UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE CIÊNCIAS SOCIAIS E HUMANAS
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

Jonathas Kilque Villanova

CONSCIÊNCIA FENOMÊNICA, REPRESENTACIONALISMO DE SEGUNDA-ORDEM E NEGOCIAÇÕES METALINGUÍSTICAS

Santa Maria, RS
2022

Jonathas Kilque Villanova

CONSCIÊNCIA FENOMÊNICA, REPRESENTACIONALISMO DE SEGUNDA-ORDEM E NEGOCIAÇÕES METALINGUÍSTICAS

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do título de **Mestre em Filosofia**.

Prof. Dr. César Schirmer dos Santos

Santa Maria, RS
2022

This study was financied in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001

Villanova, Jonathas Kilque
Consciência Fenomênica, Representacionalismo de Segunda-ordem e Negociações Metalinguísticas / Jonathas Kilque Villanova.- 2022.
125 p.; 30 cm

Orientador: César Schirmer do Santos
Dissertação (mestrado) - Universidade Federal de Santa Maria, Centro de Ciências Sociais e Humanas, Programa de Pós-Graduação em Filosofia, RS, 2022

1. Consciência 2. Representacionalismo de Segunda ordem 3. Ética conceitual 4. Negociações Metalinguísticas 5. Filosofia da Mente I. , César Schirmer do Santos II. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

Declaro, JONATHAS KILQUE VILLANOVA, para os devidos fins e sob as penas da lei, que a pesquisa constante neste trabalho de conclusão de curso (Dissertação) foi por mim elaborada e que as informações necessárias objeto de consulta em literatura e outras fontes estão devidamente referenciadas. Declaro, ainda, que este trabalho ou parte dele não foi apresentado anteriormente para obtenção de qualquer outro grau acadêmico, estando ciente de que a inveracidade da presente declaração poderá resultar na anulação da titulação pela Universidade, entre outras consequências legais.

Jonathas Kilque Villanova

CONSCIÊNCIA FENOMÊNICA, REPRESENTACIONALISMO DE SEGUNDA-ORDEM E NEGOCIAÇÕES MENTALINGUÍSTICAS

Dissertação apresentada ao Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do título de **Mestre em Filosofia**.

Aprovada em 02 de junho de 2022.

César Schirmer dos Santos, Dr. (UFSM)
(Presidente/Orientador)

Julia Telles Menezes, Dr. (UFF)

André Joffily Abath, Dr. (UFMG)

Santa Maria, RS
2022

AGRADECIMENTOS

Aos meus pais, Renato e Rojane e irmão Anderson, por me forneceram as condições materiais e afetivas para trilhar meu próprio caminho na vida. Eu não poderia estar aqui sem o apoio e compreensão de vocês, muito obrigado por tudo. Amo vocês.

À Helena Zirbes Maurer, que esteve comigo em todas as etapas do percurso e sem a qual eu não teria iniciado minha graduação em Filosofia. Teu carinho, compreensão e amor foram e continuam sendo um pilar essencial da minha trajetória acadêmica e da minha vida. Tudo o que realizei tem também a tua força. Obrigado pela paciência e pelo suporte incomensurável que tua companhia me proporciona todos os dias. Eu te amo muito, meu bem.

Aos professores e professores que fizeram parte de todos meus anos de formação. Em especial, agradeço à prof. Vera pelo letramento e ao prof. Vitor Hugo pelo primeiro contato com as ciências humanas. Na graduação, agradeço a todos as professoras e professores pelas aulas, conversas e orientações. Em especial, agradeço os prof. Marcelo Fabri e prof. Renato Fonseca pelas aulas instigantes e por despertar minha curiosidade pela leitura das obras de Brentano e pelas questões da filosofia da mente e da consciência. Ao prof. e orientador César Santos, agradeço por fornecer-me as ferramentas necessárias para realizar pesquisa em filosofia e por insistir na relevância dos métodos na filosofia.

Aos colegas e amigos que, de alguma maneira, fizeram parte da minha trajetória acadêmica até esse ponto, agradeço profundamente. Em especial, agradeço à Érica Vieira dos Santos, pela amizade e companheirismo regados a mate, que se estenderam para além dos muros da UFSM. Aos amigos Leonardo Piovesan, Danrlei Souza e Bismarck Bório, pela camaradagem e por suportarem os papos filosóficos intermináveis.

À Alexandra Elbakyan, criadora do *Library Genesis*. Sem o acesso ao conhecimento proporcionado por sua iniciativa, haveria uma gigantesca lacuna na minha formação e na de milhões de estudantes em todo planeta. Um mundo no qual o acesso ao conhecimento está disponível a todos é uma utopia que vale à pena tentarmos concretizar. Iniciativas como as de Elbakyan são um primeiro passo nessa direção, mas não podem ser as únicas – o compromisso social com o acesso ao conhecimento é um dever que se estende à toda sociedade, em especial às instituições que a compõem e os indivíduos que as representam. Aos que lutam por tornar esse ideal uma realidade, agradeço-os.

À CAPES e ao CNPq, pelo apoio financeiro que permitiu a realização desta pesquisa.

RESUMO

CONSCIÊNCIA FENOMÊNICA, REPRESENTACIONALISMO DE SEGUNDA-ORDEM E NEGOCIAÇÕES METALINGUÍSTICAS

AUTOR: Jonathas Kilque Villanova
ORIENTADOR: César Schirmer dos Santos

Esta dissertação consiste de três capítulos. O capítulo 1 aborda as variedades do conceito de consciência, nele apresentamos uma proposta taxonômica do conceito orientada pela aridade e escopo de aplicação do predicado "consciente". Através da taxonomia oferecida, apresentamos alguns esclarecimentos sobre o significado e aplicação de cada conceito, bem como discutimos relações relevantes entre eles. No capítulo 2, oferecemos uma introdução ao representacionalismo de segunda-ordem. A primeira parte do capítulo se concentra na relevância do Princípio da Transitividade para a caracterização e defesa dessa abordagem. Segundo o Princípio da Transitividade: estados mentais conscientes são estados de que um sujeito está consciente de ter. Após discutir o que faz desse princípio uma intuição pré-teórica admissível, apresentamos o argumento do representacionalismo de segunda-ordem que fundamenta sua estratégia explicativa, segundo o qual, um estado mental é (fenomênicamente) consciente apenas se há uma metarepresentação apropriada desse estado. Após apresentar esse argumento, discutimos duas objeções ao argumento centrados no princípio da transitividade. Na segunda parte do capítulo revisamos as evidências empíricas que corroboram com o representacionalismo de segunda-ordem sobre a consciência e apresentamos sucintamente algumas das consequências empíricas implicadas pela abordagem, bem como as diferenças e semelhanças com abordagens rivais. No capítulo 3, argumentamos que é possível interpretar o debate entre abordagens de primeira-ordem e segunda-ordem da consciência como sendo uma negociação metalinguística. Primeiramente, elucidamos o que são negociações metalinguísticas, e quais condições um debate deve satisfazer para ser interpretável como uma negociação metalinguística. Em seguida, argumentamos, através de evidências textuais, que o debate entre abordagens de primeira-ordem e segunda-ordem satisfazem todas as quatro condições elencadas. Portanto, o debate pode ser interpretado como uma disputa normativa sobre o que o conceito de consciência fenomênica deveria significar. Por fim, concluímos sugerindo como essa interpretação pode ser frutífera para desenvolvimentos futuros na investigação teórica e empírica da consciência.

Palavras-chave: Filosofia da Consciência; Consciência Fenomênica; Representacionalismo de Segunda-ordem sobre a Consciência; Negociações Metalinguísticas; Ética Conceitual.

