodológico-ontológica. Ou seja: “[...] o facto de fazermos uso de tarefas cognitivas nas nossas investigações para investigar a atividade neuronal do cérebro não implica ou justifica a suposição de que o próprio cérebro é um dispositivo cognitivo”⁷³ (Northoff, 2016c: 263).

⁷²Para mais exemplos de dados empíricos explicados sem o apelo ao conceito de representação, veja-se, por exemplo, Orlandi (2010, 2012, 2016) ou Hutto (2013).

⁷³No mesmo sentido ver, por exemplo, Thagard (2012) & Churchland (2012).

Considere, por exemplo, o uso de imagens funcionais do cérebro através de ressonância magnética funcional. Este método representa a atividade do cérebro através de pontos coloridos e manchas coloridas. Concluir que, a partir dos pontos e manchas coloridas, a atividade (representação) do cérebro é colorida seria cometer um erro metodológico flagrante. As propriedades representadas são relacionadas não pela ontologia específica do cérebro ou pela sua atividade neuronal, mas pelo método específico utilizado (cf. Northoff, 2016c: 9).

Este modelo “ativo” ou representacional do cérebro é uma atualização do modelo “passivo” ou ingênuo que pode incluir importantes características conceituais e empíricas. No entanto, e como acabámos de apresentar, falha em fazê-lo adequadamente e é necessário um modelo mais plausível para dar conta do fenómeno da percepção, algo que esperamos apresentar de seguida.

VII.2.3. Realismo Direto e Modelo Espaço-temporal do Cérebro

Até este momento, os modelos de percepção propostos apresentavam deficiências quanto a nível conceptual como a nível empírico. Que tipo de modelo pode acomodar todos os factos anteriormente relatados? Iremos de seguida propor uma alternativa que considere os dois primeiros modelos sem considerá-los mutuamente exclusivos, mas incorporando alguns dos seus recursos.

O modelo ingênuo do cérebro está correto em relação à sua posição realista direta. No entanto, a direção do ajuste não está entre o mundo e um cérebro passivo, mas entre uma relação mundo-corpo-cérebro. O conceito de alinhamento espaço-temporal pode explicar a direcionalidade necessária para elucidar a atividade perceptiva.

Por exemplo, Northoff (2018) mostra-nos que existe um processo de acoplamento entre o cérebro, o coração e o estômago que pressupõe uma direcionalidade específica entre eles. Ou seja: um cérebro saudável tenta alinhar-se às fases da atividade espontânea com os batimentos cardíacos e, por outro lado, a sua amplitude é acoplada ao início da fase do estômago. Em ambos os casos, pode observar-se essa direcionalidade: a atividade espontânea do cérebro alinha-se e integra-se ao corpo (ou seja, coração ou estômago), algo que também é suportada por indícios sobre a transferência de informações do estômago para o cérebro (cf. Northoff, 2018: 201-202).

Além disso, a relação entre o mundo e o cérebro não é atemporal, como previsto nos dois primeiros modelos, mas temporal. Mais importante, é bidirecional, porque o cérebro tenta conformar-se com as informações do mundo e vice-versa. Portanto, podemos falar sobre uma relação mundo-cérebro

neste terceiro modelo. A relação inversa – uma relação cérebro-mundo – colocar-nos-ia novamente no modelo representacionista da percepção: a direcionalidade dar-se-ia somente do cérebro para o mundo.

No entanto, teremos que explicar, neste momento, por que razão este terceiro modelo da percepção – o modelo espaçotemporal da percepção – tem mais poder explicativo do que os modelos anteriores.

Argumentaremos, com Johnston (2007), que o cérebro não representa o mundo, como sugere o modelo representacionista da percepção, mas, pelo contrário, o cérebro – isto é, a sua atividade neuronal espontânea e induzida – é afetada pelos eventos ou objetos do mundo externo. Essa afecção é um processo seletivo que é temporal, ao contrário do que acontece no modelo passivo de percepção: nesse caso, o que seria relevante seria somente um processo passivo, uma correspondência isomórfica cuja estrutura espaçotemporal seria individual. Todavia, não estamos a defender um processo tão passivo: como Johnston sugere, a forma como o modelo espaçotemporal da percepção recolhe informação do mundo dá-se de forma seletiva.

A principal diferença é aceitar, novamente, que o cérebro e o corpo selecionam partes específicas do ambiente enquanto descartam outras. Em poucas palavras: o ambiente contribui para os objetos da percepção, enquanto que os processos mentais contribuem com a seleção, tornando os processos dinâmicos e não-isolados.

Isso só é possível com base numa relação mundo-cérebro, distinta da relação indireta com que somos confrontados com o conceito de representação. Tal implica a seguinte diferença conceptual: neste modelo espaçotemporal da percepção, a representação é substituída pela seleção dos eventos ou objetos no ambiente. Como podemos, ainda, adotar uma posição realista? Utilizando o conceito de “amostragem seletiva”, estamos em posição de fornecer, com este modelo, uma explicação que é i) realista, ii) direta e iii) causal e que, além disso, pode explicar todos os dados empíricos acima citados.

Primeiramente, a amostragem seletiva pressupõe um papel ativo do cérebro: contra o modelo ingénuo da percepção, que pressupõe um papel passivo, neste modelo espaçotemporal a relação entre cérebro e mundo não é unidirecional – como no modelo ingénuo da percepção e no modelo representacionista da percepção – mas é bidirecional: o ambiente fornece a impressão sensorial com a qual o cérebro se relaciona e, ao mesmo tempo, seleciona as mesmas por meio da amostragem. Portanto, o ambiente contribui para os *inputs* e o cérebro com sua seleção.

Podemos finalmente detalhar de que forma o processo da “amostragem seletiva” funciona: ao selecionar os *inputs* do mundo, o cérebro conforma a sua própria atividade espontânea ao mesmo. A origem do conteúdo da percepção está agora “localizada” no próprio mundo e, mais especificamente, na

relação entre o mundo e o cérebro, nomeadamente no grau de conformidade em que o cérebro se adapta ao mundo – esse processo é baseado numa base estatística (cf. Northoff, no prelo: 6; 2016b: 219-252; 2016b: 253-277).

Estamos também numa posição mais vantajosa de explicar (ii) os casos negativos da percepção, i.e., os casos de cenários de ilusão e alucinação: tais casos acontecem exatamente porque está em falta o alinhamento adequado entre o cérebro e o mundo: o primeiro não se pode conformar – por variadas razões – ao segundo.

Dessa forma, o modelo da percepção espaçotemporal pode fazer algo epistemicamente poderoso: explicar por que razão os cenários de alucinação e ilusão acontecem. Neste modelo, podemos explicar tanto os casos artificiais como os casos naturais. Por exemplo, se um paciente esquizofrénico está “a ouvir” um som que não está presente no ambiente, podemos explicar esse acontecimento afirmando que tal é claramente um caso em que a relação entre o mundo e o cérebro se encontra desequilibrada, ou seja, o cérebro não está “associado” (i.e., o impacto na sua atividade neuronal é baixo) ao mundo e a característica espaçotemporal do cérebro perde o equilíbrio necessário causando, dessa forma, a alucinação. Como esse processo é baseado estatisticamente, é fácil compreender por que razão tal pode acontecer: um cérebro pouco saudável “olha” mais para si mesmo sem considerar as informações do ambiente para equilibrar a sua atividade espontânea.

Este terceiro modelo é também ontologicamente poderoso: não é necessário postular outra categoria metafísica, como a representação, para explicar (i) os casos positivos nem (ii) os casos negativos da percepção.

Um dos maiores problemas do representacionalismo é oferecer mais entidades ontológicas para explicar a percepção do que as necessárias. Com este modelo indireto da percepção, é necessário postular duas entidades:

(*) O objeto do mundo, independente da mente;

(**) A representação (ou ideia ou dado dos sentidos) do objeto, dependente da mente.

Seguindo o Princípio da Navalha de Ockham, quando confrontado com hipóteses concorrentes, deve-se seleccionar a solução com menor número de suposições ontológicas. No segundo modelo da percepção, se estivermos conscientes de uma mesa à nossa frente, estaremos diretamente conscientes da ideia (**) dessa mesa e indiretamente conscientes da própria mesa (*). Como afirmado por Campbell, tal é

problemático: “Parece que estamos diante da ideia de uma duplicação absurda do mundo material [...]” (Campbell, 2002: 156).

Considerando isso, o terceiro modelo absorve as lições dos dois modelos anteriores: aceita que o mundo exerce influência sobre o cérebro (como o modelo passivo), mas também aceita que o cérebro exerce influência sobre o mundo (como o modelo ativo). A percepção não abrange nem uma relação activa nem passiva com o mundo. De maneira bem diferente, essa relação é “dupla” entre o mundo e o cérebro (ou mente) e vice-versa.

Uma consequência principal do que argumentámos até agora é a de sustentar que a cognição não é essencialmente representacional⁷⁴ (contra, por exemplo, Fodor, 1981 e parcialmente a favor de, por exemplo, Brooks, 1991). Uma opção é afirmar que a estrutura conceptual representacionista é apenas uma *façon de parler*⁷⁵ teórica útil que pode ajudar os teóricos da mente a esboçar as suas teorias e orientar a pesquisa empírica. No entanto, a maioria dos investigadores, quando usam a abordagem representacionista da percepção – ou cognição em geral – parecem realmente acreditar que a representação é algo real, fazendo realmente parte da ontologia humana.

Outra opção é tentar redefinir a noção de representação (como por exemplo: Varela *et al.*, 1991; Clark & Grush, 1999 ou Gallagher, 2005). De qualquer forma, a noção clássica de representação deve ser substituída no discurso científico da Neurociência e Ciências Cognitivas em geral dado que faz parte de um esquema conceptual problemático para descrever dados empíricos relevantes.

Restará concluir com uma advertência relevante: o terceiro modelo poderá estar a cometer um erro que apontámos anteriormente, inspirado pela reflexão metodológica da Abordagem do Isolacionismo, nomeadamente, a de se confundir níveis ou categorias. Por exemplo: a afirmação de que o cérebro está em contacto com o mundo seleccionando informação do mesmo deve ser lida com cautela. Não se está a afirmar que, literalmente tal seja o caso. Como o próprio Hacker admite (cf. Hacker & Bennett, 2003: 91; 379) uma linguagem metafórica pode ser útil – e tem-no sido ao longo da história do conhecimento – para fazer avançar determinado problema.

Assim, embora o modelo espaçotemporal da percepção não resolva claramente o problema filosófico em questão, estamos perante um modelo que nos parece vir a ser mais promissor dadas as limitações dos modelos aqui investigados. Será necessário, contudo, aprimorar o esquema conceptual

⁷⁴ Alguns teóricos das teorias corpóreas da cognição argumentam o mesmo há algum tempo: “A ciência cognitiva corpórea radical (...) é a tese de que a cognição deve ser descrita em termos de dinâmica agente-ambiente, e não em termos de computação e representação” (Chemero, 2009: x). Essas “dinâmicas ambientais” são o que chamamos de características espaçotemporais do terceiro modelo agora apresentado.

⁷⁵ Embora outros pensem que a representação clássica já perdeu totalmente o seu poder explicativo – pelo menos no contexto de uma teoria pós-cognitivista da mente (cf. Ramsey, 2007).

da descrição do processo da percepção de forma a garantir tanto a plausibilidade conceptual mas também a plausibilidade empírica deste modelo.

VII.3. Conclusão:

Na primeira parte do capítulo, discutimos como os (ii) casos negativos da percepção influenciaram as ideias comuns sobre a percepção. A principal conclusão que obtivemos dessa discussão foi que as teorias indiretas da percepção são problemáticas e que a abordagem realista direta é a mais plausível. Na segunda parte do capítulo, tentámos conectar os diferentes modelos do cérebro usados na Neurociência actual e nas Ciências Cognitivas em geral com as posições específicas em relação ao problema de percepção. Concluimos, com base em alguns dados empíricos, que a visão mais radical (a posição ingénuo) do (RD) é problemática dado que assume um modelo implausível do cérebro. O mesmo aconteceu com a posição realista indireta: assume um modelo implausível do cérebro, mas por razões diferentes do primeiro – esse segundo modelo só pode explicar os dados empíricos ao propor o conceito de representação que é conceptualmente problemático.

Finalmente, concluimos com um terceiro modelo que parece incorporar as duas intuições dos dois modelos anteriores. O objetivo principal era mostrar como um modelo espaçotemporal do cérebro pode explicar a experiência perceptiva em ambos (i) os casos positivos e (ii) os casos negativos. Investigações mais detalhadas e específicas sobre como esse modelo funciona devem ser avançadas, algo que esperamos alcançar com uma investigação futura.⁷⁶

De seguida, para encerrar esta reflexão, iremos perceber de que forma a metodologia que favorecemos – a Abordagem da Neurofilosofia Não-Redutiva – pode ser utilizada para criticar alguns pressupostos à teoria do processamento preditivo, especificamente a leitura desta teoria do fenómeno da consciência.

⁷⁶ Este capítulo é uma tradução adaptada do capítulo “A neurophilosophical approach to perception” (Gouveia & Northoff, 2019) da obra *Perception, Cognition and Aesthetics* pela Routledge. Apesar de o capítulo ser em co-autoria, o autor desta tese escreveu o manuscrito inteiro do capítulo, tendo efectuado a investigação aí apresentada com supervisão científica do co-autor.

VIII. PROCESSAMENTO PREDITIVO E CONSCIÊNCIA: FALÁCIA DA PREVISÃO E A SUA RESOLUÇÃO ESPAÇOTEMPORAL

VIII.1. Introdução geral

O processamento preditivo (PP) (também chamado de “teoria cognitiva da codificação preditiva”) é uma abordagem recente que reúne várias ideias sobre o cérebro e a mente. A primeira delas é a refutação do que se poderia chamar de “doutrina tradicional do cérebro” – o cérebro é tradicionalmente visto como um mero órgão passivo, cuja atividade é determinada apenas pela percepção externa dos sentidos.

Mais importante, um papel ativo do cérebro é negado nessa doutrina tradicional: o cérebro tem como função a associação dos diversos estímulos entre si (cf. Engel *et al.*, 2001). Apesar desta forma de conceber o cérebro ser influente na investigação contemporânea sobre o cérebro e a mente, este modelo do cérebro tem levantado sérias dúvidas por variadas razões, como iremos notar. A ideia original do (PP) foi inicialmente exposta pelo fisiólogo Hermann von Helmholtz no século XIX (Von Helmholtz, 1867) embora possamos encontrar também uma certa intuição preditiva em Husserl que explica o fenómeno da percepção como um “horizonte composto” de diversas possibilidades perceptivas que poderão ser selecionadas (cf. Husserl, 1960: 44).

No sentido contrário à doutrina tradicional do cérebro, a abordagem do processamento preditivo concebe o cérebro como um mecanismo de previsão: o papel do mesmo é combinar os sinais sensoriais da realidade com as suas expectativas ou crenças anteriores sobre a maneira como o mundo é, formando o seu melhor palpite sobre qual foi a causa desses sinais (cf. Clark, 2013; Clark, 2016). Existem variados esboços que tentam explicar de que forma o processamento preditivo pode ser implementado pelo cérebro (veja, por exemplo, Engel & Singer, 2001; Friston, 2010; Wacongne *et al.*, 2011; Bastos *et al.*, 2012; Brodski *et al.*, 2015). Neste modelo, supõe-se que o cérebro não percebe som ou luz: o que é experienciado, antes, é o melhor palpite ou previsão do que possa ter causado essa sensação, algo que está situado no mundo externo.

Alguns exemplos podem ser oferecidos para entender a intuição desta abordagem para explicar a atividade do cérebro em geral. O primeiro exemplo é uma ilusão perceptiva:

- (a) E TE SE DISSESSE
- (b) QUE LESTE A FRASE DE CIMA ERRADAMENTE

O que está a acontecer ao cérebro entre a leitura de (a) e (b)? O cérebro usa projeções anteriores presentes nos circuitos de uma parte específica do mesmo (por exemplo, o córtex visual). O estímulo percebido (a) não mudou: a informação sensorial é exatamente a mesma na primeira leitura que fizemos (que não reparámos no erro) e na segunda leitura (que, com a informação presente na frase de baixo, tornou claro o erro da frase de cima). O que mudou então? De acordo com a explicação do processamento preditivo, o que mudou foi a melhor previsão efetuada pelo cérebro sobre essas impressões sensoriais. Percebemos, assim, que a própria previsão influencia diretamente aquilo que podemos perceber conscientemente ou não.

A ideia de usar este exemplo é a seguinte: quando se lê pela primeira vez (a), a maioria das pessoas não nota que há um erro (a posição do “se” está no lugar errado da frase); após a leitura de (b), recebe-se informações sobre o facto de que (a) está errada; portanto, na segunda leitura, o cérebro notará o erro indo ao mundo e atualizará o modelo anterior da frase. Na primeira leitura, o cérebro não percebeu os *inputs* factuais do mundo: antes, tentou adivinhar o que poderia estar lá, com base em informações anteriores sobre a língua portuguesa e nas suposições de que o autor da frase saberia português e não cometeria um erro gramatical tão flagrante. Na segunda leitura, o cérebro teve que ir efetivamente “ao mundo” para atualizar o seu modelo e perceber a frase real, isto é, o *input* real do mundo.

A principal consequência geral da forma como a abordagem do (PP) concebe o fenómeno da percepção é afirmar que, em vez de esta depender de informações do mundo para o cérebro (doutrina tradicional), esta teoria preditiva reverte a direcionalidade de todo o processo: as previsões perceptivas fluem do cérebro para o mundo (cf. Gregory, 1980). Na estrutura do processamento preditivo – *contra* a doutrina tradicional do cérebro – supõe-se que os seres humanos não experienciam passivamente o mundo. Pelo contrário, geram-no ativamente, conforme argumentado num dos primeiros estudos sobre esta abordagem:

As teorias clássicas do processamento sensorial concebem o cérebro como um dispositivo passivo, impulsionado por estímulos. Por outro lado, abordagens mais recentes enfatizam a natureza construtiva da percepção, concebendo-a como um processo ativo e altamente seletivo. De facto, há ampla evidência de que o processamento de estímulos é controlado por influências de cima para baixo [*top-down*] que moldam fortemente a dinâmica intrínseca das redes talamocorticais e criam constantemente previsões a respeito de eventos sensoriais futuros (Engel *et al.*, 2001: 74).

A estrutura de processamento preditivo não é apenas, como se poderia pensar, uma descrição específica da percepção. Para alguns investigadores, esta teoria pode explicar quase todos – se não mesmo todos – os mecanismos da cognição humana, incluindo o Eu e a consciência.

Veja-se um exemplo da investigação em relação ao Eu:⁷⁷

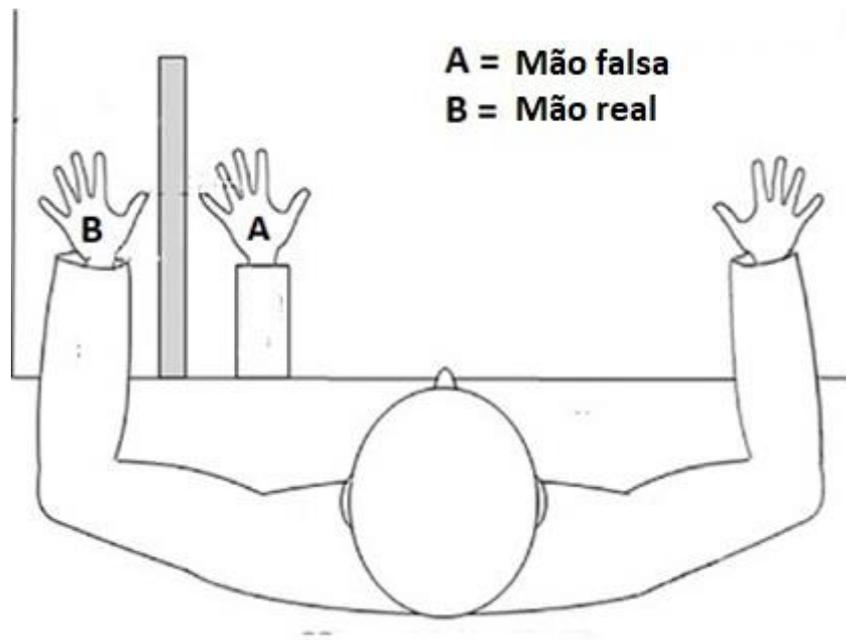


Figura 8.1.: Exemplo da ilusão da mão de borracha

A ilusão da mão de borracha é mais focada no Eu do que na percepção em si.⁷⁸ Anil Seth, por exemplo, argumenta que a semelhança entre os *inputs* de sentir um toque na mão (B) e ver um toque na mão falsa (A) que é colocada onde a outra mão (B) deveria estar é suficiente para o cérebro humano criar um modelo do corpo em que a mão falsa (A) faça parte desse corpo (cf. Seth, 2013).

Na sua opinião, a ilusão das mãos de borracha (cf. Suzuki *et al.* 2013) mostra que mesmo a experiência consciente do que é o nosso corpo é um tipo de melhor previsão, um tipo de “alucinação controlada” pelo cérebro (cf. Horn, 1980).

⁷⁷ Para uma defesa da visão do não-eu usando a estrutura do Processamento Preditivo, ver Letheby & Gerrans, 2017. Na direção oposta, uma defesa do realismo sobre o Eu pode ser encontrada em Hohwy & Michael, 2017.

⁷⁸ Nesta ilusão, o sujeito tem à sua frente uma mão falsa e, escondida da sua percepção, a sua mão real. O investigador estimula ao mesmo tempo e ao mesmo ritmo ambas as mãos (recorde-se que o sujeito apenas vê a mão de borracha a ser estimulada, mas sente a mão verdadeira a ser estimulada). Passado alguns minutos, o cérebro irá assumir que a mão de borracha é parte do seu corpo.

Assim, tal como os seres humanos podem estar enganados ou equivocados no que diz respeito à forma como percebem o mundo, também poderão estar equivocados das percepções de si mesmos. Esta tese interessante pode criar novas ideias na investigação em doença mental e psiquiátrica dado que o seu objetivo não é apenas tratar sintomas, mas também entender os mecanismos específicos relacionados ao distúrbio do Eu em condições como esquizofrenia ou transtorno de personalidade bipolar.

Alguns investigadores (por exemplo, Seth, 2013) argumentaram que faz menos sentido aplicar o (PP) aos sentidos tradicionais que transportam os *inputs* do mundo para o corpo e o cérebro (o que chamamos de exterocepção) do que aplicar esta abordagem às condições fisiológicas internas do corpo (o que chamamos de interocepção). A principal razão para tal poderá ter a ver com a sobrevivência de um organismo – pode ser mais vital evitar esbarrar em estados interoceptivos imprevisíveis do que evitar encontrar estados exteroceptivos imprevisíveis:

Um nível de oxigenação ou açúcar no sangue que é inesperadamente baixo provavelmente será uma má notícia para o organismo, enquanto que sensações exteroceptivas inesperadas (como novos *inputs* visuais) têm menos probabilidade de serem prejudiciais e, em alguns casos, são até desejáveis (Seth, 2017: 9).

Seguindo a ideia de que os cérebros são máquinas preditivas, os teóricos do processamento preditivo geralmente afirmam que a percepção e a ação são mecanismos cujo objetivo é minimizar o erro de previsão, ou seja, a discrepância entre os sinais sensoriais do mundo (externo) e as previsões sensoriais geradas internamente (interno). Um resultado dessa visão é que os mecanismos do cérebro envolvem uma interação entre o processamento de cima para baixo (*top-down*) e de baixo para cima (*bottom-up*) na cognição em geral, contra a visão tradicional (cf. Wanja & Metzinger, 2017: 1).