ABSTRACT

PHENOMENAL CONSCIOUSNESS, HIGHER-ORDER REPRESENTATIONALISM AND METALINGUISTIC NEGOTIATIONS

AUTOR: Jonathas Kilque Villanova
ADVISOR: César Schirmer dos Santos

This dissertation is divided in three chapters. Chapter 1 addresses the varieties of the concept of consciousness, in this chapter we present a taxonomic proposal of the concept of consciousness oriented by the arity and scope of application of the predicate "conscious". Through this taxonomy, we present some elucidations of the meaning and application of each concept, as well as a discussion of the meaningful relations between them. In chapter 2, we offer an introduction to higher-order representational theories of consciousness. The first part of the chapter focuses on the relevance of the Transitivity Principle to the characterization and defense of this approach to consciousness. According to the Transitivity Principle: conscious states are states one is aware of being in. After discussing what makes this principle intuitively admissible, we present the argument that justifies higher-order representational theories explanatory strategy, according to which a mental state is (phenomenally) conscious only if there is an adequate metarepresentation of that state. After presenting this argument, we discuss two objections of this argument focusing on the transitivity principle. In the second part of the chapter, we review the empirical evidence supporting second-order representationalism and briefly present some of the empirical consequences implied by the approach, as well as the differences and similarities with rival approaches. In Chapter 3, we argue that it is possible to interpret the debate between first- and second-order approaches to consciousness as a metalinguistic negotiation. First, we elucidate what metalinguistic negotiations are, and what conditions a debate must satisfy to be interpretable as a metalinguistic negotiation. Then, we argue, through textual evidence, that the debate between first-order and second-order approaches satisfy all four conditions listed. Therefore, the debate can be interpreted as a normative dispute over what the concept of phenomenal consciousness should mean. Finally, we conclude by suggesting how this interpretation can be fruitful for future developments in the theoretical and empirical investigation of consciousness.

Keywords: Philosophy of Consciousness; Phenomenal Consciousness; Higher-order Representational Theories of Consciousness; Metalinguistic Negotiations; Conceptual Ethics.

SUMÁRIO

INTRODUÇÃO 13

1. **VARIEDADES DA CONSCIÊNCIA** 18

1.1. LÓGICA E SEMÂNTICA DO PREDICADO "CONSCIENTE" 19

1.1.1 Consciência de criatura: Intransitiva 20

1.1.2 Consciência de criatura: Transitiva 21

1.1.3 Consciência de estado: Intransitiva 22

1.1.3.1 Consciência de acesso 24

1.1.3.2 Consciência fenomênica 25

1.1.4 Consciência de estado: Transitiva 28

2 REPRESENTACIONALISMO DE SEGUNDA-ORDEM: UMA VISÃO GERAL 30

2.1 ABORDAGENS DE PRIMEIRA-ORDEM E SEGUNDA-ORDEM SOBRE A CONSCIÊNCIA FENOMÊNICA 31

2.1.1 Representacionalismo de primeira-ordem e de segunda-ordem 32

2.1.1.1 A teoria PANIC de Michael Tye 34

2.1.1.1.1 Objeções à teoria PANIC 36

2.2 PRINCÍPIO DA TRANSITIVIDADE: UMA INTUIÇÃO PRÉ-TEÓRICA ACEITÁVEL? 38

2.2.1 Plausibilidade pré-teórica 39

2.2.2 Evidência fenomenológica 39

2.2.3 Utilidade teórica/metodológica na investigação empírica 40

2.2.4 Qualificações ao princípio da transitividade 41

2.3 O ARGUMENTO MESTRE DO REPRESENTACIONALISMO DE SEGUNDA-ORDEM: 43

2.4 O EXPLANANS DA ABORDAGEM REPRSENTACIONALISTA DE SEGUNDA-ORDEM 48

2.4.1 O princípio da transitividade como *explanans* da abordagem de segunda-ordem 48

2.4.2 Pensamentos e percepções de segunda-ordem 49

2.4.3 Teorias de segunda-ordem e metacognição: 51

2.5 MOTIVAÇÕES EMPÍRICAS PARA A ADOÇÃO DO REPRESENTACIONALISMO DE SEGUNDA-ORDEM 55

2.5.1 Implicações empíricas do representacionalismo de segunda-ordem 56

2.5.2 Diferenças entre o representacionalismo de segunda-ordem e a teoria da área do trabalho global 60

2.6.2.1 A função da consciência no Representacionalismo de Segunda-ordem 61

2.5.3 Correlatos neurais da consciência e a relação entre (relatos de) experiência subjetiva e o desempenho cognitivo em estudos neurocognitivos sobre a consciência 63

3 SERIA O DEBATE ENTRE ABORDAGENS DE PRIMEIRA-ORDEM E DE SEGUNDA-ORDEM DA CONSCIÊNCIA FENOMÊNICA UMA NEGOCIAÇÃO METALINGUÍSTICA? 70

3.1 INTRODUÇÃO 70

3.2 NEGOCIAÇÕES METALINGUÍSTICAS: O QUE SÃO? E COMO IDENTIFICÁ-LAS? 73

3.2.1 O que são negociações metalinguísticas 73

3.2.2 Condições para negociações metalinguísticas e justificação 76

3.3 O DESACORDO ENTRE ABORDAGENS DE PRIMEIRA E SEGUNDA-ORDEM DA CONSCIÊNCIA FENOMÊNICA: DUAS INTERPRETAÇÕES 82

3.3.1 Interpretação factual 84

3.3.2 Interpretação metalinguística 86

3.4 UMA DISPUTA SOBRE TERMOS: A CONCEPÇÃO NAGELIANA E SUA AMBIGUIDADE 88

3.4.1 A recepção da concepção nageliana da consciência fenomênica entre abordagens de primeira-ordem e de segunda-ordem 89

3.4.2 Problemas com a concepção nageliana da consciência: "algo como é" como fonte de ambiguidade 93

3.4.2.1 Ambiguidade e condições de verdade de uma proposição 93

3.4.2.2 O caso do motorista distraído e a ambiguidade da expressão "algo como é"..... 95

3.4.3 Dois sentidos para "algo como é": subjetividade mundana e subjetividade experiencial 99

3.5 A CONDIÇÃO DA NORMATIVIDADE DO SIGNIFICADO NO DEBATE ENTRE ABORDAGENS DE PRIMEIRA-ORDEM E SEGUNDA-ORDEM DA CONSCIÊNCIA 105

3.5.1 Argumento a partir da ambiguidade à semântica normativa do conceito de consciência 106

3.5.2 Uma objeção a partir da ambiguidade da concepção nageliana da consciência 109

3.6 CONCLUSÃO E DESDOBRAMENTOS..... 112

CONCLUSÃO 117

REFERÊNCIAS..... 119

INTRODUÇÃO

A consciência é um fenômeno ao mesmo tempo familiar e misterioso. Todos nós estamos familiarizados, em primeira mão, com como é sentir o calor do sol em nossa pele, a dor de uma cãimbra ou perceber a tonalidade azul do céu. Experiências como essas são fenomênicamente conscientes. Parafraseando a famosa passagem de Nagel (1974) sobre a consciência fenomênica, há algo como é, para nós, experienciarmos esses eventos mentais. Além disso, enquanto falantes competentes do português, todos nós somos capazes de utilizar o termo "consciência" ou "consciente" em diversos contextos do dia-a-dia, como quando afirmarmos que recobramos a consciência a cada manhã ao acordar ou quando perguntamos se alguém agiu sob um estado alterado de consciência.

Por outro lado, a familiaridade que temos com esses eventos mentais pode ser contrastada com as amplas lacunas no nosso conhecimento empírico sobre esses fenômenos: como o nosso corpo e, em especial, o nosso cérebro é capaz de produzir estados mentais conscientes? Por que temos estados mentais conscientes, ao invés de termos apenas estados mentais inconscientes? E o que diferencia nossos estados mentais inconscientes de nossos estados mentais que são conscientes? Essas e outras questões continuam em aberto, sendo um objeto de investigação, em parte da investigação filosófica e, em parte, da investigação empírica nas ciências cognitivas e neurociências.

A constatação de que há muito o que se fazer em termos de uma elucidação empírica da consciência pode obscurecer o fato de que também no âmbito conceitual há muito a ser feito. A resposta às questões "O que diferencia estados mentais conscientes de inconscientes?" ou "Como o cérebro pode realizar biologicamente a consciência?", assim como as demais questões envolvendo o esclarecimento das suas funções ou realização física, são relevantemente informadas por como compreendemos o significado do termo consciência. Previamente à formulação de teorias que respondam às questões que foram elencadas acima, precisamos ter clareza sobre o que queremos dizer quando utilizamos o termo "consciência" ou "consciente". Entretanto, ao contrário do que as observações de que estamos familiarizados com como é ter estados mentais conscientes e de que somos competentes em empregar o conceito de consciência no nosso discurso cotidiano possam sugerir, nós não temos um conhecimento explícito e bem fundamentado sobre o que é a consciência, nem enquanto fenômeno na natureza, nem enquanto um conceito que pode servir de base para a articulação de uma teoria.

Nossa situação com relação ao conhecimento sobre o que é a consciência pode ser comparada ao caso da discrepância entre a competência linguística e o conhecimento explícito da gramática de uma língua. A prática na construção de sentenças gramaticais é uma competência atribuível a qualquer pessoa fluente em uma língua. Porém, do fato de que uma pessoa é capaz de observar ao conjunto de regras gramaticais na construção de sentenças, não se segue que essa pessoa seja capaz de enunciar essas regras. Como observa Strawson (2002, p.19), o domínio de uma prática não implica no domínio explícito da teoria dessa prática, embora possamos conceder que envolve um domínio implícito de alguns aspectos dessa teoria (e.g., seguir as regras gramaticais na construção de sentenças sem saber explicitamente quais são essas regras).

Esse exemplo inspirado pela analogia de Strawson é comparável ao caso da consciência em dois sentidos. Num primeiro sentido, estarmos familiarizados com como é termos estados mentais conscientes não implica em um conhecimento explícito sobre o que é a consciência enquanto um fenômeno natural. Sabemos como é a sensação do calor do sol em nossa pele, como é perceber a tonalidade do céu e sentir o sabor do café. Mas, de maneira alguma isso implica no conhecimento sobre os processos ou componentes dos mecanismos responsáveis por realizar os eventos mentais conscientes. Num segundo sentido, nossa capacidade de usarmos o termo “consciente” ou “consciência” em diversos contextos cotidianos não implica, como no exemplo acima, em um conhecimento explícito e sistemático das condições que observamos para aplicar esses termos, muito menos implica em estarmos de posse de uma teoria explícita sobre a consciência. Similarmente, a competência no uso do termo consciente em contextos cotidianos não implica em termos um conhecimento dos problemas (nem das respostas a esses problemas) vinculados à noção de consciência no contexto da teorização filosófica ou científica sobre a consciência.

Sem termos clareza sobre como utilizamos esse conceito e, em especial, sobre qual é o fenômeno ou conjunto de fenômenos na realidade que buscamos designar ao utilizar esse conceito, é muito improvável que possamos ser bem-sucedidos em elucidar as questões sobre como a consciência é realizada ou o que diferencia os estados mentais conscientes dos estados mentais inconscientes, por exemplo. Tendo isso em vista, a presente dissertação busca, em diferentes contextos, responder à questão “o que é a consciência?” através de uma elucidação do(s) conceito(s) utilizado(s) na teorização sobre esse fenômeno.

O capítulo 1 inicia notando a polissemia da noção de consciência. Tendo em vista que há diferentes maneiras de utilizarmos o termo “consciente”, oferecemos uma taxonomia dos conceitos de consciência orientada pelas características lógicas e semânticas de diferentes usos

do predicado “consciente”. Neste capítulo diferenciaremos quatro variedades de conceitos de consciência: (1) consciência de criatura (1a) intransitiva e (1b) transitiva; e (2) estados mentais (2a) intransitivamente conscientes e (2b) transitivamente conscientes. Ofereceremos também uma breve caracterização de cada um desses conceitos e discutiremos algumas das relações explanatórias relevantes entre eles.

Ter uma compreensão inicial do que queremos dizer ao utilizar algum conceito de consciência é um ponto de partida relevante na investigação sobre esse fenômeno. Muito do sucesso de uma elaboração e defesa de uma teoria da consciência pode depender das concepções que temos do fenômeno que buscamos explicar. Os conceitos de que dispomos no início de um trabalho investigativo orientam a formulação de nossas perguntas, hipóteses e abordagens explanatórias. Porém, para que possamos oferecer uma explicação sobre o que é a consciência enquanto um fenômeno natural e elucidemos às questões sobre sua realização física e as funções que ela desempenha, devemos operacionalizar esses conceitos em teorias capazes de oferecer respostas aos problemas conceituais e filosóficos, bem como hipóteses empiricamente testáveis.

Uma das teorias da consciência promissoras, tanto por sua capacidade de responder aos problemas filosóficos (ver, CARRUTHERS, 2000; 2005), como por suas previsões empíricas e compatibilidade com as evidências científicas (ver, capítulo 2, seção 2.6; LAU; ROSENTHAL, 2011; BROWN; LAU; LeDOUX, 2019), é o representacionalismo de segunda-ordem. Segundo o representacionalismo de segunda-ordem, estados mentais conscientes são estados mentais de que o sujeito está, de alguma maneira, consciente.

A intuição fundamental do representacionalismo de segunda-ordem de que estados mentais conscientes são estados de que um sujeito está, de alguma maneira, consciente de ter é chamada de *princípio da transitividade*. O status desse princípio é desafiado de diversas maneiras por teorias rivais e objetores do representacionalismo de segunda-ordem. No capítulo 2 dessa dissertação exploramos diversos problemas e questões que se concentram sobre a motivação, plausibilidade e fertilidade do princípio da transitividade enquanto um ponto de partida para a teorização sobre a consciência. Nas seções inicias do capítulo (2.1) apresentamos uma caracterização geral da abordagem representacionalista de segunda-ordem e de que maneira ela contrasta com a abordagem de primeira-ordem sobre a consciência. Como uma ilustração particular de uma teoria de primeira-ordem, apresentamos esquematicamente a teoria PANIC de Michael Tye (1995; 2000) e levantamos uma crítica que sugere que mais do que os elementos contidos nessa teoria são necessários para explicar o que distingue estados mentais conscientes de estados mentais inconscientes (KRIEGEL, 2002).

Um dos principais pontos de conflito entre o representacionalismo de segunda-ordem e as demais abordagens de primeira-ordem sobre a consciência parece dizer respeito ao modo como o *princípio da transitividade* é explorado como parte do *explanans* adotado pela teoria. Com o objetivo de elucidar as motivações para adotar esse princípio como ponto de partida para a teorização sobre a consciência, a seção 2.2 apresenta as três razões para considerar o *princípio da transitividade* um ponto de partida plausível e frutífero para uma teoria da consciência. Na sequência (seção 2.3), abordamos a questão sobre como o princípio da transitividade conduz à hipótese de que estados mentais conscientes são o resultado de processos metarepresentacionais, que caracteriza o representacionalismo de segunda-ordem, através da apresentação do que chamaremos de *argumento mestre do representacionalismo de segunda-ordem*. Na seção 2.4 expandimos nossa discussão sobre como o princípio da transitividade é explorado na formulação do *explanans* da abordagem de segunda-ordem, detalhando alguns aspectos de diferentes versões de teorias pertencentes à abordagem de segunda-ordem e qualificando a sua compreensão no contexto dessa abordagem. Por fim, na seção 2.5 discutimos e revisamos algumas das motivações empíricas que corroboram com a estratégia explicativa da abordagem de segunda-ordem sobre a consciência fenomênica. Além de apresentar algumas das evidências empíricas que corroboram com as hipóteses do representacionalismo de segunda-ordem, essa seção evidencia as diferenças nas previsões empíricas dessa abordagem em contraste com as abordagens rivais. Desta maneira, o capitulo 2 engloba tanto uma caracterização geral do representacionalismo de segunda-ordem como uma ilustração de como a variedade de conceitos de consciência, em especial a relação entre estados mentais *transitivamente conscientes* e *intransitivamente conscientes* pode ser empregada na formulação de hipóteses empiricamente informadas sobre a consciência.

Teorias pertencentes à abordagem representacionalista de segunda-ordem sobre a consciência contrastam com um grupo heterogêneo de teorias chamado de abordagem de primeira-ordem sobre a consciência. À primeira vista, o que está em disputa entre esses dois grupos diz respeito ao *explanans* adotado por cada abordagem. Enquanto representacionalistas de segunda-ordem sustentam que metarepresentações são um elemento necessário da explicação sobre o que faz de um estado mental consciente, as teorias pertencentes à abordagem de primeira-ordem negam essa afirmação. Porém, uma segunda interpretação sobre a relação entre essas duas abordagens é possível e frutífera para discussão sobre a consciência.

No capítulo 3 exploramos a possibilidade de interpretarmos o debate entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência como dizendo respeito ao *explanandum* que uma teoria da consciência deveria adotar. Nesse sentido, nossa hipótese é

a de que o debate entre essas duas abordagens pode ser lido como se fosse uma negociação metalinguística sobre o conceito de consciência que figura como *explanandum* de uma teoria. A primeira parte do capítulo (seção 3.2) tem o objetivo de caracterizar o que são negociações metalinguísticas, e quais condições um debate deve satisfazer para ser interpretável como uma negociação metalinguística. Basicamente, identificamos quatro condições, que podem ser resumidas da seguinte maneira: um debate pode ser interpretado como uma negociação metalinguística quando (1) aparenta haver uma disputa entre os debatedores e (2) há um estado real de desacordo entre eles, (3) que diz respeito a compreensão ou uso de um termo e, (4) a natureza desse desacordo é sobre como tal termo deveria ser utilizado.