A sofisticação desta abordagem também está na sua formalização. De um modo geral, as estruturas de processamento preditivo podem ser descritas como defendendo as seguintes teses (cf. Ning 2019: 83-84): (1) a formação da percepção é inferencial; (2) a inferência inconsciente é preditiva; (3) a inferência realizada pela mente (PP) é (aproximadamente) de Bayes; e (4) o objetivo do cérebro é minimizar o erro de previsão. Vamos analisar com mais detalhes o que cada uma delas significa:

- (1) A tese inferencialista baseia-se na ideia de que o *input* recebido não é capturado perfeitamente: o cérebro precisa de combinar as informações ruidosas (*noisy*) do mundo (externo) com o

modelo que antecipa o que causou esse *input* específico (interno). Depois de comparar modelos diferentes – interpretações diferentes do que pode causar esse *input* específico – o cérebro deduz então o *input* “real” (cf. Kiefer 2017);

- (2) A tese preditivista deriva da suposição de que o cérebro está a antecipar a contribuição que será recebida do mundo externo. Ao contrário da explicação tradicional, a primeira etapa do processamento dos dados não é receber um *input* específico do mundo exterior, mas gerar um modelo (uma previsão) que será confirmado ou não confirmado numa segunda fase através da contribuição do mundo (cf. Anderson & Chemero, 2013: 204; Clark, 2013b: 236);
- (3) A inferência de Bayes é um método estatístico específico baseado no teorema de Bayes. O objetivo é atualizar a probabilidade de uma hipótese específica sobre o mundo à medida que mais *inputs* ou informações são processados pelo sistema. A probabilidade posterior $[P(H | E)]$ é calculada pela conjunção da probabilidade $[P(E | H)]$ com a probabilidade anterior $[P(H)]$ dividida pela probabilidade marginal $[P(E)]$, onde E significa Evidência, ou seja, os novos dados externos, e H representa uma hipótese, cuja probabilidade pode ser influenciada por E (cf. Wanja & Metzinger, 2017);
- (4) As discrepâncias entre o *input* previsto e o *input* real não são processadas passivamente, mas são ativamente aplicadas para atualizar as informações já formadas (os modelos preditivos), a fim de antecipar com mais precisão os *inputs* externos futuros. O objetivo principal das atualizações é minimizar o erro de previsão resultante das previsões anteriores feitas pelo cérebro com base em informações anteriores (cf. Friston 2010: 129).

O processamento preditivo é, então, uma teoria empírica sobre como o cérebro opera e gera a atividade induzida por estímulo ou evocada por tarefas, pressupondo, assim, um modelo preditivo do funcionamento do cérebro.

O processamento preditivo oferece uma estrutura para o estudo do cérebro e das suas características mentais, em geral. Nesse sentido, existem duas questões centrais em relação à consciência que podem ser levantadas:

- (5) O processamento preditivo é suficiente para a seleção dos conteúdos na consciência?
- (6) O processamento preditivo é suficiente para associar qualquer conteúdo à consciência?

Com base em dados empíricos sobre a previsão pré-estímulo de conteúdos conscientes subsequentes, alegaremos que, em relação a (5), o modelo preditivo do cérebro pode lidar perfeitamente bem com a seleção de conteúdos na consciência. No entanto, o processamento preditivo pode ter mais dificuldades em relação a (6), a saber, de que forma qualquer conteúdo selecionado em (5) pode ser associado à consciência, isto é, de que forma o conteúdo da percepção pode ser experienciado subjetivamente em vez de ser simplesmente inconsciente (cf. Northoff 2018: 128). Esse será o foco da próxima secção deste capítulo.

VIII.2. Processamento preditivo e falácia da previsão (argumento negativo)

Existirá alguma maneira de investigar o conteúdo da consciência? Para explorar esta questão, o grupo em torno de Andreas Kleinschmidt (Hesselmann *et al.*, 2008) utilizou imagens de ressonância magnética funcional em indivíduos humanos que estavam a ter uma experiência ilusória visual do vaso de Rubin:

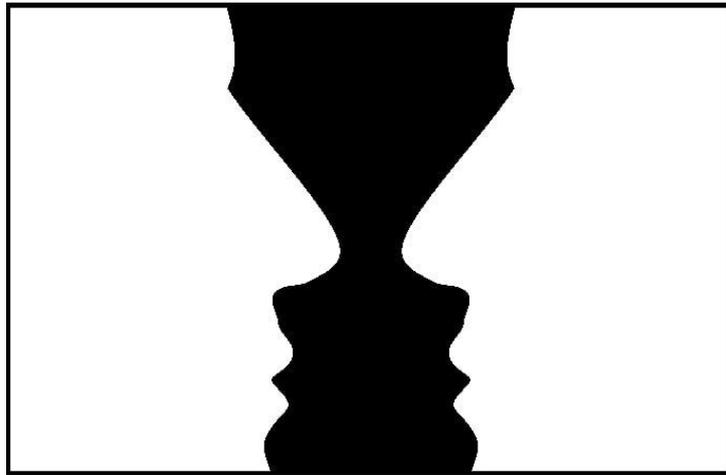


Figura 8.2.: Exemplo da ilusão do vaso de Rubin.

Esse tipo de ilusão tem a seguinte estrutura: o mesmo estímulo (figura 8.2.) pode ser percebido como tendo dois conteúdos diferentes, um como rosto e outro como vaso. Nesses casos, temos dois conteúdos perceptivos distintos criados pelo mesmo estímulo.

O fenómeno perceptivo que descreve a mudança de conteúdo na consciência é apelidado de percepção biestável ou multistável. Que lição podemos retirar deste tipo de percepção biestável sobre o conteúdo e a sua associação à consciência? A ideia relevante que é importante para o nosso debate é a seguinte: o fenómeno da percepção biestável mostra que não há uma relação direta e necessária entre um *input* específico e um conteúdo específico na consciência. Como este exemplo do vaso de Rubin mostra, os mesmos estímulos específicos podem ser interpretados de duas maneiras diferentes, ou seja, podem originar dois conteúdos diferentes, dependendo de circunstâncias neurobiológicas específicas. Mas o que está a acontecer, efetivamente, no cérebro durante as diferentes experiências perceptivas?

Os dados empíricos (cf. Hesselmann *et al.*, 2008) apontam para a existência de duas áreas principais do cérebro relevantes para o processo do *input* visual: (a) a atividade de pré-estímulo sofre alterações na área facial fusiforme (AFF) e (b) a atividade de pré-estímulo sofre alterações no córtex pré-

frontal. Ambos (a) e (b) irão impactar diretamente no conteúdo que será experienciado na consciência, mostrando que os níveis de atividade do estado de repouso adicionam algo e podem realmente manipular o *input* e, conseqüentemente, o seu conteúdo (cf. Northoff 2018: 129). Portanto, podemos declarar que é empiricamente plausível afirmar que não há relação direta e necessária entre o conteúdo da percepção, por um lado, e o conteúdo real do *input*, por outro.

De que forma é que o modelo preditivo do cérebro pode explicar o que acabamos de descrever? O grupo em torno de Kleinschmidt (Sadaghiani *et al.*, 2010) defende uma interpretação preditivista destas descobertas empíricas: se o nível de atividade dos pre-estímulos for elevada, o *input* previsto é forte e, portanto, não pode ser substituído pelo *input* real – o estímulo – que resultará num erro de previsão baixo (cf. Northoff, 2018: 133).

A única maneira de abordar esta questão com base no modelo preditivo do cérebro é cometer o que podemos descrever como a “falácia da previsão” (FP), na qual se deduz da diferença entre o processamento do conteúdo em termos do *input* previsto e do conteúdo real que existe uma relação direta com a associação desse mesmo conteúdo à consciência (cf. Northoff, 2018: 144). Todavia, não parece haver qualquer evidência empírica que apoie a associação dos conteúdos relacionados ao *input* previsto ou erro de previsão (codificação preditiva) com a consciência.

A falácia de previsão (FP) pode ser formulada da seguinte forma:

(FP): $x(\theta \vee \neg \theta)$ é consciente se $x \in (a_1, a_2, \dots, a_n)$.

Onde θ = preciso e $\neg \theta$ = impreciso e a_1, a_2, \dots, a_n = conteúdo selecionado. A formalização pode ser problemática desta forma porque seria absurdo supor que todo conteúdo selecionado seja consciente: isso não faria sentido em termos da forma como o cérebro economiza a sua energia.⁷⁹ Talvez algumas doenças mentais apresentem um cenário no qual uma versão forte de (FP) possa ser aplicado: um cérebro anormal pode processar muitas informações ao mesmo tempo, ignorando a dinâmica espaçotemporal do mundo e o seu impacto para equilibrar a atividade cerebral e, conseqüentemente, a experiência consciente. Esta interpretação acontece se for feita uma leitura forte da falácia da previsão.

No entanto, se evitarmos uma interpretação mais forte, seguir-se-á uma versão mais fraca: para cada conteúdo selecionado (a_1, a_2, \dots, a_n) existe a possibilidade de que alguns desses conteúdos

⁷⁹ A teoria preditiva do cérebro suporta o seu axioma principal através de uma explicação evolutiva: é mais económico, do ponto de vista energético e de gestão de informação sensorial, que o cérebro tenha criado mecanismos de previsão, diminuindo assim o gasto energético do organismo como um todo.

estejam conscientemente ativos. Com esta versão mais fraca de (FP), fica claro por que razão a mesma é útil para mostrar o obstáculo que a consciência representa para a abordagem do processamento preditivo.

A falácia da previsão sugere, portanto, que o modelo cognitivo da atividade induzida por estímulos, conforme pressuposto como base do modelo preditivo do cérebro, não pode ser estendido à consciência. A lição principal a ser apreendida neste ponto é que um modelo puramente cognitivo da consciência é insuficiente – tanto em termos empíricos como em termos conceptuais – para explicar o que acontece na percepção biestável. Em poucas palavras: a consciência não vem com o conteúdo, mas estende-se além do conteúdo e da nossa cognição sobre os mesmos (cf. Northoff, 2018: 150).

Assim, argumentamos que deve haver um fator adicional que permita a associação dos conteúdos selecionados com a consciência, em vez de serem simplesmente inconscientes – mesmo quando selecionados. Tal exigirá um argumento para um modelo da consciência não-cognitivo como a “consciência não-cognitiva” proposta por Cerullo (cf. Cerullo *et al.*, 2015) que introduz uma dimensão adicional e, portanto, se estende além do modelo cognitivo.

Terminamos o capítulo a argumentar que tal dimensão adicional pode ser encontrada na dinâmica espaçotemporal do cérebro defendida pela a Abordagem da Neurofilosofia Não-Redutiva. Tal levar-nos-á a propor um modelo espaçotemporal da consciência (cf. Northoff e Huang, 2017).

VIII.3. Processamento preditivo e dinâmica espaçotemporal (argumento positivo)

Como tentámos demonstrar, a única maneira pela qual o processamento preditivo (PP) pode explicar a consciência é cometendo o que apelidamos de falácia de previsão (secção 1.2.). Como vimos, o processamento preditivo pode dar conta da seleção de conteúdos. No entanto, é problemático em relação à associação da consciência com o conteúdo específico selecionado pela máquina preditiva, como afirma Northoff:

Os conteúdos interoceptivos em si não envolvem ainda a associação dos conteúdos à consciência [*sc.* dado que] podem permanecer inconscientes (como na maioria das vezes na vida quotidiana), independentemente de serem (julgados, detetados e tomados em consideração) precisos ou imprecisos (Northoff, 2018: 145).

Portanto, a nossa tese defende que a distinção entre conteúdo preciso e conteúdo impreciso e consciência do conteúdo é empiricamente suportada por dados que demonstram como a precisão e a consciência interoceptivas podem ser dissociadas da sensibilidade interoceptiva (consciência) (cf. Garfinkel *et al.*, 2015). Nesse caso, o conteúdo selecionado (preciso ou impreciso) pode sempre permanecer inconsciente: não existe forma de explicar como é que o conteúdo selecionado pode estar diretamente associado à consciência sem adicionar uma dimensão extra, como a teoria do (PP) supõe baseando-se numa falácia da previsão.

Essa dissociação entre conteúdos selecionados e consciência ocorre igualmente com os conteúdos exteroceptivos. Novamente, esses conteúdos podem ser precisos ou imprecisos em relação à maneira como o mundo é, mas sem uma dimensão extra, os conteúdos selecionados permanecerão independentes de qualquer associação com a consciência. Tal é também suportado por vários conjuntos de dados empíricos diferentes (ver Faivre & Koch, 2014; Koch *et al.*, 2016; Lamme, 2010; Northoff & Huang, 2017; Tsuchiya *et al.*, 2015; Koch & Tsuchiya, 2012).

Northoff oferece o seguinte exemplo para mostrar quão relevante é esta ideia:

Quando conduzimos o nosso carro por estradas familiares, grande parte da nossa percepção da estrada, incluindo a paisagem ao longo da mesma, permanecerá um elemento inconsciente. Tal muda quando algo desconhecido surge, como um bloqueio na estrada devido a um acidente. Nesse caso, podemos repentinamente perceber as casas ao longo

da estrada de uma maneira consciente e perceber que, por exemplo, existem aí belas mansões (Northoff, 2018: 147).

Este exemplo mostra-nos duas ideias importantes: primeiro, há uma alteração no conteúdo à medida que o primeiro modelo da “estrada” – com base nas informações *a priori* que o cérebro recolheu das viagens anteriores – se altera até à atualização desse modelo, tendo em conta o bloqueio da estrada causado pelo acidente. Esse processo pode ser bem explicado pela estrutura do (PP), à medida que o *input* previsto muda nos dois cenários. Tal, por sua vez, implica mudanças no erro de previsão e, posteriormente, no conteúdo que é percebido conscientemente.

A abordagem espaçotemporal argumenta que, além das mudanças no erro de previsão e do seu conteúdo associado, há mais mudanças no segundo cenário: essas mudanças espaçotemporais (cf. (7), (8), (9) e (10)) são centrais para permitir a associação do conteúdo à consciência. Portanto, estamos perante duas mudanças distintas: (i) baseada na cognição, relacionada ao conteúdo, e (ii) baseada na dinâmica, relacionada com a diferença entre estados inconscientes e conscientes. Nesse sentido, a abordagem espaçotemporal não exclui a abordagem do processamento preditivo, mas serve como complemento, contabilizando a associação de conteúdos com a consciência e incorporando a maquinaria de (PP) no mundo enquanto considera ainda o papel ativo do cérebro na dinâmica espaçotemporal com o mundo.

A principal suposição, então, é argumentar que, de facto, a maioria dos elementos do ambiente permanecem inconscientes na nossa percepção e cognição em geral. Da mesma forma, quando conduzimos por uma estrada familiar, na maioria das vezes, o conteúdo do nosso corpo, como os batimentos cardíacos, permanecem entidades inconscientes. A falácia de previsão ignora o facto de que alguns dos conteúdos associados ao *input* previsto e ao erro de previsão podem permanecer inconscientes no sistema nervoso. Se é factual que parte do processamento de informações que ocorre no cérebro é inconsciente – o que podemos chamar de “processamento inconsciente” (cf. de Ouden *et al.* 2009, Kok *et al.* 2013, Wacongne *et al.* 2011, Vetter, Sanders & Muckli 2014) –, então a única maneira de ignorar isso é cometer a falácia de previsão.

A principal lição a retirar deste ponto é que o processamento preditivo e o modelo preditivo do cérebro podem explicar adequadamente a distinção entre precisão e imprecisão do conteúdo consciente: (5) o problema de seleção não é um problema real para a abordagem do (PP). No entanto, argumentámos que a abordagem do (PP) não consegue explicar de forma consistente por que razão os conteúdos (precisos e imprecisos, isto é, selecionados) podem estar associados à consciência, em vez

de permanecerem como conteúdos inconscientes: (6) o problema da associação é um problema real e precisa de uma solução diferente.

A hipótese é então defender que a estrutura do (PP) necessita de ser incorporada na estrutura da maquinaria pré-preditiva mais ampla, baseada na dinâmica espaço-temporal das relações mundo-corpo-cérebro. A teoria espaço-temporal da consciência é baseada na conjunção de quatro conceitos-base principais:

(7) Aninhamento espaço-temporal (*Spatiotemporal Nestedness*) (ANET): os mecanismos neuronais do (ANET) são responsáveis pelo nível ou estado de consciência e estão relacionados com a atividade espontânea do cérebro que mostra uma organização temporal multifacetada de diferentes escalas temporais ligadas entre si. Experimentalmente, pode ser testado com paradigmas de tarefas livres em estado de repouso e está associado à predisposição neuronal da consciência (pNC);

(8) Alinhamento espaço-temporal (*Spatiotemporal Alignment*) (ALET): os mecanismos neuronais do (ALET) são responsáveis pela seleção do conteúdo e constituem a estrutura ou forma da consciência, estando relacionados com a atividade de estímulo cerebral. Experimentalmente, pode ser testado com paradigmas de pré-estímulo e está associado ao pré-requisito neuronal da consciência (preNC);

(9) Expansão espaço-temporal (*Spatiotemporal Expansion*) (EET): os mecanismos neuronais da (EET) respondem pela dimensão fenomenal da consciência e estão relacionados com a atividade inicial do cérebro induzida por estímulos. Experimentalmente, pode ser testado com paradigmas pós-estímulo sem relato e está associado aos correlatos neuronais da consciência (CNC);

(10) Globalização espaço-temporal (*Spatiotemporal Globalization*) (GET): os mecanismos neuronais da (GET) são responsáveis pela dimensão cognitiva da consciência, o relato do seu conteúdo está relacionado com a atividade cerebral induzida por estímulos tardios. Experimentalmente, pode ser testado com paradigmas de relatórios pós-estímulo e está associado às consequências neuronais da consciência (conCNC) (cf. Northoff 2018: 155).

Como discutido na secção I.2., a consciência vai além do conteúdo e necessita de dimensões ou recursos adicionais. Que dimensões podem ser essas? A hipótese aqui avançada sugere que essas dimensões consistem nos mecanismos espaço-temporais (7), (8), (9) e (10) que permitem que o conteúdo – baseado em estímulos – seja associado a uma experiência subjetiva fenomenal – consciência. Deste modo,

situando o mecanismo de processamento preditivo na dinâmica espaçotemporal da relação entre o cérebro e o mundo, prevemos que teremos maior poder explicativo para explicar a experiência consciente no ser humano.⁸⁰

⁸⁰ Mais investigações detalhadas serão feitas para especificar de que forma o modelo espaçotemporal da consciência pode explicar as diferentes relações mundo-cérebro (ver, por exemplo, Northoff *et al.*, 2019).

VIII.4. Conclusão:

O objetivo deste último capítulo foi tentar demonstrar que a consciência não pode ser determinada apenas pelo conteúdo, não se reduzindo ao mesmo, mas necessitando de dimensões adicionais para ser explicada. Tal é, do nosso ponto de vista, empiricamente suportado por vários estudos que mostram que os mecanismos neuronais subjacentes a essas várias funções cognitivas não são suficientes para associar as funções cognitivas e os seus conteúdos à consciência.

Por esse motivo, qualquer modelo plausível da consciência deve estender-se além do conteúdo e da cognição: esta é a tese deflacionária proposta neste capítulo que é independente do argumento positivo. O processamento preditivo é um excelente primeiro passo contra a doutrina tradicional do cérebro, com enorme poder explicativo. No entanto, pode não ser suficiente para lidar com o maior problema da ciência contemporânea: explicar a experiência consciente subjetiva. Para esse fim, propusemos um argumento positivo baseado na dinâmica espaçotemporal do cérebro.

Podemos, então, resumir os dois argumentos da seguinte maneira: os mecanismos de processamento preditivo são uma condição necessária e, eventualmente, suficiente do conteúdo da consciência; no entanto, esses mecanismos não são uma condição suficiente para a própria consciência (argumento negativo). Portanto, é necessária uma abordagem mais espaçotemporal, na *Gestalt* da dinâmica espaçotemporal, que possa fornecer a condição suficiente da própria consciência como distinta do seu conteúdo (argumento positivo) algo que, do nosso ponto de vista, é alcançado exatamente por essas dinâmicas espaçotemporais da relação entre o mundo, o corpo e o cérebro.⁸¹

⁸¹ Este capítulo é uma versão traduzida do capítulo "Predictive Processing and Consciousness: Prediction Fallacy and Its Spatiotemporal Resolution" (Gouveia, 2020) da obra *The Philosophy and Science of Predictive Processing* (editado por Manuel Curado, Dina Mendonça e Steven S. Gouveia) pela editora Bloomsbury.

CONCLUSÃO

O caminho trilhado até aqui procurou investigar o problema da relação metodológica entre a investigação filosófica e a investigação neurocientífica. Analisámos, no primeiro capítulo, de que forma este problema se coloca na investigação contemporânea sobre os estudos do mental e do neuronal, além de diversos aspectos adjacentes, como as principais questões de investigação realizadas em Filosofia da Neurociência, Metafilosofia e, muito brevemente, em Neurociência contemporânea.

De seguida, avançámos com a primeira resposta ao problema metodológico, a resposta isolacionista que, no seu âmago, nega que haja sequer um problema dado que, para os defensores desta abordagem, o trabalho filosófico e o trabalho neurocientífico estão claramente separados na sua essência e qualquer tentativa de conciliar os dois ao mesmo nível estará condenada ao fracasso ou à mera confusão conceptual. Fizemos notar a influência de uma certa interpretação de Wittgenstein e dos pressupostos metafilosóficos que sustentam esta abordagem. Finalizámos a análise desta posição isolacionista levantando algumas objeções que apontam para uma forma muito tradicional de definir tanto a investigação filosófica quanto a investigação neurocientífica que poderá estar desatualizada à luz do progresso nesta área de estudo.

Deste modo, passámos à análise da Abordagem do Reduccionismo, examinando de que forma a redução pode ser concebida (como tradução, derivação ou explicação), de que forma a redução é definida por Patricia Churchland e John Bickle, os dois autores defensores desta abordagem; a relação da teoria da Psicologia Popular e da necessidade de existir uma teoria (e não um fenómeno particular) que possa ser reduzida a uma teoria futura da Neurociência e, finalmente, os vários aspetos e características da redução interteorética. Ademais, dissecámos as principais objeções levantadas a esta abordagem, como a defesa de que a teoria candidata defendida por estes autores a ser reduzida, a teoria da Psicologia Popular, poder não ser considerada uma teoria de todo, o que impediria a redução de uma teoria a outra, algo que está na base da abordagem reducionista. Finalmente, demos conta de que os pressupostos metafilosóficos desta posição metodológica podem estar equivocados dado que se defende, ao contrário do isolacionismo da posição anterior, uma ligação direta entre o trabalho filosófico e o trabalho neurocientífico, nomeadamente a redução do primeiro ao segundo. Isto é algo que, para muitos autores, é filosófica e cientificamente problemático.

Um desses autores críticos da posição reducionista é Francisco Varela, o fundador da terceira abordagem analisada acima, a Abordagem da Neurofenomenologia. Observámos a influência da

fenomenologia de Husserl, a importância do corpo (em oposição ao foco somente no cerebral) que advém da investigação avançada por Merleau-Ponty ou a importância das teses defendidas pela abordagem dos 4E's e da forma como esta influencia a proposta metodológica detalhada por Varela. Fizemos notar alguns exemplos da investigação contemporânea em Neurociência para demonstrar a aplicabilidade desta abordagem, como seja o caso da rivalidade binocular ou o Eu (*self*) em Evan Thompson, assim como três lições pragmáticas que a Neurofenomenologia deve seguir para ser epistemicamente relevante. Finalizamos a análise desta Abordagem da Neurofenomenologia com algumas objeções, como a ligação dúbia ao conceito obscuro de “introspeção” desta abordagem, uma resposta pouco clara sobre como lidar com o problema da lacuna ou hiato explicativa (*gap*) ou o facto de estarmos perante uma abordagem que, no seu pressuposto-base, apresenta uma metafilosofia no mínimo inconclusiva e no máximo incoerente.