Em seguida, argumentamos, através de evidências textuais, que o debate entre abordagens de primeira-ordem e segunda-ordem satisfazem todas as quatro condições elencadas. Na seção 3.3 oferecemos duas possíveis interpretações sobre o desacordo entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência. Na primeira interpretação, chamada de *interpretação factual*, o desacordo é sobre os fatos que melhor descrevem as condições em virtude das quais um estado mental é consciente. Na segunda interpretação, chamada de *intepretação metalinguística*, o desacordo trata sobre como deveríamos compreender o conceito de consciência que figura como *explanandum* de uma teoria da consciência e, consequentemente, sobre que tipo fenômeno é denotado por esse conceito. Na seção 3.4 discutimos como é possível compreender que essas duas abordagens disputem sobre o significado do termo "consciência". Partimos de evidências de que ambas as abordagens adotam a *concepção nageliana da consciência*, segundo a qual: para um estado mental M e um sujeito S, M é um estado fenomênicamente consciente, se e somente se, há algo como é para S ter/estar em M (ver FARRELL, 2017, p.3). Porém, tal concepção revela ser ambígua (seção 3.4.2), as quais abrem espaço para a sustentação de diferentes interpretações sobre como interpretar o aspecto ambíguo da concepção nageliana (seção 3.4.3). Por fim, argumentamos que, na medida em que a concepção nageliana é ambígua, há boas evidências para sustentarmos que as abordagens estão engajadas em uma discussão sobre a normatividade do significado do conceito de consciência, isto é, sobre como deveríamos interpretar e melhorar o conceito ambíguo empregado pela concepção nageliana (seção 3.5). Por fim, realizamos alguns apontamentos sobre as consequências para a investigação da consciência geradas pela hipótese de que o debate entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência pode ser interpretado como uma negociação metalinguística (seção 3.6).

1. **VARIEDADES DA CONSCIÊNCIA**

Este capítulo aborda a polissemia do conceito de consciência e a variedade de conceitos utilizados para lidar com esta polissemia. Propõe-se neste capítulo uma taxonomia dos conceitos de consciência fundamentado sobre critérios lógicos e semânticos do termo "consciente".

"Consciência" se diz de muitas maneiras. Algumas vezes dizemos que alguém está consciente como um modo de dizer que essa pessoa não está dormindo ou sob efeito de anestésicos gerais, mas acordada e atenta. Nesses casos, o predicado "consciente" é usado como um predicado de um lugar, cuja aplicação se dá usualmente em referência a organismos complexos. Outras vezes, usamos "consciente" como um predicado de dois lugares, como quando dizemos que alguém está consciente de algo. Nesses casos, a aplicação do predicado "consciente" exige especificar não apenas um organismo que é consciente, mas também algo (e.g, um objeto, um evento ou estado) sobre o que esse organismo é consciente.

Há diferentes variedades (de conceitos) de consciência. Do ponto de vista lógico, "Consciente" utilizado no primeiro sentido introduzido acima é um predicado distinto do predicado "consciente" no segundo sentido. Do ponto de vista semântico, o que se busca dizer através do uso da noção de "consciente" em cada caso também é diferente. Cada uso chama atenção para diferentes aspectos de nossas vidas mentais e teremos diferentes coisas para explicar sobre o que é ser consciente, dependendo do sentido que o termo é utilizado. O termo "consciente" (ou "consciência"), portanto, não corresponde a um único conceito, mas a uma pequena multidão deles. Neste capítulo introduzimos uma taxonomia de conceitos de consciência e caracterizamos alguns dos seus principais aspectos.

Muito do sucesso de uma elaboração e defesa de uma teoria da consciência pode depender das concepções que temos do fenômeno que buscamos explicar. Os conceitos de que dispomos no início de um trabalho investigativo desempenham o importante papel de orientar nossas perguntas, hipóteses e abordagens explanatórias sobre os fenômenos que eles designam. Uma má formulação, compreensão ou uso de um desses conceitos pode facilmente implicar em problemas para nossas teorias. A ambiguidade ou a vagueza de um conceito pode nos conduzir a formular teorias que confunde diferentes fenômenos como sendo o mesmo e a descrição equivocada de um fenômeno pode facilmente nos conduzir por um caminho de investigação equivocado.

A fim de evitar essas adversidades e fornecer uma caracterização perspícua do alvo de nosso discurso quando falamos sobre consciência, este capitulo introduz uma taxonomia das

variedades do conceito de consciência pautada pelas diferenças e semelhanças do predicado lógico "consciente". Nesse capítulo, discutimos quatro conceitos de consciência: (1) consciência de criatura; (2) consciência de estado; (2a) consciência de acesso e (2b) consciência fenomênica. E duas subdivisões (transitiva e intransitiva) dos conceitos (1) e (2). Nenhuma das noções que figuram nessa árvore taxonômica são inéditas na literatura. Com efeito, muitas delas são canonicamente utilizadas para introduzir o objeto de investigação de uma hipótese científica ou filosófica sobre a consciência – em especial as variedades (2a) e (2b)[1]. Deste modo, o presente capítulo permitirá garantirmos uma compreensão inicial de como se relacionam e se caracterizam, em linhas gerais, os diferentes conceitos de consciência frequentemente utilizados na literatura. Além disso, reconhecer as diferenças nos âmbitos de aplicação desses quatro conceitos e suas variedades nos permitirá isolar adequadamente os alvos explicativo das teorias discutidas nos próximos capítulos.

1.1. LÓGICA E SEMÂNTICA DO PREDICADO "CONSCIENTE"

"Consciente" é um adjetivo que pode ser usado de diversas maneiras. Uma primeira maneira através da qual podemos distinguir diferentes usos de "consciente" é segundo o conjunto dos termos aos quais ele se aplica. Atribuímos consciência para dois tipos de entidades diferentes: (i) indivíduos ou grupos de indivíduos (i.e., João; os seres humanos; uma vaca; os mamíferos, etc.) e (ii) estados mentais de indivíduos (crenças, desejos, percepções, memórias, etc.). Distinguindo os usos do predicado consciente nesses dois eixos, obtemos as variedades de consciência conhecidas, respectivamente, como *consciência de criatura* e *consciência de estado*[2].

Outra maneira através da qual podemos distinguir os usos de "consciente" é segundo a aridade apresentada pelo predicado em cada contexto de uso – isto é, o número de elementos que a propriedade representada por "consciente" deve se aplicar para ser adequadamente avaliada. Também nessa distinção podemos identificar dois usos distintos do predicado, conforme se compreenda que "consciente" é um predicado unário (de aridade um) ou um predicado binário (de aridade dois). Utilizado como um predicado unário, dizemos que algo é "consciente" em um sentido *intransitivo*, pois o uso desse predicado em linguagem natural

[1] Para uma discussão sobre a relevância das variedades de conceitos de consciência (1) e (2) no design experimental e articulação teórica nas ciências, ver BAYNE, 2007.
[2] Autores que empregam alguma versão dessa divisão são: Kriegel (2009); Carruthers (2000); Tye (1995; 200); Rosenthal (2005); Bayne (2007).

corresponderia ao uso do termo “consciente” em “Eu estou consciente.” ou “João recobrou a consciência”, no qual não há outro elemento que deva ser associado ao predicado para que ele seja adequadamente utilizado na sentença. Quando utilizamos “consciente” como um predicado binário, utilizamos essa noção como uma espécie de relação entre duas coisas ou de maneira *transitiva*. Nesse caso, o uso correspondente em linguagem natural exige um complemento, como quando dizemos “Eu estou consciente *de que está chovendo*” ou “O rato está consciente *de que há um gato à espreita*”.

O par de conceitos “transitivo” e “intransitivo” resulta de uma analogia com a transitividade ou intransitividade gramatical dos verbos. Assim como verbos transitivos exigem um complemento para que expressem um sentido completo, sentenças nas quais utilizamos o predicado binário “consciente” exigem um complemento que nos diga a respeito do que um sujeito é consciente (e.g., João está consciente *da gravidade do acidente*). De posse dessas duas distinções, obtemos uma taxonomia de variedades de conceitos de consciência quaternária: (1) consciência de criatura (1a) intransitiva e (1b) transitiva; e (2) estados mentais (2a) intransitivamente conscientes e (2b) transitivamente conscientes.

Nas seções seguintes discutimos o significado e as relações explicativas entre essas variedades de conceitos da consciência. Estabelecer caracterizações iniciais desses conceitos nos auxiliará na identificação do *explanandum* das teorias da consciência discutidas nos próximos capítulos dessa dissertação. Antecipando essa discussão brevemente, veremos que as noções de consciência intransitiva de estado e consciência transitiva de estado são as noções relevantes para a formulação e avaliação de teorias *representacionalistas de segunda-ordem da consciência* – que serão o foco de nossa discussão nos próximos capítulos. A razão pela qual estas duas noções são centrais para a formulação desse grupo de teorias é o fato de que seu objetivo se concentra elucidar as condições em que um estado mental M é consciente em termos de uma relação representacional com um estado mental intencional M*[3, 4].

1.1.1 Consciência de criatura: Intransitiva

[3] Esta é uma caracterização inicial das teorias representacionalistas de segunda-ordem que retrata de maneira fiel a ampla maioria de abordagens representacionalistas de segunda-ordem. Contudo, essa caracterização não é unanimemente aceita, de que não serve de critério para avaliar se uma teoria qualquer é uma versão da abordagem representacionalista de segunda-ordem.

[4] Note que, apesar da descrição poder sugerir que M e M* são estados mentais numericamente distintos, é possível que M e M* sejam idênticos, de modo que a relação se converteria numa autorrepresentação de um estado mental. Brentano (1874/2015) supostamente defendia uma versão desta abordagem e Kriegel (2009) explicitamente sustenta o que pode se chamado de uma teoria representacional de mesma-ordem (Same-order Theory).

Em alguns contextos de senso comum, utilizamos o predicado “consciente” como um predicado unário que é atribuído a indivíduos. Quando utilizamos “consciente” nesse sentido, fazemos uso do que se chama de *consciência de criatura* em um sentido *intransitivo*.

O predicado “consciente” no sentido de consciência intransitiva de criatura é um predicado unário que tem indivíduos como termos, sejam eles seres humanos, animais, outros organismos (ou ainda sistemas artificiais). Mais comumente, utilizamos “consciente” nesse sentido para distinguir o estado mental atual de uma pessoa que está acordada e lúcida de uma pessoa que está em coma, sob o efeito de um anestésico geral ou dormindo. Por exemplo, “João está consciente, mas Maria ainda está inconsciente” ou “C-3PO é um droide consciente”. Semelhantemente, podemos utilizar o predicado unário “consciente” para dizer que um organismo tem certas capacidades mentais, como pensar, perceber e sentir (em contraste com organismos ou outros indivíduos que não possuem essas capacidades). Em tais casos, o predicado “consciente” é aplicado ou a indivíduos ou a conjuntos de indivíduos e simplesmente denota um conjunto de capacidades que um indivíduo tem, independentemente do fato de ele exercer ou não em um determinado momento essas capacidades. Assim, um ser humano é consciente, mesmo que ele esteja atualmente dormindo ou em coma, pois ele é potencialmente capaz de perceber, sentir e pensar. O primeiro sentido de uso desse predicado diz algo a respeito do estado mental atual de uma criatura, mas o segundo sentido apenas nos informa sobre as disposições para estar em determinados tipos de estados mentais que essa criatura tem.

1.1.2 Consciência de criatura: Transitiva

Alguns filósofos reconhecem também uma versão binária do predicado “consciente” aplicado a indivíduos. Nesses casos, o predicado “consciente” exige como termos um organismo e um objeto ou evento. Por exemplo: *João* está consciente de que *está chovendo*; *o rato* está consciente *do gato que está à espreita*. Este uso do predicado consciente é conhecido como *consciência transitiva de criatura*. A ideia central por trás deste conceito é denotar a capacidade de um organismo de ter percepções ou pensamentos sobre algo (eventos, estados e objetos externos ou internos).

Embora logicamente não haja nada de errado com esse conceito, adotá-lo pode motivar confusões com relação à noção de *consciência transitiva de estado*. Pois, dizer de um organismo que ele tem consciência transitiva (no sentido de consciência de criatura) é dizer que esse organismo percebe ou pensa tal-e-tal. Assim, parece ser uma alternativa mais econômica e clara formularmos nossas sentenças não em termos de instanciações de percepções,

pensamentos e outros estados mentais e nos referirmos ao objeto sobre o qual são esses estados. Pois, se a atribuição de consciência transitiva de criatura redunda em dizermos que um indivíduo tem certos estados mentais *sobre algo*, então esse uso do conceito implica que há um estado mental através do qual esse indivíduo é consciente – veremos na seção 2.3 que isso equivale à *consciência transitiva de estado* (KRIEGEL, 2009, p. 29).

Reconhecendo essa redundância, Carruthers (2005, p.37) propõe abandonar o vocabulário sugerido pela noção de *consciência transitiva de criatura*. Pois não há benefícios explicativos gerados pela adoção desse conceito. Alternativamente, podemos simplesmente utilizar o vocabulário de senso comum que sugere expressões como "o indivíduo X percebe tal-e-tal (ou pensa que tal-e-tal) " como uma maneira de descrever casos em que um indivíduo está consciente de alguma coisa. Por isso, ao prosseguirmos nossa exposição, não faremos uso desse conceito de consciência.

1.1.3 Consciência de estado: Intransitiva

Em contextos ligeiramente mais distante do senso comum, como nas ciências cognitivas e na filosofia, o adjetivo "consciente" é comumente atribuído a estados mentais como pensamentos e percepções. Utilizar o predicado consciente tendo como universo de discurso os estados mentais de um indivíduo corresponde a variedade de consciência chamada de *consciência de estado.* Note que, diferentemente do conceito de consciência de criatura, o conceito de consciência de estado permite-nos falar de partes conscientes ou inconscientes da vida mental de uma criatura, esteja ela acordada ou sob efeito de anestésicos gerais. Nesse sentido o conceito de consciência de estado está apto para descrições mais sutis e complexas de nossa vida mental.

Apesar do conceito de consciência de estado não ser explicitamente utilizado no vocabulário psicológico de senso comum, é possível encontrar alguns indícios de que adotamos uma concepção, no mínimo, similar a esse conceito quando falamos de algumas atitudes proposicionais (e.g, crer, desejar e algumas emoções). Consideramos os seguintes casos como ilustrativos de situações em que o senso comum compreende que uma atitude proposicional admite variedades conscientes e inconscientes e, portanto, no qual um conceito semelhante ao de *consciência intransitiva de estado* está em jogo:

Durante o percurso do trabalho até em casa, João frequentemente se pega em meio a devaneios sobre onde irá tirar suas férias ou o que fará no final de semana. Ao longo do trajeto que João realiza, ele para e avança corretamente nos semáforos e desvia de eventuais

obstáculos, mesmo que sua consciência esteja absorta em devaneios sobre férias e descanso. Contudo, João não recorda de fato de ter realizado as manobras necessárias com o seu veículo para evitar obstáculos e obedecer aos sinais de trânsito. Em outras palavras, João não recorda de ter tido as percepções do ambiente enquanto estava dirigindo, apesar de claramente ter sido capaz realizar os ajustes comportamentais necessários para não bater o carro. Uma maneira de fazer sentido dessas situações consiste em adotar a ideia de que há percepções que ocorrem de maneira inconsciente.

Ao jogar xadrez, um enxadrista não iniciante está consciente de diversos movimentos que suas peças podem realizar. Mas o fato de que o enxadrista é capaz de pensar nas jogadas que lhe estão disponíveis (e avaliar as melhores entre elas) não significa que o enxadrista esteja constantemente repassando em sua mente as regras do jogo (i.e., as regras de movimentação de cada peça e as jogadas legais). É evidente que o enxadrista não iniciante conhece as regras do jogo, e que de alguma maneira essa informação é utilizada para planejar e avaliar suas jogadas, mas as crenças sobre as regras do xadrez são plausivelmente inconscientes ou subconsciente na maior parte do tempo, porque não ocupam um lugar central no palco da consciência – exceções surgem, por exemplo, quando se exige uma explicação do jogo para um iniciante ou há uma questão sobre a legalidade de uma jogada (CARRUTHERS, 2000, p.19).

Ambas as situações são corriqueiras e podem ser ampliadas para outras situações nas quais há pouca ou nenhuma resistência por parte do senso comum em admitir que há aspectos da mentalidade que não são conscientes. Sendo assim, parece haver também no vocabulário psicológico do senso comum algum espaço para o uso do predicado consciente atribuído a estados mentais.