Um segundo autor crítico da posição reducionista é Georg Northoff, defensor da Abordagem da Neurofilosofia Não-Redutiva, uma posição que procura equilibrar os dois extremos do problema metodológico, isto é, procura equilibrar metodologicamente tanto o trabalho filosófico como o trabalho neurocientífico. Para tal, Northoff irá avançar uma série de princípios (princípio da assimetria, da bidirecionalidade e da circularidade transdisciplinar) que irão fundamentar a natureza idiossincrática da hipótese neurofilosófica. Fizemos notar igualmente a diferença entre esta abordagem e a abordagem reducionista, nomeadamente o facto de a segunda ser uma abordagem que se reduz ao cérebro e a primeira que se baseia somente no cérebro, mas que não reduz o seu foco de investigação ao mesmo, defendendo, assim, uma Neurociência da Primeira Pessoa em oposição a uma Neurociência da Terceira Pessoa. Avançamos com vários exemplos que mostram o carácter global e diverso que esta Abordagem da Neurofilosofia Não-Redutiva possui, tendo um poder explicativo de potencial vasto, podendo explicar, por exemplo, o Eu (*self*), as emoções e até a relação do Eu com a COVID-19. Para encerrar a análise desta abordagem, explorámos as principais objeções que podem ser levantadas à mesma, nomeadamente a tese de que, aos ombros da Abordagem Isolacionista, esta abordagem poderá estar a confundir níveis e categorias de análise; a objeção de que, embora a hipótese metodológica avançada por esta abordagem seja interessante, dado que procura ter uma excecionalidade metodológica não se reduzindo nem ao trabalho filosófico nem ao trabalho neurocientífico, mas estando no meio de ambos num constante cruzamento iterativo, poderemos estar perante uma metodologia que não poderá impactar nem a investigação filosófica nem a investigação neurocientífica, exatamente por ser diferente de ambas; finalmente, analisámos a objeção de que esta abordagem poderá apresentar uma metafilosofia imprecisa, podendo levantar dúvidas em relação à sua diferença entre as abordagens analisadas

anteriormente, como seja, por um lado, a Abordagem do Reduccionismo ou a Abordagem da Neurofenomenologia, concluindo com a defesa de que a Abordagem da Neurofilosofia Não-Redutiva nos parece ser uma abordagem que consegue incorporar os vários pontos positivos das três abordagens anteriores sem lidar com os aspetos mais negativos das mesmas.

Depois de analisadas as quatro abordagens que tentam responder ou lidar com o problema metodológico da relação entre a Filosofia e a Neurociência, procurámos compreender de que forma dois conceitos das duas áreas – do lado da Filosofia, o conceito de *qualia*, e do lado da Neurociência, o conceito de informação – são definidos por cada abordagem. O conceito de *qualia* é explorado explicitamente por cada abordagem, o que permitiu identificar de que forma os pressupostos metafilosóficos da cada abordagem influenciam diretamente as definições de *qualia* de cada abordagem. Já em relação ao conceito de informação, fizemos notar que, por ser um conceito utilizado em variadas áreas, encontramos um entrave teórico dado que as abordagens analisadas não possuíam uma definição explícita ou central deste conceito nas suas investigações. Ainda assim, explicámos de que forma o conceito de informação é utilizado no âmbito do conceito neurocientífico de “processamento de informação”, examinando a posição de cada abordagem e de como é que estas influenciam diretamente a definição desse conceito.

Para finalizar esta investigação, concluímos a reflexão metodológica e passámos à aplicação da abordagem que nos parece mais promissora – a Abordagem da Neurofilosofia Não-Redutiva – ao conceito de percepção, examinando de que forma a plausibilidade conceptual pode ser cruzada com a plausibilidade empírica. Fizemos notar que o Realismo Direto parece ser uma teoria mais plausível do que o Realismo Indireto mas que, contudo, será necessário adotar um modelo do cérebro espaço-temporal (em vez de passivo-ingénuo ou ativo-indireto) de forma a avançarmos com uma perspetiva plausível do fenómeno perceptivo.

Ultimámos esta reflexão com a apresentação de uma teoria promissora da atualidade, a teoria do processamento preditivo, que procura explicar todos os fenómenos mentais pressupondo a tese de que o cérebro – e a sua relação com o corpo e o ambiente – é uma máquina de predição. Embora esta teoria seja realmente promissora, fizemos notar que, seguindo os pressupostos da Abordagem da Neurofilosofia Não-Redutiva, é necessário enquadrar a maquinaria do processamento preditivo em estruturas espaço-temporais mais holísticas e relacionadas com o ambiente.

A grande descoberta desta investigação doutoral é que fica claro de que forma as posições metafilosóficas dos autores, muitas vezes tomadas de forma implícita ou inconsciente, influenciam direta ou indiretamente as suas posições metodológicas na relação entre a Filosofia e a Neurociência, por um

lado, e a forma como investigam os diversos fenómenos e conceitos da sua área de investigação, como o seja o conceito de *qualia* ou de informação.

CAPÍTULO 1

PAG. 3:

“Thus, unabashed methodological naturalists, their explicit doctrines notwithstanding, did not necessarily do philosophy any differently from the way it had traditionally been done: in the main without relying substantively on – let alone actively collecting – empirical data. In practice, if not always in theory, it was assumed that philosophy could, indeed perhaps should, be done from an ‘armchair’” (D’Oro & Overgaard, 2017: 1).

“A philosopher should philosophise: he should try to solve philosophical problems, rather than talk about philosophy” (Popper, 1968: 68).

“The critical inquiry into the sciences, their findings, and their methods [...] remains a characteristic of philosophical inquiry” (Popper, 1975: 53).

PAG. 4:

“If I deny a distinction, it is the still fashionable distinction between philosophy and meta-philosophy, the philosophy of philosophy. The remarks I make about philosophy [...] are [...] nothing more or less than philosophical remarks” (Cavell, 2002: xxxii).

PAG. 8:

“I believe that the function of a scientist or of a philosopher is to solve scientific or philosophical problems, rather than to talk about what he or other philosophers are doing or might do. Any unsuccessful attempt to solve a scientific or philosophical problem, if it is an honest and devoted attempt, appears to me more significant than a discussion of such a question as ‘What is science?’ or ‘What is philosophy?’ And even if we put this latter question, as we should, in the slightly better form, ‘What is the character of philosophical problems?’, I for one should not bother much about it; I should feel that it had little weight, even compared with such a minor problem of philosophy as the question whether every discussion or every criticism must always proceed from ‘assumptions’ or ‘suppositions’ which themselves are beyond argument.” (Popper, 1968: 66)

“Philosophy is not at its most interesting when it is talking about itself” (Williams, 2006: 169).

PAG. 10:

“But if we had to name anything which is the life of the sign, we should have to say that it was its use; [...] The sign (the sentence) gets its significance from the system of signs, from the language to which it belongs. Roughly: understanding a sentence means understanding a language” (Wittgenstein, 1992: 30).

⁸² Acautelando o rigor das traduções feitas no corpo do texto, listam-se de seguida todos os textos originais das mesmas.

PAG. 12:

“the tribunal of sense experience not individually but only as a corporate body” (Quine, 1980: 41).

PAG. 13:

“it is within science itself, and not in some prior philosophy, that reality is to be identified and described” (Quine, 1981: 21).

PAG. 14:

“The sort of explanation one is looking for when one is puzzled by an aesthetic impression is not a causal explanation, not one corroborated by experience or by statistics as to how people react” (Wittgenstein, 1966: 17).

“In Ethics, as in all other philosophical studies, the difficulties and disagreements, of which its history is full, are mainly due to a very simple cause: namely to the attempt to answer questions, without first discovering precisely what question it is which you desire to answer” (Moore, 1991: vii).

PAG. 15:

“Philosophy [...] one of a number of blanket terms used by deans and librarians in their necessary task of grouping the myriad topics and problems of science and scholarship under a manageable number of headings [...] are grouped under “philosophy” makes neither man responsible for the other’s topic” (Quine, 1975: 228).

CAPÍTULO 2

PAG. 29:

“The isolationist projects are based on the assumption that the findings of the natural sciences, neuroscience included, do not influence directly the practice of philosophy, as at least some aspects of philosophical reflection are independent of empirical facts investigated by scientists” (Brozek *et al.*, 2013: 16).

“The most significant contribution to philosophy of mind in recent years. It examined thoroughly and carefully the pretensions of cognitive science to have superannuated philosophical psychology” (Kenny, 2013: xvii).

PAG. 30:

“Empirical questions about the nervous system are the province of neuroscience. It is its business to establish matters of fact concerning neural structures and operations. It is the task of cognitive neuroscience to explain the neural conditions that make perceptual, cognitive, cogitative, affective and volitional functions possible. Such explanatory theories are confirmed or infirmed by experimental investigations. By contrast, conceptual questions (concerning, for example, the concepts of mind or

memory, thought or imagination), the description of the logical relations between concepts (such as between the concepts of perception and sensation, or the concepts of consciousness and self-consciousness), and the examination of the structural relationships between distinct conceptual fields (such as between the psychological and the neural, or the mental and the behavioural) are the proper province of philosophy” (Bennett & Hacker, 2003: 1).

PAG. 31:

“What you see is not what is really there; it is what your brain believes is there ... Your brain makes the best interpretation it can [...] Your brain combines the information provided [...] and settles on the most plausible interpretation of all” (Crick, 1995: 30-57).

“The brain categorize, discriminate, and recombine the various brain activities occurring in different kinds of global mappings” (Edelman, 1994: 109).

PAG. 32:

“Our brains can often decide well, in seconds, or minutes [...] they must do the marvelous job with more than just pure reason” (Damásio, 1996: 173).

“The principle that psychological predicates which apply only to human beings (or other animals) as wholes cannot intelligibly be applied to their parts” (Bennett & Hacker, 2003: 73).

PAG. 33:

“Only of a human being and what resembles (behaves like) a living human being can one say: it has sensations; it sees; is blind; hears; is deaf; is conscious or unconscious” (Wittgenstein, 1953: 281).

“But, mind proves refractory to description by physics and chemistry. It belongs to the unable-to-be-sensed, the insensible, where physical and chemical have no place” (Sherrington, 2009: 339).

“There too the body's thinking seems to be done for it, namely, in the brain” (Sherrington, 2009: 237).

“I have seen the question asked ‘why should mind have a body?’ The answer may well run, ‘to mediate between it and other mind’” (Sherrington, 2009: 261).

PAG. 34:

“Ignorance of the 'how' of the tie between the brain and the mind [...] is no fault of those who study the mind or of those who study the brain. It constitutes a disability common to both of them. A liaison between them is what each has been asking for. That there is a liaison neither of them doubts. The 'how' of it we must think remains for science as for philosophy a riddle pressing to be read” (Sherrington, 2009: 240).

PAG. 36:

“But between energy and mind science found no 'how' of give and take [...] To man's understanding the world remained obstinately double (Sherrington, 2009: 251) [...] In fact natural science repudiates it as something outside its ken” (Sherrington, 2009: 291).

“The two for [...] remain refractorily apart. They seem to me disparate; not mutually convertible; untranslatable the one into the other” (Sherrington, 2009: 312).

“After years of striving to explain the mind on the basis of brain-action alone, I have come to the conclusion that it is simpler (and far easier to be logical) if one adopts the hypothesis that our being does consist of two fundamental elements [...] because it seems to me certain that it will always be quite impossible to explain the mind on the basis of neuronal action within the brain [...] I am forced to choose the proposition that our being is to be explained on the basis of two fundamental elements” (Penfield, 1975: 80).

PAG. 37

“During epileptic interference with the function of this gray matter [...] consciousness vanishes and, with it, goes the direction and planning of behavior. That is to say, the mind goes out of action and comes into action with the normal functioning of this mechanism” (Penfield, 1975: 47).

PAG. 38

“What you see is not what is really there; it is what your brain believes is there” (Crick, 1995: 31).

“Neurons [...] have knowledge. They have intelligence, for they are able to estimate the probability of outside events [...] neurons present arguments to the brain based on specific features that they detect, arguments on which the brain constructs its hypothesis of perception” (Blackmore, 1997: 91).

“Our brains can often decide correctly” (Damásio, 1996: 173).

PAG. 39

“Conceptual questions antecede matters of truth and falsehood [...] They determine not what is empirically true or false, but rather what does and what does not make sense. Hence conceptual questions are not amenable to scientific investigation and experimentation or to scientific theorizing [...] Distinguishing conceptual questions from empirical questions is of the first importance” (Bennett & Hacker, 2003: 2).

“We do not think that empirical research can solve any philosophical problems, any more than it can solve problems in mathematics” (Bennett & Hacker, 2003: 414).

PAG. 41

“When these painful contradictions are removed, ... our minds, no longer vexed, will cease to ask illegitimate questions” (*apud* Bennett & Hacker, 2003: 5).

Esta passagem deriva do original alemão que reproduzimos:

“Sind diese schmerzenden Widersprüche entfernt, so ist zwar nicht die Frage nach dem Wesen beantwortet, aber der nicht mehr gequälte Geist hört auf, die für ihn unberechtigte Frage zu stellen” (Hertz, 1879: 9).

PAG. 43

“L'action de l'âme consiste en ce que, par cela seul qu'elle veut quelque chose, elle fait que la petite glande à qui elle est étroitement jointe se meut en la façon qui est requise pour produire l'effet qui se rapporte à cette volonté” (Descartes, 1964: I-41).

PAG. 44

“Only a specialized zone of the cerebral hemispheres is in liaison with the self-conscious mind. The term liaison brain denotes all those areas of the cerebral cortex that potentially are capable of being in direct liaison with the self-conscious mind” (Popper & Eccles, 1977: 358).

PAG. 46

“The only ways in which a conceptual investigation can assist an empirical subject are by identifying conceptual error and providing a map that will help prevent empirical researchers from wandering off the high roads of sense” (Bennett & Hacker, 2003: 7).

PAG. 47

“It cannot be the task of a conceptual investigation to propose empirical hypotheses that might solve the empirical problems faced by scientists. To complain that a philosophical investigation into cognitive neuroscience has not contributed a new neuroscientific theory is like complaining to a mathematician that a new theorem he has proved is not a new physical theory” (Bennett & Hacker, 2003: 7).

“Philosophy does not contribute to human knowledge, but to human understanding” (Hacker, 2001: 139).

PAG. 48

“Most of my scientific and medical colleagues tend to be dismissive of philosophy. Their attitude, of course, reflects a realistic assessment of the likely relevance of philosophy to their immediate professional preoccupations” (Glynn, 1999: 367).

“One of the temptations of having a mind is to try using it alone to solve the mystery of its own nature. Philosophers have attempted this since time immemorial [...] as a general method to explore the matter of the mind, it just won't do” (Edelman, 1994: 31).

“We suggest that consciousness can be considered a scientific subject and that it is not the sole province of philosophers” (Edelman & Tononi, 2000: 3).

“Philosophers have had such a poor record over the last two thousand years that they would do better to show a certain modesty rather than the lofty superiority that they usually display” (Crick, 1995: 258).

PAG. 49

“They are elucidated by connective analysis, which traces, as far as is necessary for the purposes of clarification and for the solution or dissolution of the problems and puzzles at hand, the ramifying logico-grammatical web of connections between the problematic concept and adjacent ones” (Bennett & Hacker, 2003: 400).

PAG. 50

“Inasmuch as unclarity or confusion with respect to the concept of the mind and related psychological concepts will contaminate the description and understanding of the neuroscientific problems and solutions [...] Philosophy can contribute nothing to the scientific theories about the neural basis of consciousness [...] just as neuroscience can contribute nothing to the clarification of the concept of consciousness” (Bennett & Hacker, 2003: 403).

“The purpose of the present volume is to provide materials for reflection on the most recent philosophical revolution, that of linguistic philosophy. I shall mean by “linguistic philosophy” the view that philosophical problems are problems which may be solved (or dissolved) either by reforming language, or by understanding more about the language we presently use” (Rorty, 1967: 3).

PAG. 51

“For the philosopher, as an analyst, is not directly concerned with the physical properties of things. He is concerned with the way in which we speak about them. In other words, the propositions of philosophy are not factual, but linguistic in character – that is, they do not describe the behaviour of physical, or even mental, objects; they express definitions, or the formal consequences of definitions” (Ayer, 1946: 44).

“Only with Frege was the proper object of philosophy finally established: namely, first, that the goal of philosophy is the analysis of the structure of *thought*; secondly, that the study of *thought* is to be sharply distinguished from the study of psychological process of *thinking*; and, finally, that the only proper method for analysing thought consists in the analysis of *language*. [...] The acceptance of these three tenets is common to the entire analytical school” (Dummett, 1978: 458).

“Cruder versions which credit the brain with properties which can be meaningfully ascribed only to persons or their mental states, obliging that organ to become a locus of meaning and intentionality, a repository of rules, a home for grammatically characterized internal states, and so forth” (Rundle, 1997: 25).

PAG. 54

“The word ‘picture’ denotes, among other things, abstract paintings, representational paintings and films, i.e. motion pictures. Although abstract paintings and films each have something in common with representational paintings, there seem to be no relevant features common to a Jackson Pollock drip painting and a projected image of Humphrey Bogart, and yet we have no inclination – do we? – to say that the word has one meaning in a conversation about Casablanca and another when the talk turns to Lavender Mist. (...) If someone were to construct such a pattern of use from scratch, we might find it rather puzzling. But if we see it as a process of historical development, the puzzle dissolves. In the case of ‘picture’, the original focus on representational painting naturally licenses an extension of the term’s use to photographs and thence to motion pictures; and developments in painting also made natural a different extension of the term to include canvasses of a non-representational sort” (Mulhall, 2014: n/p).

PAG. 55

“I can think of no better expression to characterize these similarities than “family resemblances”; for the various resemblances between members of a family: build, features, colour of eyes, gait, temperament,

etc. etc. – overlap and criss-cross in the same way – And I shall say: 'games' form a family" (Wittgenstein, 1953: § 67).

PAG. 56

"Since in the normal course of scientific progress the meanings of words undergo change as a result of theory change, then when a hypothesis is advanced that would result in meaning change, this is not of itself an objection to the hypothesis [...] Hence, what to ancient ears may have sounded like an odd thing to say comes to seem entirely ordinary and correct. This is because the ancient ears are informed by ancient theory. [...] The assumption that a new theory's adequacy is compromised if its terms fail to preserve synonymy with the terms of the old theory is perniciously conservative. For it implies that preservation of the status quo should override considerations of empirical adequacy" (Churchland, 1986: 274).

PAG. 57

"Philosophy is allied with, and indeed continuous with, the physical sciences" (Dennett, 1987: 5).

PAG. 58

"An analysis of our ordinary way of speaking about pains shows that no events or processes could be discovered in the brain that would exhibit the characteristics of the putative 'mental phenomena' of pain, because talk of pains is essentially non-mechanical, and the events and processes of the brain are essentially mechanical" (Dennett, 1969: 91).

PAG. 59

"The systems reply suggests, entirely correctly in my opinion, that Searle has confused different levels of explanation (and attribution). / understand English; my brain doesn't - nor, more particularly, does the proper part of it (if such can be isolated) that operates to 'process' incoming sentences and to execute my speech act intentions" (Dennett, 1980: 429).

PAG. 60

"The distinction between the personal and sub-personal levels of explanations is not new at all. The philosophy of mind initiated by Ryle and Wittgenstein is in large measure an analysis of the concepts we use at the personal level [...] the personal and sub-personal levels must not be confused. The lesson has occasionally been misconstrued, however, as the lesson that the personal level of explanation is the *only* level of explanation when the subject matter is human minds and actions. In an important but narrow sense that is true, for as we see in the case of pain, to abandon the personal level is to stop talking about pain. In another important sense it is false, and it is this that is often missed. The recognition that there are two levels of explanation gives birth to the burden of relating them, and this is a task that is not outside the philosophers' province" (Dennett, 1969: 95-96).

“Nonsense is often generated when an expression is used contrary to the rules for its use. The expression in question may be an ordinary, non-technical expression, in which case the rules for its use can be elicited from its standard employment and received explanations of its meaning. Or it may be a technical term of art, in which case the rules for its use must be elicited from the theorist’s introduction of the term and the explanation he offers of its stipulated use” (Bennett & Hacker, 2003: 6)

PAG. 62

“No matter what any philosopher may say, examining the use of words is an empirical investigation, which often yields everyday garden-variety truths and objections and is subject to correction by standard observations and objections” (Dennett, 2007: 81).

PAG. 63

“If neurophysiologists, psychologists, artificial-intelligence scientists, or philosophers wish to change existing grammar, to introduce new ways of speaking, they may do so; but their new stipulations must be explained and conditions of application laid down. What may not be done is to argue that since we know what “to think”, “to see”, or “to infer” mean and know what “the brain thinks means, therefore we must know what “the brain thinks, sees, and infers” means. For we know what these verbs mean only in so far as we have mastered their existing use, which does not license applying them to the body or its parts, save derivatively. Nor may one cross the new “technical” use with the old ones, as, for example, neuroscientists typically do in their theorizing. For this produce a conflict of rules and hence incoherence in the neuroscientists’ use of these terms” (Hacker, 1990: 148-149).

PAG. 65

“If all this is to be taken at face value, it seems to show, first, that the grammatical remark that these predicates, in their literal use, are restricted to human beings and what behaves like human beings is either wrong *simpliciter* or displays “semantic inertia” that has been overtaken by the march of science, for machines actually do behave like human beings. Secondly, if it makes literal sense to attribute epistemic and even perceptual predicates to machines which are built to simulate certain human operations and to execute certain human tasks, it seems plausible to suppose that the human brain must have a similar abstract functional structure to that of the machine design. In which case, surely it must make sense to attribute the variety of psychological predicates to the human brain after all” (Hacker, 1990: 160-161).

PAG. 66

“The behavioral evidence for mental states with the existence of the mental states themselves [...] they confuse the criterial basis for the application of the mental concepts with the mental states themselves” (Searle, 2007: 102-103).

PAG. 66

“It is a fallacy to say that the conditions for the successful operation of the language game are conditions for the phenomena in question” (Searle, 2007: 105).

PAG. 67

“Their arguments against *brain as subject* does not carry over into *brain as locus*” (Searle, 2007: 111).

PAG. 68

“As soon as definite knowledge concerning any subject becomes possible, this subject ceases to be called philosophy, and becomes a separate science” (Russell, 1967: 90).

“In the history of human inquiry, philosophy has the place of the initial central sun, seminal and tumultuous: from time to time it throws off some portion of itself to take station as a science, a planet, cool and well regulated, progressing steadily towards a distant final state. This happened long ago at the birth of mathematics, and again at the birth of physics: only in the last century we have witnessed the same process once again [...] in the birth of the science of mathematical logic, through the joint labours of philosophers and mathematicians. Is it not possible that the next century may see the birth [...] of a true and comprehensive science of language? Then we shall have rid ourselves of one more part of philosophy (there will still be plenty left) in the only way we ever can get rid of philosophy, by kicking it upstairs” (Austin, 1979: 180).

PAG. 70

“In my experience, disputes about the nature of philosophy tend to be fruitless and usually just express preferences for different research projects” (Searle, 2007: 122).

PAG. 72

“The common assumption of philosophical exceptionalism is false. Even the distinction between the *a priori* and the *a posteriori* turns out to obscure underlying similarities. Although there are real methodological differences between philosophy and the other sciences, as actually practiced, they are less deep than is often supposed” (Williamson, 2007: 3).

PAG. 73

“Like mathematics, it is a science, but not a natural science. Its methods are not primarily experimental, though it can draw on the results of natural science. Like foundational mathematics, its methods are abductive as well as deductive. As in the natural sciences, much progress in philosophy consists in the construction of better models rather than in the discovery of new laws. We should not worry about whether philosophy is *a priori* or *a posteriori*, because the distinction is epistemologically superficial” (Williamson, 2019: 19).

CAPÍTULO 3

PAG. 75

“The application of neuroscientific discoveries to traditionally philosophical concerns” (Bickle, 2009: 4).

“The facts are the facts; linguistic analysis may tell you what normal speakers believe, but it cannot be a reliable guide to facts about the world, to the real nature of things” (Churchland, 1996: 290).

PAG. 80

“An object (or concept) is said to be reducible to one or more objects if all statements about it can be transformed into statements about these other objects” (Carnap, 1967: 1).

Esta passagem deriva do original alemão que reproduzimos:

“Ein Gegenstand (oder Begriff) heißt auf einen oder mehrere andere Gegenstände „zurückführbar“, wenn alle Aussage über ihn sich umformen lassen in Aussagen über diese anderen Gegenstände” (Carnap, 1967: 1).

PAG. 84

“It is arguably analytic that reduction must simplify; after all, reductions must reduce... On this score bridge laws of the form [Something is a T1 if and only if it is a B1] apparently are wanting in various ways. Since [Something is a T1 if and only if it is a B1] is supposed to be a contingent law, the concepts [T1] and [B1] remain distinct; hence bridge laws yield no conceptual simplification. Further, since we have only a contingent biconditional “iff” connecting properties [T1] and [B1], [T1] and [B1] remain distinct properties and there is no ontological simplification.... If we want ontological simplification out of our reductions, we must somehow find a way of enhancing bridge laws... into identities” (Kim, 1998: 96-97).

PAG. 87

“The role of a theory is not to give us more facts but to organize facts into a practically manageable system” (Oppenheim & Kemeny, 1956: 11).

PAG. 89

“However, to make such identifications a requirement is to hobble severely the theoretical development of neuroscience, for it implies that only a reduction on the retentive end of the spectrum will do, and that neither will any of the variously revisionary styles of reductive relations be acceptable nor will displacement be considered even possible” (Churchland, 1986: 284).

PAG. 93

“Folk psychology is commonsense psychology - the psychological lore in virtue of which we explain behavior as the outcome of beliefs, desires, perceptions, expectations, goals, sensations, and so forth. It is a theory whose generalizations connect mental states to other mental states [...] These homey generalizations are what provide the characterization of the mental states and processes” (Churchland, 1986: 299).

PAG. 94

"[O]nce our fictitious ancestor Jones, has developed the theory that overt verbal behavior is the expression of thoughts and taught his compatriots to make use of the theory in interpreting each other's behavior, it is but a short step to the use of this language in self-description. Thus, when Tom, watching Dick has behavioral evidence which warrants the use of the sentence (in the language of the theory) "Dick is thinking 'p'" ... Dick using the same behavioral evidence, can say, in the language of the theory, "I am thinking 'p'" ... And it now turns out - need it have? - that Dick can be trained to give reasonably reliable self-descriptions using the language of the theory, without having to observe his overt behavior. Jones brings this about, roughly, by applauding utterances by Dick of "I am thinking that p" when the behavioral evidence strongly supports the theoretical statement "Dick is thinking that p"; and by frowning on utterances of "I am thinking that p," when the evidence does not support this theoretical statement. Our ancestors begin to speak of the privileged access each of us has to his own thoughts. What began as a language with a purely theoretical use has gained a reporting role" (Sellars, 1956: 320).

PAG. 97

"It doesn't matter whether it is actually "stored in the head" as a theory [...] or whether it constitutes more than just a theory (i.e., plays roles other than generating predictions and explanations of behavior). All that matters [...] is that folk psychology can be treated as the theory that our commonsense ontology of the mental belongs with" (Bickle, 1998: 43).

PAG. 98

"The fact that it makes no sense at all to speak of mental events as occurring at some point within the body has the result that the identity theory cannot be true. This is because the corresponding physical events do occur at some point within the body, and if those physical are identical with mental events, then those mental events must occur at some point within the body. But those mental events do not occur at any point within the body because any statement to the effect that they occur here, or there, would be senseless. Hence, mental events cannot meet the conditions of coexistence in space, and therefore cannot be identical with physical events" (Shaffer: 1968: 48-49).

PAG. 99

"We were driven to functionalism (hence to the autonomy of psychology) by the suspicion that there are empirical generalizations about mental states that can't be formulated in the vocabulary of neurological or physical theories: neurology and physics don't, we supposed, provide projectible kind-predicates that subtend the domains of these generalizations" (Fodor, 1981: 25).

"Every interesting discussion, that is every discussion which leads to an advance of knowledge, terminates in a situation where some decisive change of meaning has occurred... A new theory of pains will not change the pains ... it will change the meaning of 'I am in pain'" (Feyerabend, 1962: 58).

PAG. 100

"By a variety of pragmatic and social considerations [...] The whim of the central investigators [...] the desire to preserve or to break with past habits, the related opportunities for publicizing the theory, cadging grants, and attracting disciples" (Churchland, 1986: 283).

PAG. 101

“It is neither natural nor simple to formalize in first-order logical (Suppes, 1956: 248) [...] and utterly impractical” (Suppes, 1956: 250).

“Instead of saying that certain equations are a model of subatomic or economic phenomena, we propose to say that the subatomic or economic phenomena are models of the theory represented by those equations” (Balzer, Moulines & Sneed, 1987: 2).

PAG. 103

“Our goal is to suggest how cognitive psychology may begin to converge with neurobiology to yield a new perspective in the study of learning (Kandel, 1984: 376) [...] we have tried to provide neuronal versions of the Rescorla and Wagner models of conditioning” (Kandel, 1984: 388).

“In terms of my theory of reduction, the empirical base sets of Rescorla and Wagner’s theory get directly related, although heterogeneously, to base sets and relations of Hawkins and Kandel’s account. The ontologies of the two theories are about the same things – the real-world systems in their intended empirical applications overlap – although they describe these things at different theoretical levels. What we have in the reduction of Rescorla and Wagners’s theory to Hawkins and Kandel’s theory is a smooth intertheoretic reduction [...]: a reduction that vindicates the ontology of the reduced theory by virtue of structuring its amorphous, functionally specified base elements into their underlying component parts and the dynamics of these lower-level kinds” (Bickle, 1998: 179).

PAG. 105

“Intervene neuronally and track the effects on behavior” (Bickle, 2006: 425).

“When this strategy is successful, the cellular or molecular events in specific neurons into which the experiments have intervenes... directly explain the behavioral data. These explanations set aside intervening explanations, including the psychological, the cognitive/information processing, even the cognitive-neuroscientific... These successes amount to reductions of mind to molecular pathways in neurons...” (Bickle, 2006: 426).

PAG. 106

“Strict cross-theoretic entity and property identities, no. Total elimination of the caloric-fluid/phlogiston variety, again no” (Bickle, 1998: 211).

PAG. 110

“Consider what the brain-state theorist has to do to make good his claims. He has to specify a physical-chemical state such that any organism (not just a mammal) is in pain if and only if (a) it possesses a brain of a suitable physical-chemical structure; and (b) its brain is in that physical-chemical state. This means that the physical-chemical state in question must be a possible state of a mammalian brain, a reptilian brain, a mollusc's brain [...] Even if such a state can be found, it must be nomologically certain that it will also be a state of the brain of any extraterrestrial life that may be found that will be capable of

feeling pain [...] Thus it is at least possible that parallel evolution, all over the universe, might always lead to one and the same physical 'correlate' of pain. But this is certainly an ambitious hypothesis" (Putnam, 1975: 436).

PAG. 112

"Hence, the new-wave account of intertheoretic reduction allows into the fold cases involving reduced theories multiply realized in the reducing theory. The reduction relation p can relate many token potential models of the reducing theory to a single potential model of the reduced theory. This [...] is an interesting knock against the applicability of multiple realizability arguments against new-wave reductionism" (Bickle, 1998: 115).

"Two volumes of a gas might have the same temperature, but the distributions of velocities of their constituent molecules will be quite different even while their mean value is the same. To be consistent, functionalists should again deny reductive success to statistical mechanics since, as they would put it, temperature of a gas is differently realized in the two cases" (Churchland, 1986: 357).

PAG. 113

"It is smooth intertheoretic reductions that motivate and sustain statements of cross-theoretic identity, not the other way around" (Churchland, 1985: 11).

"It sees the ontology the reductionism guiding its methodology [...] but [...] ontological conclusions depend on the outcome of theory reduction" (Bickle, 1998: 146).

PAG. 114

"Scientists still use physical optics and equilibrium thermodynamics more often than phlogiston chemistry" (Bickle, 1998: 144).

PAG. 115

"Precisely because the various forms of ontological and explanatory reductionism are metaphysical theses concerning the logical of existential attributions and the logical structure of explanation, they are not open to scientific confirmation or disconfirmation. If they are to be confirmed or confuted, then it will be by analytical argument" (Bennett & Hacker, 2003: 356-357).

PAG. 117

"And yet, there are philosophers who refuse to acknowledge scientific philosophy as a philosophy, who wish to incorporate its results into an introductory chapter of science and claim that there exists an independent philosophy, which has no concern with scientific research and has direct access to truth. Such claims, I think, reveal a lack of critical judgment. Those who do not see the errors of traditional philosophy do not want to renounce its methods or results and prefer to go on along a path which scientific philosophy has abandoned. They reserve the name of philosophy for their fallacious attempts at a superscientific knowledge and refuse to accept as philosophical a method of analysis designed after the

patterns of scientific inquiry. What is required for a scientific philosophy is a reorientation of philosophic desires" (*apud* Bickle, 2003: xviii).

Esta passagem deriva do original alemão que reproduzimos:

“Und trotzdem gibt es Philosophen, welche sich weigern, die wissenschaftliche Philosophie als Philosophie anzuerkennen und sie lieber in das Einleitungskapitel eines wissenschaftlichen Lehrbuchs verweisen möchten; Philosophen, welche immer noch behaupten, daß es eine unabhängige Philosophie gibt, die nichts mit wissenschaftlicher Forschung zu tun hat, sondern ihren eigenen Zugang zur Wahrheit besitzt. Solche Ansprüche offenbaren meiner Ansicht nach das Fehlen jeder Kritik. Diejenigen, welche die Fehler der traditionellen Philosophie nicht erkennen, wollen natürlich ihre Ergebnisse und Methoden nicht aufgeben und ziehen es vor, auf einem Pfad weiterzugehen, den die wissenschaftliche Philosophie längst verlassen hat. Sie reservieren den Namen Philosophie für ihre mit Fehlschlüssen durchsetzten Versuche, eine überwissenschaftliche Erkenntnis aufzufinden, und weigern sich, eine Methode als philosophisch zu bezeichnen, die sich die wissenschaftliche Forschung zum Muster angenommen hat. Wer einer wissenschaftlichen Philosophie gerecht werden will, muß seine philosophischen Wünsche und Ziele revidieren” (Reichenbach, 1968: 305).

“I am putting together the individual pieces that science provides to make explicit the ‘bigger picture’ that most scientists leave implicit. That's "philosophy" enough for me” (Bickle, 2003: xv).

PAG. 118

“As should be clear from the discussion in the text, I am not charging cognitive neuroscientists with literally asserting nonsense!” (Bickle, 2003: 160).

“Is continuous with science, and is even part of science [...] philosophy lies at the abstract and theoretical end of it” (Quine, 1969: 83).

“We are not doing natural science, nor yet natural history [...] our considerations could not be scientific ones” (Wittgenstein, 1958: § 109).

PAG. 119

“Quine and his followers suppose that all ‘conceptualization’ is theory-ridden. So as soon as language is used, theory is involved [...]. They suggest that there is no distinction between conceptual truths [...] and empirical truths [...] – and, by implication, no difference between norms of representation (which is what most conceptual truths are) and statements of fact. Hence, too, they think that philosophy is continuous with science, and a part of the same enterprise as (they ascribe to) science: viz. the construction of theories about ‘the world’. Nothing, in our view, could be further from the truth” (Bennett & Hacker, 2003: 380).

PAG. 120

“The suggestion that the brain remembers or has knowledge or uses linguistic symbols is sometimes pilloried as a mere conceptual error that consists in taking categories appropriate to one domain and applying them to a different, inappropriate domain. But one person's category error is another person's deep theory about the nature of the universe [...] The important thing for getting at the truth about brains is not whether in customary usage ordinary humans-in -the-street do say that persons remember but do

not say that brains remember; rather, it is whether we ought to say that brains remember – whether, given the empirical facts, it is a reasonable hypothesis that brains remember” (Churchland, 1986: 273).

“The recognition that it is within science itself, and not in some prior philosophy, that reality is to be identified and described” (Quine, 1981: 21).

PAG. 121

“[...] ‘naturalism’ is a cluster concept or a family resemblance term; that is, a concept defined by a medley of qualities, none of which are essential. Naturalism in philosophy may be distinguished by a variety of shared aims, beliefs, principles, assumptions, methodological procedures, attitudes and values” (Overgaard, Gilbert & Burwood, 2013: 55).

CAPÍTULO 4

PAG. 125

“Suppose you are sitting in a garden on a lovely spring day, admiring an apple tree in bloom. The narrow phenomenological description might include something about your shifts of attention, your tacit awareness of moving your eyes, the feelings of joy that accompany your looking and so on. The broad description adds something about the colours of the blossoms, how the tree stands out against a background of other things (the lawn, hedges, a garden shed and so on) and how it is presented as having sides and profiles not currently seen (the rear side of the trunk, minute details of the blossoms that can only be seen close at hand, etc.), and the aesthetic qualities the tree is experienced as having (it is beautiful, the very picture of spring and so on)” (Overgaard, Gilbert & Burwood, 2013: 73).

PAG. 126

“One must not confuse noema (correlate) and essence” (Husserl, 1980: 73).

Esta passagem deriva do original alemão que reproduzimos:

“Man darf nicht verwechseln Noema (Korrelat) und Wesen” (Husserl, 1986: 87).

“For Husserl, phenomenological reduction [...] is the method of leading phenomenological vision from the natural attitude of the human being whose life is involved in the world of things and persons back to the transcendental life of consciousness [...] For us phenomenological reduction means leading phenomenological vision back from the apprehension of a being [...] to the understanding of the being of this being” (Heidegger, 1982: 21).

Esta passagem deriva do original alemão que reproduzimos:

“Für Husserl ist die phänomenologische Reduktion [...] die Methode der Rückführung des phänomenologischen Blickes von der natürlichen Einstellung des in die Welt der Dinge und Personen hineinlebenden Menschen auf das transzendente Bewußtseinsleben [...] Für uns bedeutet die phänomenologische Reduktion die Rückführung des phänomenologischen Blickes von der wie immer bestimmten Erfassung des Seienden [...] auf das Verstehen des Seins dieses Seienden” (Heidegger, 1975: 29).

PAG. 127

“Every mental phenomenon is characterized by what the Scholastics of the Middle Ages called the intentional (or mental) inexistence of an object, and what we might call, though not wholly unambiguously, reference to a content, direction toward an object (which is not to be understood here as meaning a thing), or immanent objectivity. Every mental phenomenon includes something as object within itself, although they do not all do so in the same way. In presentation something is presented, in judgement something is affirmed or denied, in love loved, in hate hated, in desire desired and so on” (Brentano, 1995: 68).

Esta passagem deriva do original alemão que reproduzimos:

“Jedes psychische Phänomen ist durch das charakterisiert, was die Scholastiker des Mittelalters die intentionale (auch wohl mentale) Inexistenz eines Gegenstandes genannt haben, und was wir, obwohl mit nicht ganz unzweideutigen Ausdrücken, die Beziehung auf einen Inhalt, die Richtung auf ein Objekt (worunter hier nicht eine Realität zu verstehen ist) oder die immanente Gegenständlichkeit nennen würden. Jedes enthält etwas als Objekt in sich, obwohl nicht jedes in gleicher Weise. In der Vorstellung ist etwas vorgestellt, in dem Urteile ist etwas anerkannt oder verworfen, in der Liebe geliebt, in dem Hasse gehaßt, in dem Begehren begehrt usw”(Brentano, 1973: 124-125).

PAG. 128

“The usual conception of intentionality [...] misconstrues the structure of the self-directedness toward [...] an ego or subject is supposed, to whose so-called sphere intentional experiences are then supposed to belong [...] the Dasein, is essentially such that this being, so far as it is, is always already dwelling with the extant. The idea of a subject which has intentional experiences merely inside its own sphere and is not yet outside it but encapsulated within itself is an absurdity” (Heidegger, 1982: 63-64).

Esta passagem deriva do original alemão que reproduzimos:

“Die übliche Auffassung der Intentionalität [...] verkennt sie auch die Struktur des Sichrichtens-auf [...] ein Ich, ein Sub-jekt an und läßt dessen sogenannter Sphäre intentionale Erlebnisse zugehören [...] des Daseins, ist seinem Wesen nach so, daß dieses Seiende, sofern es ist, je schon bei einem Vorhandenen sich aufhält. Die Idee eines Subjekts, das nur in seiner Sphäre intentionale Erlebnisse hat und noch nicht draußen ist, sondern in seiner Kapsel eingeschlossen, ist ein Unbegriff” (Heidegger, 1975: 89-90).

“It is the body which ‘catches’ and ‘comprehends’ movement. The acquisition of a habit is indeed the grasping of a significance, but it is the motor grasping of a motor significance [...] it is the body which ‘understands’” (Merleau-Ponty, 1989: 143-144).

Esta passagem deriva do original francês que reproduzimos:

“C’est le corps qui ‘attrape’ et qui ‘comprend’ le mouvement. L’acquisition de l’habitude est bien la saisie d’une signification, mais c’est la saisie motrice d’une signification motrice [...] c’est le corps qui ‘comprend’” (Merleau-Ponty, 1945: 167, 169).

PAG. 129

“Every perceptual habit is still a motor habit” (Merleau-Ponty, 1989: 153).

Esta passagem deriva do original francês que reproduzimos:

“Toute habitude perceptive est encore une habitude mortice” (Merleau-Ponty, 1945: 178).

“In a way our problem is that we have been looking for consciousness where it isn't. We should look for it where it is. Consciousness is not something that happens inside us. It is something we do or make. Better: it is something we achieve. Consciousness is more like dancing than it is like digestion” (Noë: 2009: xii).

“Inside and outside are inseparable. The world is wholly inside and I am wholly outside myself” (1989: 407).

Esta passagem deriva do original francês que reproduzimos:

“L'intérieur et l'extérieur sont inséparables. Le monde est tout au dedans et je suis tout hors de moi” (Merleau-Ponty, 1945: 467).

PAG. 132

“To think of perception as something that happens only in the brain, is to ignore the contribution of embodiment to sensory “pre-processing” (Gallagher, 2008: 164).

“In the case of auditory sensation, for example, the shape and location of the ears determine directional information by amplifying or filtering specific inputs” (Gallagher, 2008: 164).

“To adapt the Heideggerian analysis to embodied cognition, one could say that the organism doesn't simply find itself deeply situated in an environment presented as one possibility rather than another. Rather, it is part of the very nature of human existence that being in the environment is one of its necessary, existential characteristics” (Gallagher, 2008: 165).

PAG. 133

“Cognition involves the manipulation of information-bearing elements within the system, part of which is “out there,” in the world—so it doesn't happen merely “in the head” (Gallagher, 2008: 169).

“The general tendency of human reasoners to lean heavily on environmental supports. Thus consider the use of pen and paper to perform long multiplication, the use of physical re-arrangements of letter tiles to prompt word recall in Scrabble, the use of instruments such as the nautical slide rule, and the general paraphernalia of language, books, diagrams, and culture. In all these cases the individual brain performs some operations, while others [i.e., other operations] are delegated to manipulations of external media” (Clark & Chalmers, 1998: 7).

PAG. 134

“Perception is not something that happens to us, or in us. It is something we do” (Noë, 2004: 1).

“A mode of exploration of the environment drawing on implicit understanding of sensorimotor regularities” (Noë, 2004: 29-30).

“To designate a quest to marry modern cognitive science and a disciplined approach to human experience, thus placing myself in the lineage of the continental tradition of phenomenology” (Varela, 1996: 330).

PAG. 136

“Ultimately the weight of a proof is its convincing nature, the immediacy of the evidence which is imposed on us, beyond the logical chains of symbolic reasoning. This is the nature of intuitive evidence” (Varela, 1996: 339).

PAG. 138

“To move forward in our understanding of consciousness, we need to give up the internal, neural microfocus [...] The locus of consciousness is the dynamic life of the whole, environmentally plugged-in person or animal. Indeed, it is only when we take up this holistic perspective on the active life of the person or animal that we can begin to make sense of the brain’s contribution to conscious experience” (Noë, 2009: xiii).

PAG. 140

“The findings to date already strongly suggest that visual awareness cannot be thought of as the end product of such a hierarchical series of processing stages. Instead it involves the entire visual pathway as well as the frontal parietal areas, which are involved in higher cognitive processing. The activity of a significant minority of neurons reflects what is consciously seen even in the lowest levels we looked at, V1 and V2; it is only the proportion of active neurons that increases at higher levels in the pathway” (Logothetis, 1999: 74).

PAG. 141

“Neurophenomenology combines the careful study of experience from within with investigations of the brain and behavior from without. It uses descriptions of direct experience to guide the study of the brain processes relevant to consciousness” (Thompson, 2015: 28).

PAG. 142

“The self is a process, not a thing or an entity. The self isn’t something outside experience, hidden either in the brain or in some immaterial realm. It is an experiential process that is subject to constant change (Thompson, 2015: xxxi).”

“Given that scientific models of the world are always distillations of our conscious experience as observers, it makes no sense to try to reduce consciousness to one or another of our scientific models, including our models of brain activity” (Thompson, 2015: 101).

PAG. 143

“There is no such thing as a self. Contrary to what most people believe, nobody has ever been or had a self [...] To the best of our current knowledge there is no thing, no indivisible entity, that is us, neither in the brain nor in some metaphysical realm beyond this world” (Metzinger, 2009: 1).

PAG. 144

“The illusion—or delusion—is taking the self to have an independent existence, like taking the mirror image to be really in the mirror. Notice that the image as such isn’t an illusion; it’s the taking of the image to exist in the mirror that’s the illusion. Similarly, it’s not the appearance of the self as such that’s the illusion; it’s taking the self to exist independently that’s the illusion” (Thompson, 2015: 365).

PAG. 146

“Logistically, these take time. Scheduling regular meetings for brain storming and teaching, as well as creating a glossary and living conceptual framework document are simple practices and tools for successful NP implementation. The outcome should not only be a productive experimental development collaboration, but also lead to synthesized results, not individualistic pieces in interpretation from each domain’s perspective. Instead the results are the whole or a big chunk of puzzle” (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 6).

“High standards of experimental design and execution to achieve variable control, reliability, generalizability, and replication of results” (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 6).

PAG. 147

“Interplay between the phenomenological project and some techniques of cognitive science can generate a unique integrative methodology. As such, neurophenomenology should take the best of the practices of cognitive science, while contributing its unique techniques that have not been a common part of the cognitive science toolbox” (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 7).