Se simplesmente dizemos de um determinado estado mental que ele é consciente, então usamos “consciente” como um predicado unário, cujo termo é um estado mental. Esse uso de consciência equivale ao conceito chamado de *consciência intransitiva de estado*. Ao atribuir consciência nesse sentido do termo aos estados mentais, frequentemente buscamos apenas sustentar que tais estados mentais desfrutam de atributos distintos de outros estados mentais que não são conscientes. Basicamente, há dois sentidos em que se diz que estados mentais são intransitivamente conscientes (ou inconscientes):

Estados mentais são conscientes se o conteúdo desses estados está disponível para processos inferenciais, controle racional do comportamento ou para processos introspectivos e de expressão verbal. Pensamentos e percepções são exemplos paradigmáticos de estados mentais que são conscientes nesse sentido (e inconscientes se o conteúdo desses estados não

está disponível para processos inferenciais, controle racional do comportamento ou para processos introspectivos e de expressão verbal).

Estados mentais são conscientes se são experienciados de alguma maneira pelo sujeito que os tem. Nagel caracteriza este modo dos estados mentais serem conscientes da seguinte maneira: "[...] fundamentalmente, um organismo tem estados mentais conscientes se e somente se há algo como é *ser* esse organismo – algo como é *para* o organismo" (NAGEL, 1974, p. 436, tradução nossa, ênfases do autor). Exemplos paradigmáticos de estados mentais conscientes nesse sentido são nossas experiências sensoriais, mas também percepções e (sob uma interpretação) alguns pensamentos.

Cada um dos sentidos distinguidos acima corresponde a um conceito técnico distinto, na seção 1.1.3.1 examinaremos o conceito de *consciência de acesso*, que identifica a consciência à diferentes formas de disponibilidade cognitiva dos estados mentais. Na seção 1.1.3.2 examinaremos brevemente a noção de *consciência fenomênica*. A consciência fenomênica é rodeada de controvérsias. Há um consenso de que nós temos estados mentais que são experienciados subjetivamente ou "possuem algo como é" estar em um desses estados, mas para além desses fatos há pouco consenso sobre a natureza e funcionamento da consciência fenomênica[5]. Devido a sua ampla aceitação e uso, introduziremos essa noção através da concepção de Nagel (1974), mas anteciparemos também algumas qualificações uteis para discussões posteriores. As discussões encontradas no capítulo 3 se valem dessas caracterizações e distinções e aprofundam a temática da consciência fenomênica no contexto das teorias representacionalistas de segunda-ordem.

1.1.3.1 Consciência de acesso

O uso da noção de consciência no sentido (1) corresponde ao que Ned Block (1995) chama de *consciência de acesso*. *Consciência de acesso* é um conceito técnico de consciência funcionalmente definido (BLOCK, 2007, p.280). Ou seja, a identificação de um estado mental M com tendo a propriedade designada pelo conceito consciência de acesso tem como condição que M desempenhe determinados papeis causais e inferenciais no sistema cognitivo do sujeito. Segundo a definição de Block (1995, p.231), um estado é acesso-consciente quando seu conteúdo pode ser (i) *inferencialmente promíscuo (*essa expressão é devida a Stich, 2011, p.42):

[5] A título de qualificação, pode ser notado que nenhuma abordagem sobre a consciência nega que existam estados mentais que exibem uma fenomenologia. Mas há disputas sobre a riqueza e disponibilidade do conteúdo sensorial de diversos tipos de estados mentais conscientes (ver, DENNETT,

quando seu conteúdo está disponível para servir como premissa em raciocínio teórico ou em processos de formação de crença a partir de inferências; (ii) disponibilizado para o uso, controle e/ou planejamento racional de ações; ou (iii) disponível para expressão verbal.

Conforme esclarece Block (1995, p.231), um estado mental não precisa satisfazer todos esses requisitos de "disponibilidade" ou reportabilidade para que seja considerado um estado acesso-consciente. Com efeito, apesar da capacidade de um estado de ser verbalmente expresso ser um guia frequente em designs experimentais que visam investigar processos e estados conscientes, este é de longe o critério de menor peso na caracterização de um estado mental como acesso-consciente (BLOCK, 2007, p. 169).

Note que todas as três condições mencionadas acima são aspectos disposicionais de estados mentais. A consequência dessa caracterização da consciência de acesso é a de que ter a propriedade de ser acesso-consciente é algo que depende menos de aspectos "internos" aos estados mentais do que as relações que esse estado tem com o restante do sistema cognitivo do sujeito. Pois, considere um caso hipotético no qual um estado mental M em um organismo O_1 é capaz de manifestar todas as três disposições mencionadas. Nesse caso podemos atribuir a propriedade de ser acesso-consciente a M. Mas, quando "transplantada" para a mente de outro organismo O_2, o mesmo estado M poderia deixar de ser um estado acesso-consciente, caso não desempenhe nesse contexto os papéis causais e inferenciais relevantes. A ausência da propriedade de ser acesso-consciente nessa segunda situação é possível ainda que M não fosse alterado em nenhuma de suas propriedades constitutivas ou internas[6].

1.1.3.2 Consciência fenomênica

A noção de consciência no sentido caracterizado em (2) – isto é, no sentido de estados mentais experienciados como tendo uma perspectiva subjetiva e certas qualidades fenomenológicas – corresponde a noção de *consciência fenomênica*. O conceito de consciência fenomênica é um conceito que diz algo a respeito de como nossos estados mentais se apresentam para nós.

Ao introduzir o conceito de consciência fenomênica, Ned Block (1995) identifica o conceito com à noção popular de experiência. Nós temos diversos tipos de experiências, como

[6] Compare essa situação com a propriedade de realizar a função F que uma linha de código desempenha em um determinado contexto. Agora imagine que essa mesma linha de código é transferida para outro contexto, no qual ela é incapaz de executar F (por exemplo, porque nesse contexto uma linguagem de programação distinta é empregada). Internamente a linha de código nada é alterado durante a transferência, mas sua função é eliminada devido a diferenças no modo como ela se relaciona com as demais linhas de código do sistema.

experiências perceptuais, somatossensoriais e emocionais. O que todos esses tipos de experiência têm em comum, e o que as caracteriza enquanto experiências, é o fato de tê-las envolve "ter algo como é" (NAGEL, 1974, p. 436) para um organismo vivencia-los. A interpretação adequada da locução "ter algo como é" é motivo de controvérsias, veremos algumas delas no capitulo 3. Sendo o nosso propósito aqui introduzir uma caracterização inicial do conceito de consciência fenomênica, optaremos por ignorar momentaneamente essas disputas. Porém, são indispensáveis alguns comentários sobre a caracterização de Nagel da consciência fenomênica.

Desde o artigo seminal de Nagel (1974), tornou-se popular o uso da expressão "algo como é" (*something it is like* ou também *what-it-is-like-ness*) para se fazer referência aos estados mentais fenomênicamente conscientes. A versão completa da passagem de Nagel utiliza a expressão "algo como é" para identificar, através de condições necessárias e suficientes, quais estados mentais são conscientes em sentido fenomênico: "[...] fundamentalmente, um organismo tem estados mentais conscientes se e somente se há algo como é *ser* esse organismo – algo como é *para* o organismo" (NAGEL, 1974, p. 436, tradução nossa, ênfases do autor)[7].

Há dois aspectos da caracterização de Nagel que podem ser conceitualmente distinguidos (cf. CARRUTHERS, 2000, p. 128; KRIEGEL, 2009, p. 8; LEVINE, 2001, p. 6-7). Por um lado, estados mentais fenomênicamente conscientes exibem uma fenomenologia: um conjunto de qualidades como, por exemplo, parecer verde, frio ou molhado. Este aspecto é, por vezes, compreendido como a propriedade denotado pela expressão "algo como é" na passagem de Nagel. Há diversos conceitos cunhados na literatura para se referir a esse aspecto de nossos estados mentais fenomênicamente conscientes, entre os quais as noções de *caráter qualitativo* (Kriegel, 2009), *subjetividade mundana* (CARRUTHERS, 2000) e *propriedades sensoriais* (Rosenthal, 2005).[8]

O caráter qualitativo de uma experiência se refere às propriedades qualitativas (os *qualia*) ou fenomenológicas que um estado mental possui. Quando eu olho para o muro da casa à frente, o muro se parece para mim como exibindo um particular tom de azul, mas essa experiência se pareceria distinta para mim se eu não enxergasse cores ou se as condições de iluminação fossem diferentes. As diferenças entre as aparências que essas experiências teriam

[7] Um ponto de disputa difícil de ser ignorado, mas que frequentemente não é explicitamente tematizado é o fato de que a caracterização de Nagel da consciência fenomênica faz menção direta não aos estados mentais, mas ao sujeito. Nesse sentido há espaço para uma interpretação alternativa da consciência fenomênica como uma propriedade vinculada a organismos (i.e., como um tipo de consciência de criatura).

[8] Utilizamos os conceitos introduzidos por Kriegel (2006; 2009) meramente por conveniência. As caracterizações do caráter subjetivo e do caráter qualitativo adotadas nesta seção são independentes da teoria oferecida por Kriegel para explicá-las.

são diferenças nas propriedades que compõem o caráter qualitativo da experiência em cada caso[9].

Por outro lado, os estados mentais fenomênicamente conscientes segundo a passagem de Nagel, também se apresentam "*para o organismo"*. Para alguns, é em razão dessa apresentação de uma determinada qualidade ou fenomenologia "para o organismo" que os estados mentais adquirem sua perspectiva subjetiva, a perspectiva através da qual nós experienciamos, em primeira pessoa, determinados estados mentais. Mais uma vez, há diversos conceitos cunhados para designar essa característica dos estados mentais fenomênicamente conscientes, entre os quais se destacam as noções de *subjetividade experiencial* (Carruthers, 2000); *fenomenalidade espessa* (Rosenthal, 2002); e *caráter subjetivo* (Kriegel, 2009). A noção de um caráter subjetivo dos estados mentais (ou noções equivalentes) chama a atenção para a ideia de que organismos individuais ocupam uma perspectiva particular (i.e, subjetiva e em primeira-pessoa) na experiência de seus próprios estados mentais ocorrentes.

Conforme esclarece Kriegel (2009, p.1) há uma divisão de trabalho explanatório a ser desempenhado por cada uma dessas noções. O caráter qualitativo determina que *tipo de experiência* eu estou tendo – e.g, uma experiência visual do azul, do amarelo ou ainda uma experiência auditiva do som de um trovão ou uma experiência somatossensorial do calor. Já o caráter subjetivo pode ser considerado como o que caracteriza um estado mental enquanto um estado do *tipo experiencial*[10] (i.e, o tipo de estado que pode ser vivenciado como tendo qualidades auditiva, somatossensorial, etc. por alguém).

Em casos paradigmáticos de estados mentais com consciência fenomênica, tais estados co-instanciam o caráter subjetivo e o caráter qualitativo (KRIEGEL, 2009, p.53). Minha percepção consciente do azul do céu, por exemplo, consiste na instanciação simultânea da

---

[9] Kriegel (2009, p.8) caracteriza o caráter qualitativo de uma experiência como determinado pela soma ou conjunto das propriedades qualitativas que uma experiência instancia durante um tempo. Entretanto introduzir a noção desta maneira é conceder mais do que uma descrição do fenômeno a ser explicado, pois prevê que as propriedades qualitativas serão uma parte constitutiva do caráter subjetivo, se esse não puder ser instanciado independentemente do caráter qualitativo de uma experiência. Mais do que uma descrição do fenômeno, a caracterização de Kriegel acaba por retirar conclusões dedutivas sobre a constituição metafísica da consciência.

[10] Essa perspectiva parece corroborada pelo nosso uso do termo experiência em contextos cotidianos. Em nosso uso cotidiano do termo experiência, está implícita a ideia de que uma experiência é algo vivenciado por alguém. Assim como não há dores flutuando no vazio, não existem experiências somatossensoriais, auditivas, visuais ou de quaisquer outras modalidades que não pertençam a algum sujeito. Nosso uso do conceito parece ser, no entanto, ainda mais restritivo no sentido de que ele exige mais do que apenas um portador para as experiências. É preciso, além disso, que as qualidades experienciais de, digamos, uma dor seja vivenciada pelo sujeito que a tem. Do contrário, não estaríamos sequer dispostos a atribuir uma experiência de dor ao sujeito. Pois, considere como soa sem sentido dizermos que experimentamos uma dor, mas não a sentimos. Pode ser o caso de que sabemos que temos alguma lesão, mas se essa lesão não se apresenta para nós de uma maneira dolorosa, não há nessa situação nenhuma dor ocorrendo.

“azulidade” do céu, isto é, de uma propriedade qualitativa particular e do modo como essa qualidade particular se parece para mim.

São questões em aberto se, o caráter qualitativo e subjetivo são mais do que aspectos conceitualmente distintos, podendo também ser instanciados independentemente na mente de um organismo; se eles são aspectos explicáveis por um processo psicológico comum; e se algum desses aspectos é explanatoriamente mais relevante do que o outro para uma teoria da consciência.

1.1.4 Consciência de estado: Transitiva

Não parece haver uma maneira gramaticalmente correta de expressar em linguagem natural o uso do predicado “consciente” enquanto um predicado binário (de aridade dois) cujos termos são um estado mental e um evento, objeto ou estado. Não é gramatical, por exemplo, a sentença: “o pensamento de João está consciente de que está chovendo”. Esta constatação poderia levar à desconfiança de que não há um uso estabelecido no senso comum para o predicado consciente quando concebido como um predicado binário.

Contudo, o fato de que não há uma correspondência em linguagem natural para o uso do predicado binário “consciente” enquanto aplicado a estados mentais não deve ser considerado nem como um indício que não há um uso correspondente no senso comum, nem que tal predicado é supérfluo.

Nós não falamos normalmente de estados mentais que são conscientes de algo, mas essa não é o sentido que tem a noção de consciência transitiva de estado. Estados mentais de que se atribui consciência transitiva são estados mentais *através dos quais somos conscientes de algo*. Assim, por exemplo, podemos dizer que João está consciente de que está chovendo porque ele tem um pensamento ocorrente cujo conteúdo é a assertiva “está chovendo*”* ou porque ele tem uma percepção de um evento que aparente ser a chuva caindo lá fora. Quando utilizamos a noção de um estado mental consciente em sentido transitivo (consciência transitiva de estado) o aspecto de nossas vidas mentais que buscamos capturar é o de que um estado mental M nos permite estar conscientes de algo.

Kriegel oferece a seguinte definição do que é a propriedade denotada pelo conceito de consciência transitiva de estado: “podemos usar ‘consciência transitiva de estado’ para a propriedade que estados mentais exibem quando, e somente quando, seus sujeitos são transitivamente conscientes de algo em virtude de estar nesses estados” (KRIEGEL, 2009, p.27-28, nossa tradução, adaptado). Note que, apesar da referência aos indivíduos transitivamente

consciente na sua definição, a propriedade de fazer com que um sujeito seja consciente é ainda uma propriedade dos estados mentais. São os estados mentais transitivamente conscientes que tem a capacidade de fazer com que indivíduos sejam conscientes de algo e, por isso, trata-se de uma variedade de consciência de estado e não de criaturas.

Este é um ponto relevante, pois estados mentais que exibem consciência transitiva não precisam ser, eles mesmos, estados conscientes *para os sujeitos* (ainda que façam com que os sujeitos sejam conscientes de algo). Entretanto, consideramos que a terminologia é extremamente infeliz em esclarecer este ponto. O termo "consciência transitiva de estado" e sentenças como "M é um estado transitivamente consciente" sugerem que aquilo sobre o que somos conscientes é M e, portanto, que poderíamos esperar algum tipo de acesso consciente ou experiência do estado mental M. Porém, como sustenta Rosenthal (2006, p.4) "não é de nenhuma maneira óbvio que um estado deva, ele mesmo, ser consciente para fazer que alguém seja consciente de algo". Considere o rato que está consciente do gato via um estado mental perceptual – digamos que através de uma percepção olfativa. O que é conscientemente saliente para o rato é a presença do gato, devido a sua fenomenologia e devido à maneira como o comportamento do rato é diferencialmente modelado pelo reconhecimento da presença do gato. Mas, não é óbvio que para além da consciência da presença do gato, o rato esteja consciente de que está tendo uma percepção olfativa de um gato. Com efeito, parece ser uma exigência cognitiva demasiada para a percepção consciente do rato que ele deva ter algo como o conceito de uma experiência ocorrente ou de si mesmo como um sujeito de experiências.