PAG. 148

“If the interviewer is working with Buddhist monks, she may not need to receive a great deal of training and may be able to simply tell the participant the focus of the study. Conversely, if the same interviewer is working with undergrads at any given university in the West, she may need to pull from a collection of tools and techniques to give the participant the capacity to access the thoughts and feelings experienced” (Bockelman, Reinerman-Jones & Gallagher, 2013: 7).

PAG. 149

“Traditional cognitive science often tries to force the participant into an objective mold of predesigned conceptual frameworks. In contrast, the interactions of the interview allow a switch of directions altogether and temporarily allows the researcher to participate in the participant’s subjectivity” (Bockelman, Reinerman-Jones & Shaun Gallagher, 2013: 8).

PAG. 152

“Instead of finding “extra ingredients” to account for how consciousness emerges from matter and brain, my proposal reframes the question to that of finding meaningful bridges between two irreducible phenomenal domains. In this specific sense neurophenomenology is a potential solution to the hard problem by casting in an entirely different light on what “hard” means [...] What is needed are precisely

the connecting structures provided by PhR since they are both immediately pertinent for experience (by their very nature) and at the same time sufficiently intersubjective to serve as constructive counterparts for external analysis” (Varela, 1996: 340-341).

PAG. 153

“Turing’s great contribution was to show us that Kant’s question could be recast as an engineering question. Turing showed us how we could trade in the first-person perspective of Descartes and Kant for the third-person perspective of the natural sciences and answer all the questions” (Dennett, 2001: 455).

PAG. 155

“Phenomenology has failed to find a single, settled method that everyone could agree upon” (Dennett, 1991: 44).

PAG. 156

“It is, on the contrary a basic human ability which operates constantly in daily life, and that has been widely discussed in studies of creativity. Think about mathematics: ultimately the weight of a proof is its convincing nature, the immediacy of the evidence which is imposed on us, beyond the logical chains of symbolic reasoning. This is the nature of intuitive evidence: born not of argument but from the establishment of a clarity that is fully convincing” (Varela, 1996: 339).

PAG. 157

“There is a clear methodological ground leading to a communal validation and shared knowledge” (Varela, 1999: 340).

“I find nothing here that reveals how the phenomenological reduction differs from introspection as such” (Bayne, 2004: 352).

PAG. 158

“But this is not a betrayal of science: it is a necessary extension and complement. Science and experience constrain and modify each other as in a dance. This is where the potential for transformation lies. It is also the key for the difficulties this position has found within the scientific community. It requires us to leave behind a certain image of how science is done, and to question a style of training in science which is part of the very fabric of our cultural identity” (Varela, 1996: 347).

“Researcher sought not to accept the elimination of first-personal processes and unique experiences because of the rigidities of scientific methods; rather science can expand its methods to include the phenomenological practices that capture, investigate, and explain what may be irreducible to strictly neural processes” (Bockelman, Reinerman-Jones & Gallagher, 2013: 4).

PAG. 159

“To understand consciousness – the fact that we think and feel and that a world shows up for us – we need to look at a larger system of which the brain is only one element. Consciousness is not something the brain achieves on its own. Consciousness requires the joint operation of brain, body and world. Indeed, consciousness is an achievement of the whole animal and its environmental context. I deny in short, that you are your brain [...] To have a mind [...] requires more than a brain. Brains don't have minds; people (and other animals) do” (Noë, 2009: 10).

PAG. 160

“The aims of philosophy and of science are not different: to achieve understanding of the problems that matter to us” (Noë, 2009: 6).

“Science and humanistic styles of thinking must engage each other” (Noë, 2009: xv).

“The idea that science and philosophy, or the humanities more generally, are separate spheres with their own standards and criteria is itself a bit of a questionable ideology, a relic of enthusiasm of an earlier modern age” (Noë, 2009: xv).

“Philosophy and science share a common aim: understanding. Science and philosophy must work together to advance toward understanding” (Noë, 2009: xv).

“The empirical questions must be guided by first-person evidence” (Varela, 1996: 344).

“The novelty of my proposal is that disciplined first-person accounts should be an integral element of the validation of a neurobiological proposal, and not merely coincidental or heuristic information” (Varela, 1996: 344).

PAG. 161

“Cartesian neuroscience has no evidence to support its basic assumption that conscious experience is exhaustive of neuronal phenomena” (Noë, 1999: 263).

PAG. 163

“Is a natural phenomenon and that the cognitive complexity of consciousness increases as a function of the increasing complexity of living beings” (Thompson, 2015: 103).

“physical” no longer means essentially nonmental [...] physical being and experiential being imply each other or derive from something that is neutral between them” (Thompson, 2015: 105).

PAG. 164

“The introduction of reciprocal causation might be perceived as a subtle expression of dualism” (Lutz, 2002: 150).

“The phenomenal might just be the physical described differently – under a different guise or mode of representation” (Kirchhoff & Hutto, 2016: 353).

CAPÍTULO 5

PAG. 168

“The concept of domain concerns the field or area of investigation and thus particular issues or questions as independent of and distinguished from the kind of methodological strategy by means of which these questions are investigated [...] domain concerns the subject matter whereas the concept of discipline refers more to the method that is applied to investigate that subject matter” (Northoff, 2014a: 13).

PAG. 170

“A standard principle, illustrated by the split-brain results, is that the definition of the phenomenon to be explained coevolves with experimental discoveries. In the early stages of the scientific attack on any problem, accurate definition of the phenomenon is hampered precisely because not enough is known to permit an accurate definition. A pragmatic strategy is to begin by studying those cases agreed to be obvious examples of the phenomenon. Powered by this agreement, provisional, rough characterization can leverage the science’s first stages, with refinements in the phenomenon’s definition emerging as the surrounding facts become clear” (Churchland, 2002b: 309).

PAG. 175

“Logical conditions refer to all possible i.e. logically conceivable worlds. They include both natural and non-natural worlds with only the former underlying our physical and biological laws. Natural conditions, in contrast, refer only to the natural world and thus to the respective physical and biological laws. Since logical conditions comprise both natural and non-natural worlds, they necessarily include natural conditions” (Northoff, 2004a: 95).

PAG. 176

“The mind-brain problem refers to logical conditions including both natural and non-natural worlds. In contrast, neuroscience refers to natural conditions including the natural world only. Direct application of and inference from empirical findings in neuroscience to the mind-brain problem may thus falsely equate non-natural worlds with the natural world. This, however, may lead to false conclusions since non-natural worlds include a wider range of conditions than the natural world. The neuroscientist therefore commits a “conditional fallacy” when he directly applies and infers from his empirical findings to the mind-brain problem” (Northoff, 2004a: 97).

“Logical conditions may not necessarily ‘match’ with natural conditions” (Northoff, 2004a: 97).

PAG. 182

“Due to inclusion of ontological/epistemological assumptions i.e. theoretical aspects, the meaning of the term ‘hypothesis’ is broadened in the case of ‘neurophilosophical hypothesis’, as compared to ‘empirical i.e. neuroscientific hypothesis’. This broadened meaning is reflected in the necessity of consideration of both types of falsification, ‘empirical and logical falsification’” (Northoff, 2004a: 114).

PAG. 183

“Carnap has recognized that he is able to preserve a double standard for ontological questions and scientific hypotheses only by assuming an absolute distinction between the analytic and the synthetic; and I need not say again that this is a distinction which I reject” (Quine, 1951: 43).

PAG. 186

“First-Person Neuroscience uses methods for the systematic examination and evaluation of mental states by themselves and their contents as experienced in first-person perspective and links them with data about neuronal states as obtained in third-person perspective” (Northoff, & Heinzl, 2006: 2).

PAG. 187

“Such a direct relation is possible since the emotional experience has been transformed via a visual analogue scale into numerical values that can be correlated with the values resulting from the fMRI measurement. It has to be considered that the numerical values are not identical with the emotional experience itself. In a next step one might then compare the fMRI results from both approaches, the categorical or third-person versus the parametric or first-person analysis. If both analyses reveal different regions, the neuronal correlates of third-person categorization of emotions can be separated from the first-person (parametric) experience. As such the neuronal correlates specifically underlying first-person experience and thus mental states can be revealed and separated from those associated with third-person observation” (Northoff & Heinzl, 2006: 4-5).

PAG. 191

“We can see that even in animals, there is a most basic subcortical fear-anxiety system. That very same system receives plenty of input from especially the own body – we are thus intimately connected with our body and that relationship can transform into fear and anxiety if discrepancies in the inputs, i.e., between intero- and extero-ceptive stimuli, occur. Fear and anxiety can then be seen as strategies of the brain to re-balance and ultimately ‘normalize’ its intero- and extero-ceptive input – the brain establishes relationship to the body, which is manifest in fear/anxiety. More deeply, it tells us that even an emotion as basic as fear is not isolated and locked-up in the brain or the organism. Instead, the data tell us that fear is intrinsically relational, something constituted neuro-vegetatively as based on the interoceptive input from the body. That, in turn, establishes relationship to the environment and the world – taken in a more general sense, emotions like fear/anxiety are thus about the balance between world, body, and brain and are thus genuinely relational and neuro-ecological” (Scalabrini et al., 2020: 47).

PAG. 193

“The threat of the coronavirus presents us with a double threat – collapse of our subjectivity as due to the risk of our body being infected and collapse of our inter-subjectivity as by social isolation of our self from others. The double threat of both subjectivity and inter-subjectivity provides a depth dimension to our emotions including our fears and anxieties resulting in existential fear. Such intimate connection of

fear, self, and existence, is due to the way our brain and its spontaneous activity are organized, that is, it stabilizes itself by aligning to body and world in very much the same way we stabilize our movements during dancing by aligning to the rhythm of the music. That very same alignment of our brain to body and world, i.e., temporospatial alignment is threatened if not disrupted by the corona virus crisis – in the same, one dances erratically when being detached from the music's rhythms, we become existentially threatened and erratic when the brain's and ultimately our self's temporo-spatial neuroecological alignment to the world are threatened by both the virus and the defense measures of physical and social distancing" (Scalabrini *et al.*, 2020: 53).

PAG. 194

"We feel sorry because we cry, angry because we strike, afraid because we tremble, and not that we cry, strike, or tremble, because we are sorry, angry, or fearful, as the case may be" (James, 1884: 190).

PAG. 195

"If the relationship between emotional feeling and environment is direct and therefore constitutional, i.e., non-instrumental, changes in the environment should be able to impact and constitute emotional feelings independently of the body's sensorimotor representation. The environment itself may then directly involved in constituting emotional feelings" (Northoff, 2008: 506).

PAG. 196

"It is the relationship between person and environment, the actual difference or balance between subjects's a and b rewards, that seemed to determine emotional feelings and neural activity. It is such constitutional, i.e., non-instrumental, dependence of emotional feelings on the environment and its relationship to the person" (Northoff, 2008: 506).

PAG. 198

"The connectivity pattern thus argues strongly in favour of the intero-exteroceptive relational concept of emotional feeling which seems to make isolated interoceptive processing and thus an interoceptive-based concept of emotional feeling rather unlikely" (Northoff, 2008: 512).

PAG. 201

"Only if there is no overlap between natural and logical conditions, characterization of brain and mind as different categories can be maintained. However, there is some overlap between natural and logical conditions with respect to brain and mind. The brain as a "physical brain", underlying natural conditions by itself, must be considered as a necessary natural condition for the possibility of generating logical conditions. We remain unable to perform philosophical reasoning and thus unable to account for logical conditions without our brain. While logical conditions may remain independent from the brain by themselves, the brain must at least be regarded as a necessary natural condition for the possibility of their creation. The brain may then serve as a 'bridge between natural and logical conditions' and thus as a 'window to the mind'. The distinction between natural and logical conditions with respect to brain and mind can therefore no longer be considered as mutually exclusive" (Northoff, 2004: 122).

PAG. 202

“Since ‘neurophilosophical hypothesis’ requires ‘empirical and logical consistency’, contingency and necessity may co-occur and co-exist so that they are no longer mutually exclusive. Introduction of traces of contingency does therefore not necessarily eliminate necessity completely – ‘neurophilosophical hypothesis’ cannot be regarded as necessarily invalid in ontological regard. Since the same remains true in the case of ‘a priori’ and ‘a posteriori’, ‘neurophilosophical hypothesis’ cannot be regarded as necessarily invalid in epistemic regard either” (Northoff, 2004: 122).

PAG. 203

“The development of ‘neurophilosophical hypotheses’ may be regarded as irrelevant and non-necessary for both philosophy and neuroscience. Since ‘neurophilosophical hypotheses’ are ‘crude and arbitrary mixtures’ between empirical hypotheses and theoretical assumptions, they remain unable to make significant contributions to either philosophy or neuroscience. If, however, ‘neurophilosophical hypothesis’ cannot contribute to either discipline, they must be regarded as irrelevant in general” (Northoff, 2004: 122).

PAG. 204

“‘Arbitrary’ implies that there are no rules and strategies that serve as guidance for the generation of ‘linkage’. There are however various principles i.e. the principles of transdisciplinary methodology which establish concrete strategies for ‘linkage’ between empirical hypothesis i.e. natural conditions and theoretical assumptions i.e. logical conditions. This is reflected best in the “principle of transdisciplinary circularity” which defines specific steps for their ‘linkage’. Accordingly, the ‘linkage’ between empirical hypotheses and theoretical assumptions may be regarded as ‘systematic’ rather than ‘arbitrary’” (Northoff, 2004: 123).

“Methodological strategies for the linkage between first- and third-person data about mental and neuronal states need to be more refined and developed in the future” (Northoff & Heinzel, 2006).

PAG. 205

“I see philosophy not as an a priori propaedeutic or groundwork for science, but as continuous with science. I see philosophy and science as in the same boat – a boat which, to revert to Neurath's figure as I so often do, we can rebuild only at sea while staying afloat in it [...] All scientific findings, all scientific conjectures that are at present plausible, are therefore in my view as welcome for use in philosophy as elsewhere” (Quine, 1969: 126).

PAG. 208

“My aim is to develop specific neurophenomenal hypotheses that show how the brain’s intrinsic features, that is, its resting-state activity and neural code, predispose the intrinsic features of consciousness and its phenomenal features in a necessary and unavoidable way” (Northoff, 2014d: xii).

PAG. 209

“Though the terms “neurophenomenological” and “neurophenomenal” seem to be very close, the two approaches must nevertheless be distinguished from each other. They make different presuppositions about both theoretical-conceptual background and the empirical approach to the brain [...] Empirically, the neurophenomenological approach claims for sensorimotor functions to be central in linking the brain to the environment and to thereby constitute subjectivity and consciousness. In contrast [...] my neurophenomenal approach does not consider any specific function whether affective, sensorimotor, or cognitive, to be a necessary and predisposing condition of consciousness. This is so because my neurophenomenal approach focuses on those neuronal processes and mechanisms that must precede the constitution and differentiation of these different functions. Metaphorically speaking, the neurophenomenal approach focuses on the very ground itself, the brain’s intrinsic activity, upon which different columns, the different functions, are erected. In contrast, the neurophenomenological approach focuses on one of these columns, the sensorimotor functions” (Northoff, 2014d: 538).

CAPÍTULO 6

PAG. 215

“There are recognizable qualitative characters of the given, which may be repeated in different experiences, and are thus a sort of universals; I call these "qualia." But although such qualia are universals, in the sense of being recognized from one to another experience, they must be distinguished from the properties of objects. Confusion of these two is characteristic of many historical conceptions, as well as of current essence-theories. The quale is directly intuited, given, and is not the subject of any possible error because it is purely subjective” (Lewis, 1929: 121).

PAG. 216

“It seems to me that there is not a clear consensus about how the term ‘qualia’ should be understood, and to this extent the contemporary problem of consciousness is not well-posed. The difficulty here can be vividly brought out at first by considering the fact that there seems to be a real dispute about whether qualia exist at all” (Crane, 2000: 2).

PAG. 217

“I think that there are certain features of the bodily sensations especially, but also of certain perceptual experiences, which no amount of purely physical information includes” (Jackson, 1982: 273).

PAG. 219

“Every conscious state has a certain qualitative feel to it, and you can see this clearly if you consider examples. The experience of tasting beer is very different from hearing Beethoven’s Ninth Symphony, and both of those have a different qualitative character from smelling a rose or seeing a sunset. These examples illustrate the different qualitative features of conscious experiences” (Searle, 2000: 560).

“Qualia are the simple sensory qualities to be found in the blueness of the sky or the tone of sound produced by a cello, and the fundamental components of the images in the movie metaphor are thus made of qualia” (Damásio, 1999: 9).

“Is perhaps the most daunting problem of consciousness [...] each differentiable conscious experience represents a different quale, whether it is primarily a sensation, an image, a thought, or even a mood” (Edelman & Tononi, 2000: 157).

“Although qualia are most obviously associated with sensations and perceptions, they are also found in other mental states, such as beliefs, desires, hopes and fears, during the conscious episodes of these states” (Glynn, 1999: 392).

PAG. 220

“For any conscious state [...] there is something that it qualitatively feels like to be in that state” (Searle, 1997: xiv).

PAG. 222

“But to find experience E to be F (pleasant, unpleasant, etc.) is not a mark of experience as such, or of those experiences had while conscious, or of experiences of which one is conscious, let alone of transitive consciousness in general (which encompass far more than experiences)” (Bennett & Hacker, 2003: 278).

PAG. 223

“It would be quite possible for everyone to have something different in his box” (Wittgenstein: 1953: §293).

PAG. 227

“Consider now the possibility of learning to describe, conceive, and introspectively apprehend the teeming intricacies of our inner lives within the conceptual framework of a matured neuroscience” (Churchland, 1985: 16).

PAG. 228

“Even a utopian knowledge of neuroscience must leave Mary hopelessly in the dark about the subjective qualitative nature of sensations not yet enjoyed” (Churchland, 1985: 25).

PAG. 229

“How can I assess what Mary will know and understand if she knows everything there is to know about the brain? Everything is a lot, and it means, in all likelihood, that Mary has a radically different and deeper understanding of the brain than anything barely conceivable in our wildest flights of fancy” (Churchland, 1986: 332).

PAG. 230

“If a student really understands the nature of pregnancy by learning all there is to know about the causal nature of pregnancy, no one would expect the student to become pregnant thereby” (Churchland, 2005b: 289).

“‘Deadbies’ are things composed of cells with membranes, nuclei with DNA, the usual organelles, and so forth. ‘Deadbies’ reproduce, digest, respire, metabolize, manufacture proteins, grow, and so forth, just as organisms on Earth do. Unlike us, however, ‘deadbies’ are not really alive. This is a logical possibility. So life is explanatorily independent of biology [...] imagine a planet where the velocity of molecules in a gas increases, but lo and behold, its temperature does not. Does this tell us that temperature is explanatorily independent of mean molecular kinetic energy? Certainly not. What does this tell us about the actual relation between mean molecule kinetic energy and temperature in a gas? Not a single thing” (Churchland, 2002: 178).

PAG. 231

“The mysteriousness of a problem is not a fact about the problem, it is not a metaphysical feature of the universe ó it is an epistemological fact about us. It is about where we are in current science, it is about what we can and cannot understand [...]. It is not a property of the problem itself” (Churchland, 1996: 406).

PAG. 232

“Jackson is expecting, quite wrongly, that one form of knowledge should constitute a quite different form of knowledge. He is expecting that explicit/discursive/scientific Knowledge should somehow constitute subjective knowledge of visual experiences. But that expectation is unreasonable on its face. As well expect that your exhaustive discursive knowledge of the micro-organization of a professional golfer’s motor cortex would constitute actual practical ‘knowledge,’ on your part, of how to hit a golf ball 200 yards down the middle of the fairway, even if you have never swung a golf club before in your life” (Churchland, 2011: 30).

PAG. 234

“The claim is not that colour (to take the standard example), as experienced, exists purely in structure in the world (which we must learn how to “enact,” i.e., behave appropriately towards). Nor is it that colour exists purely in our minds and not really in the world [...]. Rather [...] it is that colour exists in our interaction with the world” (Beaton, 2013: 310).

“The overall claim, then, is that terms such as “phenomenal mind” and “subjective experience” (and, indeed, terms such as “belief,” “desire,” “knowledge,” “emotion” – all personal-level mental terms) are not about brain dynamics; rather they are about the [...] active relation to the world. If we keep this normative dimension in mind, then [...] we may find qualia” (Beaton, 2013: 310).

PAG. 236

“Conscious experiences come in unified fields. In order to have a visual experience, a subject has to be conscious already and the experience is a modification of the field” (Searle, 2000: 572).

“Neurophenomenology is consistent with the unified field model. As Searle notes, according to the unified field approach, “what we have to look for is some massive activity of the brain capable of producing a unified holistic conscious experience”. One of the working hypotheses of neurophenomenology is that this massive activity is large-scale integration via phase synchrony (or more complex forms of generalized synchrony). This global self-organization is metastable (new patterns constantly arise without the system settling down into any one of them), rapid (occurring on a time scale of 100 to 300 milliseconds), and constitutes a fundamental pole of integration for the selection and modulation of local neuronal activity” (Thompson, Lutz & Cosmelli, 2005: 48).

PAG. 238

“The methodological moral to be drawn is that we should not pursue our investigations in the philosophy of mind by advancing a priori arguments about what science will and will not be able to tell us about our experience and the experiences of other sentient creatures. The only way we will ever know what science can tell us is by doing the science. Doing the science does not mean abandoning the philosophy, but it does mean that our philosophical explorations should not involve unconstrained thought experiments and unargued appeals to intuitions about mysterious properties. Knowledge of the qualitative features of conscious experience, whether scientific or otherwise, can never be had unless we begin from a more accurate and faithful portrayal than that found in Jackson's knowledge argument and many current discussions of qualia” (Thompson, 1992: 345).

PAG. 240

“To put it metaphorically, the level of spatial and temporal noise and fuzz in the resting state signifying rest–rest interaction seems to be significantly reduced in VS. The resting state is, metaphorically speaking, more silent in VS. There is only one lonely child silently playing in the room of the brain's resting state in VS as compared to the loud chatter of the 20 children playing in the room of the resting state in the healthy brain” (Northoff, 2014d: 422).

PAG. 242

“How is the ‘difference-based coding hypothesis of consciousness’ related to the ‘biophysical spectrum hypothesis of consciousness’? I proposed the degree of global metabolism and thus the energy supply to be central in determining the resting state's ‘position’ relative to the brain's biophysical-computational [...]. The more metabolism, the more likely the resting state's operation may be ‘located’ in the middle of the biophysical-computational spectrum. At the same time, the metabolism and the energy supply were supposed to be relevant in predisposing the degree of possible difference-based coding; higher degrees of energy supply are supposed to accompany higher degrees of difference-based coding and the consequent encoding of larger spatial and temporal differences into neural activity. And, as already said before, the larger the encoded differences, the higher the level of consciousness that can possibly be assigned to the newly resulting neural activity” (Northoff, 2014d: 434).

PAG. 243

“Due to the choice of the “right” starting points, the brain’s encoding strategy and the spatiotemporal structure of its intrinsic activity, the question of the “explanatory gap” cannot even be raised anymore” (Northoff, 2014d: 483).

PAG. 244

“This means that [...] qualia and consciousness are supposed to result from this statistically and spatiotemporal based “virtual” linkage between brain, body, and environment. Qualia are thus intrinsically relational; this implies that qualia would be impossible without such statistically and spatiotemporal based “virtual” linkage between brain, body, and environment” (Northoff, 2014d: 528).

PAG. 246

“My skepticism about a definitive analysis of information acknowledges the infamous versatility of information. The notion of information has been taken to characterize a measure of physical organization (or decrease in entropy), a pattern of communication between source and receiver, a form of control and feedback, the probability of a message being transmitted over a communication channel, the content of a cognitive state, the meaning of a linguistic form, or the reduction of an uncertainty. These concepts of information are defined in various theories such as physics, thermodynamics, communication theory, cybernetics, statistical information theory, psychology, inductive logic, and so on. There seems to be no unique idea of information upon which these various concepts converge and hence no proprietary theory of information” (Bogdan, 1994: 53).

PAG. 247

“If we want to understand the nature of the information processing that underlies such functions as thinking and sensorimotor control, our theories must be constrained by how neurons are in fact orchestrated, and we cannot understand that without knowing a good deal about neurons themselves, about their connections to other neurons, and about how they form these connections” (Churchland, 1986: 36).

PAG. 248

“The brain [...] doesn’t know anything [...] it has no capacity to make inferences about anything, let alone inferences about the remote environmental [...]. We – adult humans and other animals – think; we see; we feel; we infer” (Noë, 2009: 161).

“The level of consciousness of a physical system is related to the repertoire of causal states (information) available to the system as a whole (integration)” (Tononi & Koch, 2008: 253).

CAPÍTULO 7

PAG. 272

“[...] the fact that we probe the brain's neural activity by using our cognitive tasks does not entail or justify the assumption that the brain itself is a cognitive device” (Northoff, 2016c: 263).

PAG. 275

“We seem to be faced with the idea of an absurd duplication of the material world [...]” (Campbell, 2002: 156).

PAG. 276

“Radical embodied cognitive science [...] is the thesis that cognition is to be described in terms of agent-environment dynamics, and not in terms of computation and representation” (Chemero, 2009: x).

CAPÍTULO 8

PAG. 280

“Classical theories of sensory processing view the brain as a passive, stimulus-driven device. By contrast, more recent approaches emphasize the constructive nature of perception, viewing it as an active and highly selective process. Indeed, there is ample evidence that the processing of stimuli is controlled by top-down influences that strongly shape the intrinsic dynamics of thalamocortical networks and constantly create predictions about forthcoming sensory events” (Engel *et al.*, 2001: 74).

PAG. 282

“A level of blood oxygenation or blood sugar that is unexpected is likely to be bad news for an organism, whereas unexpected exteroceptive sensations (like novel visual inputs) are less likely to be harmful and may in some cases be desirable” (Seth, 2017: 9).

PAG. 288

“The interoceptive contents themselves do not yet entail the association of contents with consciousness—they can well remain as unconscious (as is most often the case in daily life) irrespective of whether they are (judged and detected and become aware as) accurate or inaccurate” (Northoff, 2018: 145).

“When we drive our car along familiar routes, much of our perception of the route including the landscape along it will remain an unconscious element. That changes once some unknown not yet encountered obstacle, such as a street blockade due to an accident, occurs. In such a case, we may suddenly perceive the houses along the route in a conscious way and become aware that, for instance, there are some beautiful mansions” (Northoff, 2018: 147).

REFERÊNCIAS

- Adriaans, P. (2020) "Information" in *Stanford Encyclopedia of Philosophy* (E. Zalta, ed.), Fall Edition, URL = <<https://plato.stanford.edu/archives/fall2020/entries/information/>>.
- Anderson, M. & Chemero, T. (2013) "The problem with brain GUTs: Conflation of different senses of 'prediction' threatens metaphysical disaster", *Behavioral and Brain Sciences*, 36 (3): 204-205.
- Anscombe, E. (1965) "The intentionality of sensation: A grammatical feature", in *Analytic Philosophy* (R. J. Butler, ed.), Oxford: Blackwell, pp. 158-180.
- Austin, J. (1979) *Philosophical Papers*, Oxford & New York: Oxford University Press.
- Ayer, A. (1946) *Language, Truth and Logic*, New York: Penguin Books.
- Balzer, W., Moulines, C. & Sneed, J. (1987) *An Architectonic for Science*, Dordrecht: D. Reidel.
- Bastos, A., Usrey, W., Adams, R., Mangun, G., Fries, P. & Friston, K. (2012) "Canonical microcircuits for predictive coding", *Neuron*, 76 (4): 695-711.
- Bayne, T. (2004) "Closing the gap? Some questions for neurophenomenology", *Phenomenology and the Cognitive Sciences*, 3: 349-364.
- Beaney, M. (2007) "The analytic turn in early twentieth-century philosophy" in *The Analytic Turn. Essays in Early Analytic Philosophy and Phenomenology* (M. Beaney, ed.), New York & London: Routledge, pp. 1-38.
- Beaton M. (2013) "Phenomenology and embodied action", *Constructivist Foundations*, 8 (3): 298-313.
- Bechara, A. (2004) "The role of emotion in decision-making: evidence from neurological patients with orbitofrontal damage", *Brain and Cognition*, 55 (1): 30-40.
- Bechtel, W. & Richardson, R. C. (1993) *Discovering Complexity: Decomposition and Localization as Scientific Research Strategies*, Princeton: Princeton University Press.

- Beer, J., John, O., Scabini, D., & Knight, R. (2006) "Orbitofrontal cortex and social behavior: Integrating self-monitoring and emotion-cognition interactions", *Journal of Cognitive Neuroscience*, 18: 871-879.
- Bennett, M. & Hacker, P. (2003) *Philosophical Foundations of Neuroscience*, Oxford: Blackwell.
- Bennett, M. & Hacker, P. (2013) *History of Cognitive Neuroscience*, Oxford: Wiley-Blackwell.
- Bickle, J. (1998) *Psychoneural Reduction: The New Wave*, Cambridge, MA: MIT Press.
- Bickle, J. (2003) *Philosophy and Neuroscience: A Ruthlessly Reductive Account*, Dordrecht: Kluwer Academic Publishers.
- Bickle, J. (2006) "Reducing mind to molecular pathways: explicating the reductionism implicit in current cellular and molecular neuroscience", *Synthese*, 151: 411-434.
- Bickle, J. (ed.) (2009) *The Oxford Handbook of Philosophy and Neuroscience*, New York: Oxford University Press.
- Bickle, J., Silva, A. & Landreth, A. (2014) *Engineering the Next Revolution in Neuroscience: The New Science of Experiment Planning*, New York: Oxford University Press.
- Blackmore, C. (1977) *Mechanics of the Mind*, Cambridge: Cambridge University Press.
- Blackmore, S. (2002) "There is no stream of consciousness", *Journal of Consciousness Studies*, 9 (5-6): 17-28.
- Block, N. (1990) "Inverted earth", in *Philosophical Perspectives* (J. Tomberlin, ed.) (Vol. 4), Atascadero, CA: Ridgeview Publishing Company, pp. 53-79.
- Block, N. (2005) "Two neural correlates of consciousness", *Trends in Cognitive Sciences*, 9 (2): 46-52.
- Bockelman, P., Morrow, P. & Fiore, S. (2013) "Team cognition: coordination across individuals and machines" in *The Oxford Handbook of Cognitive Engineering* (J. Lee & A. Kirlik, eds.), Oxford, MA: Oxford University Press, pp. 200-215.
- Bockelman, P., Reinerman-Jones, L. & Gallagher, S. (2013) "Methodological lessons in neurophenomenology: review of a baseline study and recommendations for research approaches", *Frontiers in Human Neuroscience*, 7: 608.

Bogdan, R. (1994) *Grounds for Cognition: How Goal-guided Behavior shapes the Mind*, Hillsdale, NJ: Lawrence Erlbaum.

Boly, M., Tshibanda, L., Vanhaudenhuyse, A., Noirhomme, Q., Schnakers, C., Ledoux, D., Boveroux, P., Garweg, C., Lambermont, B., Philips, C., Luxen, A., Moonen, G., Bassetti, C., Maquet, P. & Laureys, S. (2009) "Functional connectivity in the default network during resting state is preserved in a vegetative but not in a brain dead patient", *Human Brain Mapping*, 30 (8): 2393-2400.

Brentano, F. (1973) *Psychologie vom empirischen Standpunkt*, Hamburg: Felix Meiner Verlag.

Brentano, F. (1995) *Psychology from an Empirical Standpoint* (O. Kraus, ed.; A. Rancurello, D. Terrell & L. McAlister, trad.), London: Routledge.

Broad, C. (1925) *The Mind and its Place in Nature*, London: Routledge & Kegan.

Brodski, A., Paasch, G. F., Helbling, S. & Wibrat, M. (2015) "The faces of predictive coding", *The Journal of Neuroscience*, 35 (24): 8997-9006.

Brooks, R. (1991) "Intelligence without representation", *Artificial Intelligence*, 44: 139-160.

Brozek, B., Stelmach, J. & Kurk, L. (eds.) (2013) *Philosophy in Neuroscience: Three Modes of Interactions*, Kraków: Copernicus Center Press.

Campbell, J. (2002) *Reference and Consciousness*, Oxford: Oxford University Press.

Cannon-Bowers, J., Salas, E. & Converse, S. (1993) "Shared mental models in expert team decision making" in *Individual and Group Decision Making: Current Issues* (N. J. Castellan Jr., ed.), Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 221-246.

Capurro, R. & Hjørland, B. (2003) "The concept of information" in *Annual Review of Information Science and Technology* (B. Cronin, ed.), Medford, NJ: Information Today, pp. 343- 411.

Capurro, R. (1978) *Information*, Munich: Saur.

Carnap, R. (1928) *The Logical Structure of the World and Pseudoproblems in Philosophy* (R. George, trad.) Berkeley, CA: University of California Press.

Carnap, R. (1934) *The Unity of Science* (M. Black, trad.), London: Kegan Paul, Trench, Trubner, and Co.

- Cavell, S. (2002) *Must We Mean What We Say? A Book of Essays*, Cambridge: Cambridge University Press.
- Cerullo, M., Metzinger, T. & Mangun, G. (2015) "The problem with phi: A critique of integrated information theory", *PLoS Computational Biology*, 11 (9): e1004286.
- Cessario, R. (2003) *A Short History of Thomism*, Washington, D.C.: Catholic University of America Press.
- Chalmers, D. (1996) *The Conscious Mind*, New York, NY: Oxford University Press.
- Chalmers, D. (2000) "What is a neural correlate of consciousness?", in *Neural Correlates of Consciousness* (T. Metzinger, ed.), Cambridge, MA: MIT Press, pp. 18-39.
- Chemero, A. (2009) *Radical Embodied Cognitive Science*, Cambridge, MA: MIT Press.
- Child, W. (2013), *Wittgenstein* (R. Pich, trad.), Porto Alegre: Editora Penso.
- Churchland, P. M. (1985) "Reduction, qualia, and the direct introspection of brain states", *The Journal of Philosophy*, 82 (1): 8-28.
- Churchland, P. M. (1988) *Matter and Consciousness*, Cambridge, MA: MIT Press.
- Churchland, P. M. (1989) "Knowing qualia: A reply to Jackson", in *A Neurocomputational Perspective* (Y. Nagasawa, P. Ludlow & D. Stoljar, eds.), Cambridge, MA: MIT Press, pp. 163-178.
- Churchland, P. M. (1994) "Folk psychology", in *A Companion to the Philosophy of Mind* (S. Guttenplan, ed.), Oxford: Blackwell, pp. 308-316.
- Churchland, P. M. (2011) "Consciousness and the introspection of 'qualitative simples'", *Eidos*, 15: 12-45.
- Churchland, P. M. (2012) *Plato's Camera: How the Physical Brain Captures a Landscape of Abstract Universals*, Cambridge, MA: MIT Press.
- Churchland, P. S. (1986) *Neurophilosophy. Toward a Unified Science of the Mind/Brain*, Cambridge: MIT Press.

- Churchland, P. S. (1996a) "Do we propose to eliminate consciousness?", in *The Churchlands and Their Critics* (R. McCauley, ed.), New York: Blackwell, pp. 297-301.
- Churchland, P. S. (1996b) "The Hornswoggle problem", *Journal of Consciousness Studies*, 3 (5-6): 402-408.
- Churchland, P. S. (2002a) "Epistemology", in *Brain-Wise: Studies in Neurophilosophy*, Cambridge: MIT Press, pp. 239-271.
- Churchland, P. S. (2002b) "Self-representation in nervous systems", *Science*, 12 (296): 308-310.
- Churchland, P. S. (2002c) *Brain-Wise: Studies in Neurophilosophy*, Cambridge, MA: MIT Press.
- Churchland, P. S. (2005a) "Cleansing science", *Inquiry*, 48: 464-477.
- Churchland, P. S. (2005b) "A neurophilosophical slant on consciousness research", *Progress in Brain Research*, 149: 285-293.
- Clark, A. (2013a) "Whatever next? Predictive brains, situated agents, and the future of cognitive science", *Behavioral and Brain Science*, 36: 181-253.
- Clark, A. (2013b) "Are we predictive engines? Perils, prospects, and the puzzle of the porous perceiver", *Behavioral and Brain Sciences*, 36 (3): 233-253.
- Clark, A. (2016) *Surfing Uncertainty: Prediction, Action and the Embodied Mind*, Oxford: Oxford University Press.
- Clark, A. & Chalmers, D. (1998) "The extended mind", *Analysis*, 58 (1): 7-19.
- Clark, A. & Grush, R. (1999) "Towards a cognitive robotics", *Adaptive Behavior*, 7 (1): 5-16.
- Colombetti, G. (2005) "Appraising valence", *Journal of Consciousness Studies*, 12 (8-10): 103-126.
- Colombetti, G. & Thompson, E. (2005) "Enacting emotional interpretations with feeling", *Behavioral and Brain Sciences*, 28 (2): 200-201.

Colombetti, G. & Thompson, E. (2007) "The feeling body: Towards an enactive approach to emotion", in *Body in Mind, Mind in Body: Developmental Perspectives on Embodiment and Consciousness* (W. F. Overton, U. Müller, & J. Newman, eds.), Mahwah, NJ: Lawrence Erlbaum, pp. 45-68.

Cornman, J. (1968) "On the elimination of 'sensations' and sensations", *Review of Metaphysics*, XXII: 15-35.

Cosmelli, D., David, O., Lachaux, J.P., Martinerie, J., Garnero, L., Renault, B. & Varela, F. (2004) "Waves of consciousness: Ongoing cortical patterns during binocular rivalry", *NeuroImage*, 23 (1): 128-140.

Cosmelli, D. & Thompson, E. (2007) "Mountains and valleys: Binocular rivalry and the flow of experience", *Consciousness and Cognition*, 16: 623-641.

Craig, A. (2002) "How do you feel? Interoception: The sense of the physiological condition of the body", *Nature Review Neuroscience*, 3: 655-666.

Crane, T. (2000) "The origins of qualia", in *The History of the Mind-Body Problem* (T. Crane & S. Patterson, eds.), London: Routledge, pp. 169-194.

Crane, T. (2005) "What is the problem of perception?", *Synthesis Philosophica*, 20: 237-264.

Crane, T. (2015), "The mental states of persons and their brains", in *Mind, Self and Person, Royal Institute of Philosophy* (Anthony O'Hear, ed.), Supplement 76, Cambridge: Cambridge University Press, pp. 253-270.

Crick, F. (1995) *The Astonishing Hypothesis: The Scientific Search for the Soul*, New York, NY: Charles Scribner's Sons.

Crick, F. & Koch, C. (1998) "Consciousness and neuroscience", *Cerebral Cortex*, 8 (2): 97-107.

Crick, F. & Koch, C. (2003) "A framework for consciousness", *Nature Neuroscience*, 6 (2): 119-126.

Critchley, H. (2005) "Neural Mechanisms of autonomic, affective, and cognitive integration", *Journal of Comparative Neurology*, 493: 154-166.

Critchley, H., Wiems, S., Rothstein, P., Oehman, A. & Dolan, R. (2004) "Neural systems supporting interoceptive awareness", *Nature Neuroscience*, 7 (2): 189-195.

D'Oro, G. & Overgaard, S. (eds.) (2017) *The Cambridge Companion to Philosophical Methodology*, Cambridge: Cambridge University Press.

Damásio, A. (1996) *Descartes' Error – Emotion, Reason and the Human Brain*, London: Papermac.

Damásio, A. (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, San Diego, CA: Harcourt Brace.

Deleuze, G. & Guattari, F. (1994) *What is Philosophy?* (H. Tomlinson & G. Burchell III, trad.), New York, NY: Columbia University Press.

Dennett, D. (1969) *Content and Consciousness*, New York, NY: Routledge & Kegan Paul.

Dennett, D. (1980) "The milk of human intentionality (commentary on Searle)", *Behavioral and Brain Sciences*, 3: 428-430.

Dennett, D. (1986) "The self as a center of narrative gravity", *Philosophia*, 15: 275-88.

Dennett, D. (1987) *The Intentional Stance*, Cambridge: MIT Press.

Dennett, D. (1991) *Consciousness Explained*, Boston: Little, Brown, and Co.

Dennett, D. (2001) "The fantasy of first-person science", in *The Map and the Territory*, Berlin: Springer, pp. 455-473.

Dennett, D. (2007) "Philosophy as naïve anthropology", in *Neuroscience & Philosophy: Brain, Mind & Language*, New York, NY: Columbia University Press, 73-96.

de Ouden, H., Friston, K., Daw, N., McIntosh, A. & Stephan, K. (2009) "A dual role for prediction error in associative learning", *Cerebral Cortex*, 19 (5): 1175-1185.

Descartes, R. (1964) *Passion de L'âme*, Paris: Librairie Philosophique J. Vrin.

Descartes, R. (1978) *Princípios de Filosofia* (A. Ferreira, trad.), Lisboa: Guimarães & C.ª Editores.

Dummett, M. (1978) *Truth and Other Enigmas*, London: Duckworth.

Duncan, N., Hayes, D., Wiebking, C., Tiret, B., Pietruska, K., Chen, D., Rainville, P., Marjanska, M., Ayad, O., Doyon, J., Hodaie, M. & Northoff, G. (2015) "Negative childhood experiences alter a prefrontal-insular motor cortical network in healthy adults: a multimodal rsfMRI-fMRI-MRS-dMRI study", *Human Brain Mapping*, 36 (11): 4622-4637.

Eccles, J. (1979) *The Human Mystery*, Berlin: Springer-Verlag.

Edelman, G. (1994) *Bright Air, Brilliant Fire – On the Matter of the Mind*, Harmondsworth: Penguin.

Edelman, C. & Tononi, G. (2000) *Consciousness: How Matter Becomes Imagination*, London: The Penguin Press.

Enç, B. (1983) "In defense of the identity theory", *The Journal of Philosophy*, 80 (5): 279-298.

Engel, A., Fries, P. & Singer, W. (2001) "Dynamic predictions: Oscillations and synchrony in top-down processing", *Nature Review Neuroscience*, 2 (10): 704-716.

Engel, A. & Singer, W. (2001) "Temporal binding and the neural correlates of sensory awareness", *Trends in Cognitive Sciences*, 5 (1): 16-25.

Esfeld, M. & Lam, V. (2008) "Moderate structural realism about space-time", *Synthese*, 160: 27-46.

Esfeld, M. & Lam, V. (2010) "Holism and structural realism", in *Worldviews, science and us: Studies of analytical metaphysics. A selection of topics from a methodological perspective* (R. Vanderbeeken & B. D'Hooghe, eds.), Singapore: World Scientific, pp. 10-31.

Faivre, N. & Koch, C. (2014) "Temporal structure coding with and without awareness", *Cognition*, 131 (3): 404-414.

Feyerabend, P. (1962) "Explanation, reduction, and empiricism", in *Scientific Explanation, Space, and Time – Minnesota Studies in the Philosophy of Science* (H. Feigl & G. Maxwell, eds.), 3: 28-97.

Flanagan, O. (1996) *Self Expressions: Mind, Morals, and the Meaning of Life*, New York, NY: Oxford University Press.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. & Falk, A. (2007) "Social comparison affects reward-related brain activity in the human ventral striatum", *Science*, 318 (5854): 1305-1308.

Fodor, J. (1974) "Special sciences, or the disunity of science as a working hypothesis", *Synthese*, 28: 97-115.

Fodor, J. (1981) *Representations*, Cambridge, MA: MIT Press.

Fonseca, J. (2006) "Uma abordagem anti-realista contra a formulação do problema mente-corpo segundo uma perspectiva reducionista-interteórica" in *Análises: Actas do 2º Encontro Nacional de Filosofia Analítica* (S. Miguens, J. Pinto & C. Mauro, eds.), Porto: Faculdade de Letras da Universidade do Porto, pp. 124-133.

Foster, J. (2000) *The Nature of Perception*, Oxford: Oxford University Press.

Franklin, S. (1995) *Artificial Minds*, Cambridge, MA: MIT Press.

Friston, K. (2010) "The free-energy principle: A unified brain theory?", *Nature Reviews: Neuroscience*, 11: 127-138.

Gallagher, S. (2000) "Philosophical conceptions of the self: Implications for cognitive science", *Trends in Cognitive Sciences*, 4: 14–21.

Gallagher, S. (2005) *How the Body Shapes the Mind*, Oxford: Oxford University Press.

Gallagher, S. (2008) "Intersubjectivity in perception", *Continental Philosophy Review*, 41 (2): 163-178

Ganzetti, M. & Mantini, D. (2013) "Functional connectivity and oscillatory neuronal activity in the resting human brain", *Neuroscience*, 240: 297-309.

Gärdenfors, P. (1999) "Cognitive science: From computers to anthills as models of human thought", *Human ZT*, 3 (2): 243-260.

Gardner, H. (1983) *Frames of Mind: The Theory of Multiple Intelligences*, New York: Basic Books.

Gardner, H. (1985) *The Mind's New Science: A History of the Cognitive Revolution*, New York: Basic Books.

- Garfinkel, S., Seth, A., Barrett, A., Suzuki, K. & Critchley, H. (2015) "Knowing your own heart: Distinguishing interoceptive accuracy from interoceptive awareness", *Biological Psychology*, 104: 65–74.
- Giere, R. (1988) *Explaining Science*, Chicago: University of Chicago Press.
- Glock, H. (2003) *Quine and Davidson on Language, Thought and Reality*, Cambridge: Cambridge University Press.
- Glynn, I. (1999) *An Anatomy of Thought*, London: Weidenfeld and Nicolson.
- Gold, I. & Roskies, A. (2008) "Philosophy of neuroscience" in *The Oxford Handbook of Philosophy of Biology* (M. Ruse, ed.), New York, NY: Oxford University Press.
- Goldman, A. (1992) "In defense of the simulation theory", *Mind and Language*, 7: 104–119.
- Goldman, A. (1999) "A priori warrant and naturalistic epistemology" in *Philosophical Perspectives* (J. Timberlin, ed.), Cambridge, UK: Blackwell, pp. 1-28.
- Gonçalves, J. (2018) "Why are delusions pathological?" in *Schizophrenia and Common Sense Explaining the Relation Between Madness and Social Values* (I. Hipólito, J. Gonçalves & J. G. Pereira, eds.), Cham: Springer, pp. 163-174.
- Gopnik, A. & Wellman, H. (1992) "Why the child's theory of mind really is a theory", *Mind and Language*, 7: 145-171.
- Gordon, R. (1986) "Folk psychology as simulation", *Mind and Language*, 1: 158-171.
- Gordon, S. (ed.) (2013) *Neurophenomenology and its Applications to Psychology*, New York, NY: Springer.
- Gouveia, S. (2018) *Reflexões Filosóficas: Arte, Mente e Justiça* (J. F. Teixeira, pref.), Famacão: Editora Húmus.
- Gouveia, S. & Neiva, D. (2017) "The problem of consciousness on the mind uploading hypothesis", in *Philosophy of Mind: Contemporary Perspectives* (S. Gouveia & M. Curado, eds.), Newcastle: Cambridge Scholars Publishing, pp. 180-208.
- Greene, J., Nystrom, L., Engell, A., Darley, J. & Cohen, J. (2004) "The neural bases of cognitive conflict and control in moral judgment", *Neuron*, 44: 389-400.

- Greenwood, J. (1991) *The Future of Folk Psychology*, Cambridge: Cambridge University Press.
- Gregory, R. (1980) "Perceptions as hypotheses", *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 290 (1038): 181-197.
- Grupe, D. & Nitschke, J. (2013) "Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective", *Nature Reviews Neuroscience*, 14 (7): 488-501.
- Gutting, G. (2009) *What Philosophers Know: Case Studies in Recent Analytic Philosophy*, Cambridge: Cambridge University Press.
- Hacker, P. (1990) *Wittgenstein: Meaning and Mind*, 1st. ed., London: Blackwell.
- Hacker, P. (1992) "Private language argument", in *A Companion to Epistemology* (J. Dancy & E. Sosa, eds.), Oxford: Blackwell, pp. 368-374.
- Hacker, P. (2001) "An orrery of intentionality", *Language & Communication*, 21 (2): 119-141.
- Hacker, P. (2002) "Is there anything it is like to be a bat?", *Philosophy*, 77 (300): 157-174.
- Hacker, P. (2018) "Wittgenstein's legacy: The principles of the private language arguments", *Philosophical Investigations*, 41 (2): 123-140.
- Hacker, P. (2019) *Wittgenstein: Meaning and Mind*, 2nd. ed., London: Blackwell.
- Hacker, P., Bennett, M., Searle, J., Dennett, D. & Robinson, D. (2007) *Neuroscience and Philosophy: Brain, Mind and Language*, New York: Columbia University Press.
- Han, S. & Northoff, G. (2009) "Understanding the self: A cultural neuroscience approach", *Progress in Brain Research*, 178: 203-212.
- Hanson, S., Matsuka, T. & Haxby, J. (2004) "Combinatorial codes in ventral temporal lobe for object recognition: Haxby (2001) revisited: Is there a 'face' area?", *Neuroimage*, 23 (1): 156-166.
- Hardcastle, V. (1999) "What we don't know about brains", *Studies in History and Philosophy of Biological and Biomedical Sciences*, 5C: 69-89.

- Hardcastle, V. & Stewart, C. (2002) "What do brain data really show?", *Philosophy of Science*, 69: S72-82.
- Hawking, S. (2010) *The Grand Design*, London: Bantam Publishing.
- Haxby, J., Gobbini, M., Furey, M., Ishai, A., Schouten, J. & Pietrini, P. (2001), "Distributed and overlapping representations of faces and objects in ventral temporal cortex", *Science*, 293: 2425-2430.
- Heidegger, M. (1962) *Being and Time* (J. Macquarrie & E. Robinson, trad.), Oxford: Blackwell.
- Heidegger, M. (1975) *Die Grundprobleme der Phänomenologie*, Frankfurt am Main: Vittorio Klostermann.
- Heidegger, M. (1982) *The Basic Problems of Phenomenology* (A. Hofstadter, trad.), Bloomington, IN: Indiana University Press.
- Hertz, H. (1879) *Prinzipien der Mechanik*, Leipzig: Metzger & Wittig
- Hesselmann, G., Kell, C., Eger, E. & Kleinschmidt, A. (2008) "Spontaneous local variations in ongoing neural activity bias perceptual decisions", *Proceedings of the National Academy of Sciences of USA*, 105 (31): 10984-10989.
- Hohwy, J. (2013) *The Predictive Mind*, Oxford: Oxford University Press.
- Hohwy, J. (2014) "The self-evidencing brain", *Noûs*, 50 (2): 259-285.
- Hohwy, J. & Michael, J. (2017) "Why should any body have a self?", in *The Subject's Matter: Self-Consciousness and the Body* (F. de Vignemont & A. J. T. Alsmith, eds.), Cambridge, MA: MIT Press, pp. 363-390.
- Hooker, C. (1981) "Towards a general theory of reduction. Part I: Historical and scientific setting. Part II: Identity in reduction. Part III: Cross-Categorical reduction", *Dialogue*, 20: 38-59, 201-236, 496-529.
- Horgan, T. & Woodward, J. (1985) "Folk psychology is here to stay", *Philosophical Review*, 94: 197-226.
- Horn, B. (1980) "Derivation of invariant scene characteristics from images", *Proceedings of the National Computer Conference*, 371-376.

- Huang, Z., Wuhei, X., Weng, X., Laureys, S. & Northoff, G. (2013) "The self and its resting state in consciousness: An exploratory investigation in vegetative state", *Human Brain Mapping*, 35 (5): 1997-2008.
- Huang, Z., Zhang, J., Wu, J., Qin, P., Wu, X., Wang, Z., Dai, R., Li, Y., Liang, W., Mao, Y., Yang, Z., Zhang, J., Wolff, A. & Northoff, G. (2016) "Decoupled temporal variability and signal synchronization of spontaneous brain activity in loss of consciousness: an fMRI study in anesthesia", *Neuroimage*, 124 (Pt A): 693-703.
- Huaping, W. & Xiaoming, S. (2007) "Cooperative naturalism", *Frontiers of Philosophy in China*, 2: 601-613.
- Husserl, E. (1960) *Cartesian Meditations: An Introduction to Phenomenology* (D. Cairns, trad.), The Hague: Martinus Nijhoff.
- Husserl, E. (1969) *Formal and Transcendental Logic* (D. Cairns, trad.), The Hague: Martinus Nijhoff.
- Husserl, E. (1970) *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*, (D. Carr, trad.) Evanston, IL: Northwestern University.
- Husserl, E. (1973) *Cartesianische Meditationen und Pariser Vorträge*, The Hague: Martinus Nijhoff.
- Husserl, E. (1980) *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Book III* (T. Klein & W. Pohl, trad.), Dordrecht: Kluwer Academic Publishers.
- Husserl, E. (1983) *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Book I* (F. Kersten, trad.), Dordrecht: Kluwer Academic Publishers.
- Husserl, E. (1986) *Die Phänomenologie und die Fundamente der Wissenschaften*, Hamburg: Felix Meiner Verlag.
- Husserl, E. (2001) *Logical Investigations* (D. Moran, trad.), London: Routledge.
- Hutto, D. (2013) "Exorcising action-oriented representations: Ridding cognitive science of its Nazgûl", *Adaptive Behavior*, 21 (3): 142-150.

- Hutto D. & Myin E. (2012) *Radicalizing Enactivism: Basic Minds without Content*, Cambridge, MA: MIT Press.
- Jackson, F. (1982) "Epiphenomenal qualia", *Philosophical Quarterly*, 32: 127-136.
- James, T. & Gauthier, I. (2006) "Repetition-induced changes in BOLD response reflect accumulation of neural activity", *Human Brain Mapping*, 27 (1): 37-46.
- James, W. (1884) "What is an emotion?", *Mind*, 9: 188-205.
- Johnston, M. (2004) "The obscure object of hallucination", *Philosophical Studies*, 120 (1-3): 113-183.
- Johnston, M. (2007) "Objective mind and the objectivity of our minds", *Philosophy and Phenomenological Research*, 75 (2): 233-268.
- Joll, N. (2010) "Contemporary metaphilosophy", in *Internet Encyclopedia of Philosophy* (J. Fieser & B. Dowden, eds.), URL = <<http://www.iep.utm.edu/con-meta>>.
- Kandel, E. & Hawkins, R. (1984) "Is there a cell-biological alphabet for simple forms of learning?", *Psychological Review*, 91: 375-391.
- Kandel, E. & Schwartz, J. (eds.) (1981) *Principles of Neural Science*, New York: Elsevier/North-Holland.
- Kanwisher, N. (2001) "Neural events and perceptual awareness", *Cognition*, 79 (1-2): 89-113.
- Karpatschof, B. (2000) *Human Activity: Contributions to the Anthropological Sciences from a Perspective of Activity Theory*, Copenhagen: Dansk Psykologisk Forlag.
- Kelleey, B. (2009) "The early history of the quale and its relation to the senses" in *Routledge Companion to Philosophy of Psychology* (J. Symons & P. Calvo, eds.), London: Routledge, pp. 71-89.
- Kelley, W., Macrae, C., Wyland, C., Caglar, S., Inati, S. & Heatherton, T. (2002) "Finding the self? An event related fMRI study", *Journal of Cognitive Neuroscience*, 14: 785-794.
- Kenny, A. (2013) "Foreword", in *History of Cognitive Neuroscience*, Oxford: Wiley-Blackwell, pp. xvii-xix.
- Kerr, F. (ed.) (2002) *After Aquinas: Versions of Thomism*, Malden, MA: Blackwell.

Kiefer, A. (2017) "Literal perceptual inference", in *Philosophy and Predictive Processing* (T. Metzinger & J. M. Windt, eds.), Frankfurt am Main: MIND Group.

Kim, J. (1998) *Mind in a Physical World*, Cambridge, MA: MIT Press.

Kirchhoff, M. & Hutto, D. (2016) "Never mind the gap: Neurophenomenology, radical enactivism, and the hard problem of consciousness", *Constructivist Foundations*, 11 (2): 346-353.

Kitcher, P. (1984) "In defense of intentional psychology", *Journal of Philosophy*, 81: 89-106.

Klagge, J. (1989) "Wittgenstein and neuroscience", *Synthese*, 78 (3): 319-343.

Klein, C. (2014) "The brain at rest: What it is doing and why that matters", *Philosophy of Science*, 81 (5): 974-985.

Knobe, J. (2014) "Philosophers are doing something different now: Quantitative data", *Cognition*, 135: 36-38.

Koch, C., Massimini, M., Boly, M. & Tononi, G. (2016) "Neural correlates of consciousness: Progress and problems", *Nature Reviews: Neuroscience*, 17 (5): 307-321.

Koch, C. & Tsuchiya, N. (2012) "Attention and consciousness: Related yet different", *Trends in Cognitive Sciences*, 16 (2): 103-105.

Kok, P., Brouwer, G., van Gerven, M., & de Lange, F. (2013) "Prior expectations bias sensory representations in visual cortex", *Journal of Neuroscience*, 33 (41): 16275-16284.

Kornblith, H. (2009) "Review of *The Philosophy of Philosophy* by T. Williamson", *Analysis*, 69: 109-116.

Kornhuber, H. (1984) "Attention, readiness for action and the stages of voluntary decision", *Experimental Brain Research*, Supplement 9: 420-429.

Kripke, S. (1980) *Naming and Necessity*, Oxford: Blackwell.

Kuhn, T. (1970) *The Structure of Scientific Revolutions*, Chicago, IL: The University of Chicago Press.

Ladyman, J., Ross, D., Collier, J. & Spurrett, D. (2007) *Everything Must Go: Metaphysics Naturalized*, Oxford: Oxford University Press.

- Lamme, V. (2010) "How neuroscience will change our view on consciousness", *Cognitive Neuroscience*, 1 (3): 204-220.
- Laureys, S. & Schiff, N. (2012) "Coma and consciousness: Paradigms (re)framed by neuroimaging", *NeuroImage*, 61 (2): 478-491.
- Lazerowitz, M. (1970) "A note on 'metaphilosophy'", *Metaphilosophy*, 1 (1): 91.
- Letheby, C. & Gerrans, P. (2017) "Self unbound: ego dissolution in psychedelic experience", *Neuroscience of Consciousness*, (1): 1-11.
- Levine, J. (1983) "Materialism and qualia: The explanatory gap", *Pacific Philosophical Quarterly*, 64: 354-361.
- Levinstein, B. (2007) "Facts, interpretation and truth in fiction", *The British Journal of Aesthetics*, 47 (1): 64-75.
- Lewis, C. (1929) *Mind and the world-order: Outline of a Theory of Knowledge*, New York, NY: Dover Publications.
- Lewis, D. (1969) "Art, mind, and religion", *Journal of Philosophy*, 66 (1): 22-27.
- Lewis, D. (1978) "Truth in fiction", *American Philosophical Quarterly*, 15 (1): 37-46.
- Libet, B. (1993a) "Unconscious cerebral initiative and the role of conscious will in voluntary action" in *Neurophysiology of Consciousness*, Boston, MA: Birkhäuser, pp. 269-306.
- Libet, B. (1993b) "Epilogue: I. Some implications of 'time-on theory'" in *Neurophysiology of Consciousness*, Boston, MA: Birkhäuser, pp. 385-392.
- Livingstone, M. (2002) *Vision and Art: The Biology of Seeing*, New York, NY: Abrams.
- Lloyd, D. (2000) "Terra cognita: From functional neuroimaging to the map of the mind", *Brain and Mind*, 1: 93-116.
- Loemker, L. (1969) (ed. & trans.) *G. W. Leibniz: Philosophical Papers and Letters*, 2nd ed., Dordrecht: D. Reidel.

- Logothetis, N. (1999) "Vision: a window on consciousness", *Scientific American*, 281: 68-75.
- Logothetis, N., Leopold, D. & Sheinberg, D. (1996) "What is rivaling during binocular rivalry?", *Nature*, 380: 621-624.
- Logothetis, N., Murayama, Y., Augath, M., Steffen, T., Werner, J. & Oeltermann, A. (2009) "How not to study spontaneous activity", *Neuroimage*, 45 (4): 1080-1089.
- Lutz, A. (2002) "Toward a neurophenomenology as an account of generative passages: A first empirical case study", *Phenomenology and the Cognitive Sciences*, 1: 133-167.
- Lutz, A. & Thompson, E. (2003) "Neurophenomenology: Integrating subjective experience and brain dynamics in the neuroscience of consciousness", *Journal of Consciousness Studies*, 10 (9-10): 31-52.
- Machamer, P. & Systma, J. (2009) "Philosophy and the brain sciences", *Iris: European Journal of Philosophy and Public Discourse*, 1: 65-86.
- Mandik, P., Betchel, W., Mundale, J. & Stufflebeam, R. (2001) (eds.) *Philosophy and the Neurosciences*, Oxford: Blackwell.
- Markus, H. & Kitayama, S. (1991) "Culture and the self: Implication for cognition, emotion and motivation", *Psychological Review*, 98: 224-253.
- Martin, M. (1997) "The reality of appearances" in *Thought and Ontology* (M. Sainsbury, ed.), Milan: Franco Angeli, pp. 173-214.
- McCauley, R. (2001) "Explanatory pluralism and co-evolution of theories of sciences", in *Philosophy and the Neurosciences* (W. Bechtel, P. Mandik, J. Mundale & R. Stufflebean, eds.), Oxford: Basil Blackwell, pp. 17-47.
- McCloskey, M. (1983) "Intuitive physics", *Scientific American*, 248 (4): 122-130.
- McCulloch, G. (1995) *The Mind and Its World*, London: Routledge.
- McGinn, C. (2002) *The Making of a Philosopher*, New York, NY: HarperCollins.
- McGinn, C. (2004) *Consciousness and Its Objects*, Oxford: Oxford University Press.

- Merleau-Ponty, M. (1945) *Phénoménologie de la Perception*, Paris: Éditions Gallimard.
- Merleau-Ponty, M. (1968) *The Visible and the Invisible* (A. Lingis, trad.), Evanston, IL: Northwestern University Press.
- Merleau-Ponty, M. (1989) *Phenomenology of Perception* (C. Smith, trad.), London: Routledge.
- Metzinger, T. (2009) *The Ego Tunnel: The Science of the Mind and the Myth of the Self*, New York: Basic Books.
- Metzinger, T. (ed.) (2000) *Neural Correlates of Consciousness*, Cambridge, MA: MIT Press.
- Mitchell, J., Banaji, M. & Macrae, C. (2005a) "The link between social cognition and self-referential thought in the medial prefrontal cortex", *Journal of Cognitive Neuroscience*, 17: 1306-1315.
- Mitchell, J., Macrae, C. & Banaji, M. (2005b) "Forming impressions of people versus inanimate objects: Social-cognitive processing in the medial prefrontal cortex", *NeuroImage*, 26: 251-257.
- Mitchell, J., Macrae, C. & Banaji, M. (2006) "Dissociable medial prefrontal contributions to judgments of similar and dissimilar others", *Neuron*, 50: 655-663.
- Moore, G. (1953) *Some Main Problems of Philosophy*, London: Allen and Unwin.
- Moore, G. (1991) *Principia Ethica*, Cambridge: Cambridge University Press.
- Mormann, T. (1991) "Husserl's philosophy of science and the semantic approach", *Philosophy of Science*, 58 (1): 61-83.
- Mulhall, S. (2014) *Stanton Lectures*, University of Cambridge (não publicado), acesso via podcast, URL = <<https://sms.cam.ac.uk/collection/1637674>>.
- Mundale, J. (2002) "Concepts of localization: Balkanization in the brain", *Brain and Mind*, 3: 313-330.
- Nachev, P. & Hacker, P. (2014) "The neural antecedents to voluntary action: A conceptual analysis", *Cognitive Neuroscience*, 5 (3-4): 193-208.
- Nagel, E. (1961) *The Structure of Science: Problems in the Logic of Scientific Explanation*, New York: Harcourt, Brace, and World.

- Nagel, T. (1974) "What Is It Like to Be a Bat?", *The Philosophical Review*, 83 (4): 435-450.
- Nagel, T. (1979) *Mortal Questions*, Cambridge: Cambridge University Press.
- Nagel, T. (2012) *Mind and Cosmos*, New York, NY: Oxford University Press.
- Neurath, O. (1983) *Philosophical Papers, 1913-1946*, Dordrecht: Reidel.
- Neurath, O., Carnap, R. & Hahn, H. (1996) "The scientific conception of the world: the Vienna Circle" in *The Emergence of Logical Empiricism: from 1900 to the Vienna Circle* (S. Sarkar, ed.), New York: Garland Publishing.
- Ney, A. (2008) "Reductionism", *Internet Encyclopedia of Philosophy* (J. Fieser & B. Dowden, eds.), URL = <<https://iep.utm.edu/red-ism/>>.
- Niedenthal, P. (2007) "Embodying emotion", *Science*, 316 (5827): 1002-1005.
- Ning, Z. (2019) "Representing the world with Occam's razor: an informational teleosemantics on the basis predictive processing paradigm" in *Automata's Inner Movie: Science and Philosophy of Mind* (S. Gouveia & M. Curado, eds.), Wilmington, DL: Vernon Press, pp. 81-102.
- Nisbett, R. & Wilson, T. (1977) "Telling more than we can know: Verbal reports on mental processes", *The Psychological Review*, 84 (3): 231-258.
- Noë, A. & Thompson, E. (2004) "Are there neural correlates of consciousness?", *Journal of Consciousness Studies*, 11: 3-28.
- Noë, A. (1999) "Thought and experience", *American Philosophical Quarterly*, 36 (3): 257-265.
- Noë, A. (2004) *Action in Perception*, Cambridge, MA: MIT Press.
- Noë, A. (2009) *Out of our Heads: Why you are not your Brain, and other Lessons from the Biology of Consciousness*, New York: Hill and Wang.
- Northoff, G. (2001) *Personale Identität und operative Eingriffe in das Gehirn*, Paderborn: Mentis Publisher.
- Northoff, G. (2004a) "What is neurophilosophy? A methodological account", *Journal for General Philosophy of Science*, 35: 91-127.

Northoff, G. (2004b) *Philosophy of the Brain: The Brain Problem*, Amsterdam: John Benjamins Publishing.

Northoff, G. (2008) "Are our emotional feelings relational? A neurophilosophical investigation of the James–Lange theory", *Phenomenology and Cognitive Science*, 7 (4): 501-527.

Northoff, G. (2012) "Psychoanalysis and the brain – why did Freud abandon neuroscience?", *Frontiers of Psychology*, 3: 71.

Northoff, G. (2014a) *Minding the Brain: A Guide to Philosophy & Neuroscience*, New York: Red Globe Press.

Northoff, G. (2014b) "How is our self altered in psychiatric disorders? A neurophenomenal approach to psychopathological symptoms", *Psychopathology*, 47 (6): 365-376.

Northoff, G. (2014c) *Unlocking the Brain: Coding*, Vol. I, Oxford: Oxford University Press.

Northoff, G. (2014d) *Unlocking the Brain: Consciousness*, Vol. II, Oxford: Oxford University Press.

Northoff, G. (2014e) "Are auditory hallucinations related to the brain's resting state activity? A 'neurophenomenal resting state hypothesis'", *Clinical Psychopharmacology and Neuroscience*, 12 (3): 189-195.

Northoff, G. (2015) "Is schizophrenia a spatiotemporal disorder of the brain's resting state?", *World Psychiatry*, 14 (1): 34-35.

Northoff G. (2016a) "Is the self a higher-order of fundamental function of the brain? The "basis model of self-specificity" and its encoding by the brain's spontaneous activity", *Journal of Cognitive Neuroscience*, 7 (1-4): 203-222.

Northoff, G. (2016b) "Neuroscience and Whitehead I: Neuro-ecological model of brain", *Axiomathes*, 26 (3): 219-252.

Northoff, G. (2016c) "Neuroscience and Whitehead II: Process-Based ontology of brain", *Axiomathes*, 26 (3): 253–277.

Northoff, G. (2018) *The Spontaneous Brain: From the Mind–Body to the World–Brain Problem*, Cambridge, MA: MIT Press.

Northoff, G. (no prelo) “Is perception neuro temporal and neuro-ecological? – Whitehead’s non-traditional theory of perception” in *Consciousness: bridging the gap between science and philosophy* (A. Gomez-Marín & R. Hemsell, eds.), Anoka, MN: Process Century Press, s/p.

Northoff, G. & Bermpohl, F. (2004) “Cortical midline structures and the self”, *Trends in Cognitive Science*, 8: 102-107.

Northoff, G. & Heinzl, A. (2006) “First-Person neuroscience: A new methodological approach for linking mental and neuronal states”, *Philosophy, Ethics, and Humanities in Medicine*, 1 (3): E3.

Northoff, G., Heinzl, A., de Greck, M., Bermpohl, F., Dobrowolny, H. & Panksepp, J. (2006) “Self-referential processing in our brain – A meta-analysis of imaging studies on the self”, *NeuroImage*, 31: 440-457.

Northoff, G. & Huang, Z. (2017) “How do the brain’s time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC)”, *Neuroscience and Biobehavioral Reviews*, 80: 630-645.

Northoff, G. & Quin, P. (2011) “How can the brain's resting state activity generate hallucinations? A ‘resting state hypothesis’ of auditory verbal hallucinations”, *Schizophrenia Research*, 127: 202-214.

Northoff, G., Quin, P. & Nakao, T. (2010) “Rest-stimulus interaction in the brain: a review”, *Trends in Neuroscience*, 33 (6): 277-284.

Northoff, G., Wainio-Theberge, S. & Evers, C. (2019) “Is temporo-spatial dynamics the “common currency” of brain and mind? In quest of spatiotemporal neuroscience”, *Physics Life Reviews*, 33: 34-54.

Ochsner, K., Beer, J., Robertson, E., Cooper, J., Gabrieli, J., Kihlstrom, J. & D’Esposito, M. (2005) “The neural correlates of direct and reflected self-knowledge”, *NeuroImage*, 28: 797-814.

Oppenheim, P. & Kemeny, J. (1956) “On reduction”, *Philosophical Studies*, 7: 6-19.

O’Regan, K. & Noë, A. (2001) “A sensorimotor account of vision and visual consciousness”, *Behavioral and Brain Sciences*, 24 (5): 883-917.

- Orenstein, A. (2002) *W. V. Quine*, Chesham, UK: Acumen.
- Orlandi, N. (2010) "Are sensory properties represented in perceptual experience?", *Philosophical Psychology*, 23 (6): 721-740.
- Orlandi, N. (2012) "Embedded seeing-as: Multi-stable visual perception without interpretation", *Philosophical Psychology*, 25 (4):1-19.
- Orlandi, N. (2016) "Bayesian perception Is ecological perception", *Philosophical Topics*, 44 (2): 327-351.
- Overgaard, S., Gilbert, P. & Burwood, S. (2013) *An Introduction to Metaphilosophy*, Cambridge: Cambridge University Press.
- Panksepp, J. (2003) "At the interface of the affective, behavioral, and cognitive neurosciences: Decoding the emotional feelings of the brain", *Brain and Cognition*, 52: 4-14.
- Papineau, D. (2009) "The poverty of analysis", *Proceedings of the Aristotelian Society Supplementary*, 83: 1-30.
- Parfit, D. (1989) *Reasons and Persons, Vol. 2*, Oxford: Oxford University Press.
- Paulus, M., Feinstein, J. & Khalsa, S. (2019) "An active inference approach to interoceptive psychopathology", *Annual Review of Clinical Psychology*, 15 (1): 97-122.
- Peacocke, C. (1983) *Sense and Content*, Oxford: Oxford University Press.
- Peirce, C. (1982) "Lowell Lecture" (ix) in *Writings of Charles S. Peirce: A Chronological Edition* (M. H. Fisch, ed.), Indiana: Indiana University Press
- Penfield, W. (1975) *The Mystery of the Mind: a Critical Study of Consciousness and the Human Brain*, Princeton, NJ: Princeton University Press.
- Petitmengin, C. (2006) "Describing one's subjective experience in the second person: an interview method for the science of consciousness", *Phenomenology and the Cognitive Sciences*, 5: 229-269.
- Petitot, J., Varela, F., Pachoud, B. & Roy, J.M. (1999) *Naturalizing Phenomenology: Issues in Contemporary Phenomenology and Cognitive Science*, Stanford, CA: Stanford University Press.

- Place, U. (1956) "Is consciousness a brain process?", *British Journal of Psychology*, 47 (1): 44-50.
- Platek, S., Keenan, J., Gallup, G. & Mohamed, F. (2004) "Where am I? The neurological correlates of self and other", *Cognitive Brain Research*, 19: 114-122.
- Pollatos, O., Herbert, B., Matthias, E. & Schandry, R. (2007) "Heart rate response after emotional picture presentation is modulated by interoceptive awareness", *International Journal of Psychophysiology*, 63 (1): 117-124.
- Pollatos, O., Kirsch, W. & Schandry, R. (2005) "On the relationship between interoceptive awareness, emotional experience and brain processes", *Brain Research Cognitive Brain Research*, 25 (3): 948-962.
- Pollatos, O., Traut-Mattausch, E., Schroeder, H. & Schandry, R. (2007) "Interoceptive awareness mediates the relationship between anxiety and the intensity of unpleasant feelings", *Journal of Anxiety Disorders*, 21 (7): 931-943.
- Popper, K. (1959) *The Logic of Discovery*, London: Hutchinson.
- Popper, K. (1968) *Conjectures and Refutations: The Growth of Scientific Knowledge*, New York: Harper & Row.
- Popper, K. (1975) "How I see philosophy" in *The Owl of Minerva: Philosophers on Philosophy* (C. Bontempo & S. Odell, eds.), New York: McGraw-Hill.
- Popper, K. & Eccles, J. (1977) *The Self and Its Brain*, Berlin: Springer Verlag.
- Pöyhönen, S. (2014) "Intentional concepts in cognitive neuroscience", *Philosophical Explorations*, 17 (1): 93-109.
- Praetorius, N. (2000) *Principles of Cognition, Language, and Action*, Dordrecht: Kluwer Academic Publisher.
- Putnam, H. (1975) "The nature of mental states", *Mind, Language, and Reality: Philosophical Papers, Vol. II.*, Cambridge: Cambridge University Press.
- Putnam, H. (1976) "'Two dogmas' revisited" in *Contemporary Aspects of Philosophy* (G. Ryle, ed.), Stocksfield: Oriel Press.

- Quin, P., Duncan, N., Chen, D., Chen, C., Huang, L., Huang, Z., Lin, C., Wiebking, C., Yang, C., Northoff, G. & Lane, T. (2018) "Vascular-metabolic and GABAergic inhibitory correlates of neural variability modulation. A combined fMRI and PET study", *Neuroscience*, 379: 142-151.
- Quin, P. & Northoff, G. (2011) "How is our self related to midline regions and the default-mode network?", *NeuroImage*, 57: 1221-1233.
- Quine, W. (1951) "Two Dogmas of Empiricism", *The Philosophical Review*, 60 (1): 20-43.
- Quine, W. (1960) *Word and Object*, Cambridge, MA: MIT Press.
- Quine, W. (1969) "Epistemology naturalized" in *Ontological Relativity and Other Essays*, New York, NY: Columbia University Press, pp. 69-90.
- Quine, W. (1975) "A letter to Mr. Ostermann", in *The Owl of Minerva: Philosophes on Philosophy* (C. Bontempo & S. Odell eds.), New York: McGraw-Hill, pp. 227-230.
- Quine, W. (1980) *From A Logical Point of View*, Cambridge, MA: Harvard University Press.
- Quine, W. (1981) *Theories and Things*, Cambridge, MA: Harvard University Press.
- Quine, W. (1995) *From Stimulus to Science*, Cambridge, MA: Harvard University Press.
- Raichle, M. (2010) "Two views of brain function", *Trends in Cognitive Science*, 14 (4): 180-190.
- Raichle, M. (2015) "The restless brain: how intrinsic activity organizes brain function", *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370 (1668): 20140172.
- Rainville, P. & Price, D. (2003) "Hypnosis phenomenology and the neurobiology of consciousness", *International Journal of Clinical and Experimental Hypnosis*, 51 (2): 105-129.
- Ramsey, W. (2007) *Representation Reconsidered*, Cambridge: Cambridge University Press.
- Ratcliffe, M. (2006) "Folk psychology' is not folk psychology", *Phenomenology and the Cognitive Sciences*, 5: 31-52.
- Rawls, J. (2001) *Justice as Fairness: a Restatement*, Cambridge, MA: Harvard University Press.

- Rees, G., Krieman, G. & Koch, C. (2002) "Neural correlates of consciousness in humans", *Nature Reviews Neuroscience*, 3: 261-270.
- Reichenbach, H. (1968) *Der Aufstieg der wissenschaftlichen Philosophie*, Braunschweig: Friedr. Vieweg & Sohn.
- Rescher, N. (1991) "Thought experiments in presocratic philosophy", in *Thought Experiments in Science and Philosophy* (T. Horowitz & G. Massey, eds.), Lanham: Rowman & Littlefield, pp. 31-42.
- Rescher, N. (2001) *Philosophical Reasoning: A Study in the Methodology of Philosophizing*, Oxford: Blackwell.
- Rescher, N. (2006) *Philosophical Dialectics. An Essay on Metaphilosophy*, Albany, NY: State University of New York Press.
- Rescorla R. & Wagner, A. (1972) "A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement", in *Classical Conditioning. II: Current Research and Theory* (A. Black & W. Prokasy, eds.), New York: Appleton-Century-Crofts, pp. 12-33.
- Robinson, D. (org.) (2007) *Neuroscience & Philosophy: Brain, Mind & Language*, New York, NY: Columbia University Press
- Rorty, R. (ed.) (1967) *The Linguistic Turn: Recent Essays in Philosophical Method*, Chicago, IL: The University of Chicago Press.
- Roskies, A. (2002) "Neuroethics for the new millennium", *Neuron*, 35: 21-23.
- Rundle, B. (1997) *Mind in Action*, Oxford: Clarendon Press.
- Russell, B. (1967) *The Problems of Philosophy*, Oxford: Oxford University Press.
- Russell, B. (1992) *A Critical Exposition of the Philosophy of Leibniz*, London: Routledge.
- Ryle, G. (2009) *Collected Papers, Volume II: Collected Essays 1929–1968*, London: Routledge.
- Sadaghiani, S., Hesselmann, G., Friston, K. & Kleinschmidt, A. (2010) "The relation of ongoing brain activity, evoked neural responses, and cognition", *Frontiers in Systems Neuroscience*, 4 (20): s/pag.

- Sadaghiani, S., Hesselmann, G. & Kleinschmid, A. (2009) "Distributed and antagonistic contributions of ongoing activity fluctuations to auditory stimulus detection", *Journal of Neuroscience*, 29 (42): 13410-13417.
- Sadaghiani, S., Poline, J., Kleinschmidt, A. & D'Esposito, M. (2015) "Ongoing dynamics in large-scale functional connectivity predict perception", *Proceedings of the National Academy of Sciences of the United States of America*, 112 (27): 8463-8468.
- Sainsbury, R. (2010) *Fiction and Fictionalism*, New York, NY: Routledge.
- Scalabrini, A., Mucci, C., Angeletti, L., Northoff, G. (2020) "The self and its world: a neuroecological and temporo-spatial account of existential fear", *Clinical Neuropsychiatry*, 17 (2), 46-58.
- Schaffner, K. (1967) "Approaches to reduction", *Philosophy of Science*, 34: 137-147.
- Schiff, N. (2009) "Central thalamic deep-brain stimulation in the severely injured brain: Rationale and proposed mechanisms of action", *Annals of the New York Academy of Sciences*, 1 (157): 101-116.
- Schiff, N. (2010) "Recovery of consciousness after severe brain injury: The role of arousal regulation mechanisms and some speculation on the heart-brain interface", *Cleveland Clinic Journal of Medicine*, 77 (Suppl. 3): 27-33.
- Schmitz, T., Kawahara-Baccus, T. & Johnson, S. (2004) "Metacognitive evaluation, self-relevance, and the right prefrontal cortex", *NeuroImage*, 22: 941-947.
- Schwitzgebel, E. (2011) *Perplexities of Consciousness*, Cambridge, MA: MIT Press.
- Searle, J. (1980) "Minds, Brains and Programs", *Behavioral and Brain Sciences*, 3 (3): 417-457.
- Searle, J. (1997) *The Mystery of Consciousness*, London: Touchstone.
- Searle, J. (1999) "The future of philosophy", *Philosophical Transactions of the Royal Society*, 354 (1392): 2069-2080.
- Searle, J. (2000) "Consciousness", *Annual Reviews*, 23: 557-578.
- Searle, J. (2004a) *Mind: A Brief Introduction*, Oxford: Oxford University Press.

- Searle, J. (2004b) "Comments on Noë & Thompson 'Are there NCCs?'" , *Journal of Consciousness Studies*, 11: 79-82.
- Searle, J. (2007) "Putting consciousness back in the brain" in *Neuroscience & Philosophy: Brain, Mind & Language*, New York, NY: Columbia University Press, pp. 97-124.
- Searle, J. (2015) *Seeing Things as They Are*, Oxford: Oxford University Press.
- Seger, C., Stone, M. & Keenan, J. (2004) "Cortical activations during judgments about the self and another person", *Neuropsychologia*, 42: 1168-1177.
- Sellars, W. (1956) "Empiricism and the philosophy of mind", in *Minnesota Studies in the Philosophy of Science* (H. Feigl & M. Scriven, eds.), Minneapolis, MN: University of Minnesota Press, pp. 253-329.
- Sellars, W. (1963) *Science, Perception and Reality*, London: Routledge and Kegan Paul.
- Seth, A. (2013) "Interoceptive inference, emotion, and the embodied self", *Trends in Cognitive Sciences*, 17 (11): 565-573.
- Seth, A. (2017) "The cybernetic Bayesian brain: From interoceptive inference to sensorimotor contingencies" in *Philosophy and Predictive Processing* (T. Metzinger & J. M. Windt, eds.), Frankfurt am Main: MIND Group, pp. 1-24.
- Seth, A., Suzuki, K. & Critchley, H. (2012) "An interoceptive predictive coding model of conscious presence", *Frontiers in Psychology*, 2 (395): 1-16.
- Shaffer, J. (1968) *Philosophy of Mind*, Englewood Cliffs: Prentice-Hall.
- Shepard, R. (2008) "The step to rationality: The efficacy of thought experiments in science, ethics, and free will", *Cognitive Science*, 32: 3-35.
- Sherrington, C. (1947) *The Integrative Action of the Nervous System*, Cambridge, Cambridge University Press.
- Sherrington, C. (2009) *Man on his Nature*, Cambridge: Cambridge University Press.

Silva, S., Alacoque, X., Fourcade, O., Samii, K., Marque, P., Woods, R., Mazziotta, J., Chollet, F. & Loubinoux, I. (2010) "Wakefulness and loss of awareness: Brain and brainstem interaction in the vegetative state", *Neurology*, 74 (4): 313-320.

Sklar, L. (1967) "Types of inter-theoretic reduction", *The British Journal for the Philosophy of Science*, 18: 109-124.

Slaney, K. & Maraun, M. (2005) "Analogy and metaphor running amok: An examination of the use of explanatory devices in neuroscience", *Journal of Theoretical and Philosophical Psychology*, 25: 153-173.

Smart, J. (1959) "Sensations and brain processes", *The Philosophical Review*, 68 (2): 141-156.

Smit, H. & Hacker, P. (2014) "Seven misconceptions about the mereological fallacy: A compilation for the perplexed", *Erkenntnis*, 79 (5): 1077-1097.

Smith, D. (2007) *Husserl*, London: Routledge.

Smith, D. & Thomasson, A. L. (eds.) (2005) *Phenomenology and Philosophy of Mind*, Oxford: Clarendon Press.

Sneed, J. (1971) *The Logical Structure of Mathematical Physics*, Dordrecht: D. Reidel.

Snowdon, P. (2007) "G. E. Moore on sense-data and perception", in *Themes From G. E. Moore: New Essays in Epistemology and Ethics* (S. Nuccetelli & G. Seay, eds.), Gloucestershire: Clarendon Press, pp. 119-140.

Sommers, T. (2003) "Galen Strawson (interview)", in *Believer Magazine* (J. Shenk, ed.), consultado no 10 de Julho de 2018, URL = <<https://believermag.com/an-interview-with-galen-strawson/>>.

Sperry, R. (1961) "Cerebral organization and behavior: The split brain behaves in many respects like two separate brains, providing new research possibilities", *Science*, 133 (3466): 1749-1757.

Stich, S. (1983) *From Folk Psychology to Cognitive Science*, Cambridge, MA: MIT Press.

Stich, S. & Nichols, S. (2003) "Folk psychology" in *The Blackwell Guide to Philosophy of Mind* (S. Stich and T. Warfield, eds.), Oxford: Blackwell, pp. 235-255.

Stich, S. & Ravenscroft, I. (1992) "What is folk psychology?", *Cognition*, 50: 447-468.

Stuart, M., Fehige, Y. & Brown, J. (eds.) (2018) *The Routledge Companion to Thought Experiments*, London: Routledge.

Sui, J., Liu, C. & Han, S. (2009) "Cultural difference in neural mechanisms of self-recognition", *Social Neuroscience*, 4 (5): 402-411.

Suppe, F. (1974) *The Structure of Scientific Theories*, Urbana, IL: University of Illinois Press.

Suppes, P. (1956) *Introduction to Logic*, Princeton, NJ: van Nostrand.

Suzuki, K., Garfinkel, S., Critchley, H. & Seth, A. (2013) "Multisensory integration across exteroceptive and interoceptive domains modulates self-experience in the rubber-hand illusion", *Neuropsychologia*, 51 (13): 2909-2917.

Symon, J. (2002) *On Dennett*, Belmont, CA: Wadsworth.

Sytsma, J. (2007) "Language Police Running Amok", *Journal of Theoretical and Philosophical Psychology*, 27: 89-103.

Thagard, P. (2012) *Cognitive Science of Science: Explanation, Discovery and Conceptual Change*, Cambridge, MA: MIT Press.

Thoma, J. (2016) "On the hidden thought experiments of economic theory", *Philosophy of the Social Sciences*, 46: 129-146.

Thompson, E. (1992) "Novel colours", *Philosophical Studies*, 68 (3): 321-349.

Thompson, E. (2007) *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*, Cambridge, MA: Harvard University Press.

Thompson, E. (2015) *Waking, Dreaming, Being: Self and Consciousness in Neuroscience, Meditation and Philosophy*, New York, NY: Columbia University Press.

Thompson, E., Lutz, A. & Cosmelli, D. (2005) "Neurophenomenology: an introduction for neurophilosophers", in *Cognition and the Brain: the Philosophy and Neuroscience Movement* (A. Brook & K. Akins, eds.), Cambridge: Cambridge University Press, pp. 40-97.

- Thompson, E., Noë, A. & Pessoa, L. (1999) "Perceptual completion: A case study in phenomenology and cognitive science", in *Naturalizing Phenomenology* (J. Petitot, F. Varela, B. Pachoud & J. Roy, eds.), Stanford, CA: Stanford University Press, pp. 161-195.
- Thompson, E. & Varela, F. (2001) "Radical embodiment: Neural dynamics and consciousness", *Trends in Cognitive Sciences*, 5 (10): 418-445.
- Tononi, G. (2004) "An information integration theory of consciousness", *BMC Neuroscience*, 5: 42.
- Tononi, G. (2008) "Consciousness as integrated information: A provisional manifesto", *The Biological Bulletin*, 215 (3): 216-242.
- Tononi, G. & Edelman, G.M. (1998) "Consciousness and complexity", *Science*, 282: 1846-1851.
- Tononi, G. & Koch, C. (2008) "The neural correlates of consciousness: An update", *Annals of the New York Academy of Sciences*, 1 (124): 239-261.
- Torin, A. (2003) "Churchland on arguments against physicalism" in *Consciousness In and Out: Phenomenology, Neuroscience, and the Nature of Experience* (R. Brown, ed.), Berlin: Springer.
- Tsuchiya, N. & Adolphs, R. (2007) "Emotion and consciousness", *Trends in Cognitive Sciences*, 11 (4): 158-167.
- Tsuchiya, N., Block, N. & Koch, C. (2012) "Top-down attention and consciousness: Comment on Cohen *et al.*", *Trends in Cognitive Sciences*, 16 (11): 527.
- Tsuchiya, N., Wilke, M., Frässle, S. & Lamme, V. (2015) "No-report paradigms: Extracting the true neural correlates of consciousness", *Trends in Cognitive Sciences*, 19 (12): 757–770.
- Tucker, A. (1834) *The Light of Nature Pursued*, London: Thomas Tegg & Son.
- Tye, M. (1986) "The subjective qualities of experience", *Mind*, 95: 117.
- van Dyck, M. (2003) "The roles of one thought experiment in interpreting quantum mechanics: Werner Heisenberg meets Thomas Kuhn", *Philosophica*, 72: 1-21.
- van Fraassen, B. (1972) "A formal approach to the philosophy of science", in *Paradigms and Paradoxes* (R. Colodny, ed.), Pittsburgh, PA: University of Pittsburgh, 303-351.

van Gelder, T. (1999) "Wooden iron? Husserlian phenomenology meets cognitive science", in *Naturalizing Phenomenology* (J. Petitot, F. Varela, B. Pachoud and J. Roy, eds.), Stanford, CA: Stanford University Press, pp. 245-265.

van Rijnsoever, F. & Hessels, L. (2011) "Factors associated with disciplinary and interdisciplinary research collaboration", *Research Policy*, 40: 463-472.

Vanhaudenhuyse, A., Noirhomme, Q., Tshibanda, L., Bruno, M., Boveroux, P., Schnakers, C., Soddu, A., Perlberg, V., Ledoux, D., Brichant, J., Moonen, G., Maquet, P., Greicius, M., Laureys, S. & Boly, M. (2010) "Default network connectivity reflects the level of consciousness in non-communicative brain-damaged patients", *Brain*, 133 (Pt.1): 161-171.

Varela, F. (1996) "Neurophenomenology: A methodological remedy for the hard problem", *Journal of Consciousness Studies*, 3 (4): 330-349.

Varela, F. (1998) "Neurophenomenology: A methodological remedy for the hard problem" in *Explaining Consciousness – the Hard Problem* (J. Shear, ed.), Cambridge, MA: MIT Press, pp. 337-358.

Varela, F. & Shear, J. (1999) "First-person methodologies: what, why, how", *Journal of Consciousness Studies*, 6: 1-14.

Varela, F., Thompson, E. & Rosch, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Cambridge, MA: MIT Press.

Vetter, P., Sanders, L. & Muckli, L. (2014) "Dissociation of prediction from conscious perception", *Perception*, 43 (10): 1107-1113.

von Eckardt, B. (1994) "Folk psychology", in *A Companion to Philosophy of Mind* (S. Guttenplan, ed.), Oxford: Blackwell, pp. 317-319.

von Helmholtz, H. (1867) *Handbuch der physiologischen Optik*, Leipzig: Leopold Voss.

Wacongne, C., Labyt, E., Wassenhove, V., Bekinschtein, T., Naccache, L. & Dehaene, S. (2011) "Evidence for a hierarchy of predictions and prediction errors in human cortex", *Proceedings of the National Academy of Sciences of the USA*, 108 (51): 20754-20759.

- Walker, R. (1989) *The Coherence Theory of Truth: Realism, Anti-realism, Idealism*, London/New York: Routledge.
- Wanja, W. & Metzinger, T. (2017) "Vanilla PP for philosophers: A primer on predictive processing" in *Philosophy and Predictive Processing* (T. Metzinger & J. M. Windt, eds.), Frankfurt am Main: MIND Group.
- Williams, B. (2006) *Philosophy as a Humanistic Discipline* (A. Moore, ed.), Princeton, NJ: Princeton University Press.
- Williamson, T. (2007) *The Philosophy of Philosophy*, Malden, MA: Blackwell.
- Williamson, T. (2019) "Armchair philosophy", *Epistemology & Philosophy of Science*, 56 (2): 19-25.
- Wittgenstein, L. (1953) *Philosophical Investigations* (G. Anscombe, trad.), Oxford: Blackwell.
- Wittgenstein, L. (1961) *Tractatus Logico-Philosophicus* (D. Pears & B. McGuinness, trad.), Routledge: London.
- Wittgenstein, L. (1966) *Lectures and Conversations on Aesthetics, Psychology and Religious Belief* (C. Barret, ed.) Oxford: Blackwell.
- Wittgenstein, L. (1967) *Zettel* (G. Anscombe & G. von Wright, eds., trad.), Berkeley, CA: University of California Press.
- Wittgenstein, L. (1968) "Notes for lectures on 'private experience' and 'sense data'", *Philosophical Review*, 77 (3): 275-320.
- Wittgenstein, L. (1969) *The Blue Book*, 2nd. ed. New York: Blackwell-Publishing.
- Wittgenstein, L. (1977) *Remarks on Colour* (G. Anscombe, ed.; L. McAlister, trad.), Berkeley: University of California Press.
- Ye, X., Zhu, R., Zhou, X., He, S. & Wang, K. (2019) "Slower and less variable binocular rivalry rates in patients with bipolar disorder, OCD, major depression, and schizophrenia", *Frontiers of Neuroscience*, 13: 514.
- Ylikoski, P. (2003) "Thought experiments in science studies", *Philosophica*, 72: 35-59.

Zahavi, D. (2003) *Husserl's Phenomenology*, Stanford, CA: Stanford University Press.

Zambrano, Y. (2012) *Neuroepistemology: What the Neurons Knowledge Tries to Tell Us*, Baltimore, MA: Phy Psi K'a Publishing Co.

Zangwill, N. (1992) "Variable reduction not proven", *Philosophical Quarterly*, 42: 214-218.

Zhu, Y. & Han, S. (2008) "Cultural differences in the self: From philosophy to psychology and neuroscience", *Social and Personality Psychology Compass*, 2: 1799-1811.