Em nossa discussão sobre o conceito de consciência transitiva de criatura (seção 1.1.2) sugerimos que é possível substituirmos, em alguns contextos, o conceito de consciência transitiva de criatura pelo conceito de consciência transitiva estados mentais. Em particular sugerimos que a atribuição de consciência transitiva de criatura redunda em dizermos que um indivíduo tem certos estados mentais *sobre algo* e, portanto, esse uso do conceito implica que há um estado mental através do qual esse indivíduo é consciente. A partir da definição de Kriegel, fica claro que esta é uma consequência analítica: o que é para um indivíduo ser transitivamente consciente de algo é ter um estado mental que se relaciona apropriadamente com aquilo de que ele é consciente. E estados mentais transitivamente conscientes são precisamente os estados mentais através dos quais um sujeito é capaz de se relacionar com algo de tal forma que o sujeito se torne consciente desse algo. Assim, por definição, um indivíduo exibe consciência transitiva de criatura apenas se ele está em um estado mental que instancia a propriedade de consciência transitiva (de estado) (KRIEGEL, 2009, p. 29).

2 REPRESENTACIONALISMO DE SEGUNDA-ORDEM: UMA VISÃO GERAL

Este capítulo trata sobre uma abordagem filosófica que busca explicar a consciência (fenomênica, dos estados mentais) através da ideia de que mecanismos representacionais de segunda-ordem são uma condição necessária para que estados mentais sejam conscientes (em sentido fenomênico). Essa abordagem é conhecida como *representacionalismo de segunda-ordem*[11]. Nosso objetivo neste capitulo será oferecer uma caracterização geral da abordagem e uma avaliação crítica do seu argumento central.

Na seção 2.1 caracterizamos as abordagens de primeira-ordem e de segunda-ordem sobre a consciência, localizando seu principal ponto de divergência no modo como as teorias pertencentes à cada abordagem se posicionam sobre o chamado *princípio da transitividade*. Além disso, apresentamos brevemente uma versão de uma teoria representacionalista de primeira-ordem, de modo a apresentar uma teoria rival que, apesar de concordar com a proposta representacionalista, nega a adoção do princípio da transitividade como um ponto de partida para a teorização sobre a consciência. Na seção 2.2 discutimos as motivações para a adoção do princípio da transitividade como um ponto de partida para a teorização sobre a consciência. Na seção 2.3 apresentamos e discutimos um argumento que se fundamenta na adoção do princípio da transitividade para motivar a adoção de alguma versão do *representacionalismo de segunda-ordem*. Chamamos este argumento de *argumento mestre do representacionalismo de segunda-ordem*. Após apresenta-lo, discutimos algumas objeções dirigidas ao argumento mestre, especialmente com relação a verdade da premissa (1), que supõe a verdade do princípio da transitividade. Na seção 2.4 expandimos nossa discussão sobre como o princípio da transitividade é explorado na formulação do *explanans* da abordagem de segunda-ordem, detalhando alguns aspectos de diferentes versões de teorias pertencentes à abordagem de segunda-ordem. Por fim, na seção 2.5 discutimos e revisamos algumas das motivações empíricas que corroboram com a estratégia explicativa da abordagem de segunda-ordem sobre a consciência fenomênica. Além de apresentar algumas das evidências empíricas que corroboram com as hipóteses do representacionalismo de segunda-ordem, essa seção evidencia as diferenças nas previsões empíricas dessa abordagem em contraste com as abordagens rivais. O resultado dessas discussões é uma defesa e crítica amplas sobre a estratégia explicativa do representacionalismo de segunda-ordem.

---

[11] Ao longo do capítulo, alternamos entre o uso dos rótulos “abordagem de segunda-ordem”, “teorias de segunda-ordem” e “representacionalismo de segunda-ordem”. Todos esses rótulos se referem ao mesmo grupo de teorias.

2.1 ABORDAGENS DE PRIMEIRA-ORDEM E SEGUNDA-ORDEM SOBRE A CONSCIÊNCIA FENOMÊNICA

A divisão entre abordagens de primeira-ordem e segunda-ordem divide as teorias sobre a consciência fenomênica em dois grandes grupos. Teorias pertencentes ao grupo da abordagem de segunda-ordem formam um conjunto amplamente homogêneo de teorias, na medida em que elas compartilham das mesmas motivações teóricas e estratégia explicativa. Todas as teorias pertencentes à abordagem de segunda-ordem são motivadas pelo que Rosenthal (2005, p.4) chama de *princípio da transitividade*, segundo o qual estados mentais conscientes são estados de que um sujeito está, de alguma maneira, consciente. E todas teorias pertencentes à abordagem de segunda-ordem sustentam como estratégia explicativa que um estado mental é consciente porque ele implementa, de alguma maneira apropriada, o princípio da transitividade.

Em contrapartida, teorias pertencentes à abordagem de primeira-ordem são largamente heterogêneas com relação às motivações teóricas e as estratégias explicativas adotadas (GOTTLIEB, 2018, p.320). Teorias pertencentes à abordagem de primeira-ordem podem ser dualistas de propriedade (e.g., CHALMERS, 1997); realistas ingênuos (MARTIN, 2004) e representacionalistas (TYE, 1995; DRETSKE, 1995). O que une essas teorias sob a mesma abordagem é, sobretudo, um posicionamento contrário às motivações teóricas e estratégia explicativa da abordagem de segunda-ordem. Particularmente, elas são unidas pela crença de que são falsas ou injustificadas as motivações e estratégia explicativa característica da abordagem de segunda-ordem. Em outras palavras, teorias pertencentes à abordagem de primeira-ordem estão unidas pela crença de que é falso que a implementação do princípio da transitividade é uma condição necessária para que um estado mental seja fenomênicamente consciente.

Sendo assim, a abordagem de primeira-ordem é melhor definida negativamente (adaptado de GOTTLIEB, 2016, p. 353): uma teoria T é uma teoria de primeira-ordem da consciência se, de acordo com T, um estado mental M satisfazer o princípio da transitividade não é uma condição necessária para que M seja consciente. Dretske sintetiza nas seguintes palavras o posicionamento contrário das teorias pertencentes à abordagem de primeira-ordem ao princípio da transitividade:

> Estados mentais conscientes – experiências, em particular – são estados *com os quais* nós somos conscientes, não estados *dos quais* somos conscientes. Eles são estados que nos fazem conscientes, não estados que nós tornamos conscientes por estarmos conscientes deles (DRETSKE, 1995, p. 100-101, tradução nossa, itálicos do autor).

Assim como outros defensores de abordagens de primeira-ordem, Dretske rejeita a tese de que é uma condição necessária para que estados mentais sejam fenomênicamente consciente que sejamos, de alguma maneira apropriada, conscientes desses estados. Porém, ao rejeitar esta tese e sustentar que estados mentais conscientes são aqueles estados através dos quais nos tornamos conscientes (*aware*) de algo (i.e., o objeto representacional do estado mental), Dretske e outros que rejeitem o princípio da transitividade, enfrentam uma dificuldade em explicar casos de percepção subliminar entre outros casos envolvendo percepções inconscientes (ROSENTHAL; WEISBERG, 2008; ver também, CARRUTHERS 2000).

O problema pode ser formulado da seguinte maneira. Percepções subliminares são percepções através das quais um sujeito é capaz de perceber algo, digamos, um estimulo visual brevemente apresentado, mas sobre o qual o sujeito relata não ter consciência de ter experienciado. Mas, apesar de não haver uma experiência associada ao estímulo, os estados subliminares desempenham a maioria das funções que uma percepção consciente desempenharia, como influenciar o comportamento subsequente do sujeito (RO et al, 2009). Em casos como esse, o que constitui uma dificuldade para teorias de primeira-ordem é oferecer uma distinção funcional adequada entre um estado subliminar e um estado consciente que não apele para alguma versão do princípio da transitividade (ROSENTHAL, 2012, p. 1425)

Com efeito, a resposta de Dretske (2006) a esse problema pode ser interpretada como um uso encoberto do princípio da transitividade. Em sua resposta ao problema, Dretske estipula que a diferença crucial entre estados subliminares e estados conscientes é que, nos casos de estados mentais conscientes, os sujeitos podem mencionar o conteúdo de seus estados mentais como justificação para uma determinada ação, o que não ocorre com os estados subliminares ou inconscientes. Mas, como nota Rosenthal (2012, p.1425), um sujeito não pode citar como justificação o conteúdo de um estado mental de que ele não está consciente, portanto, a solução de Dretske para a demarcação entre estados mentais conscientes e inconscientes/subliminares apela para uma versão do princípio da transitividade.

2.1.1 **Representacionalismo de primeira-ordem e de segunda-ordem**

Em geral, teorias de segunda-ordem são teorias representacionalistas sobre a consciência. Teorias representacionalistas, sejam elas de primeira-ordem ou segunda-ordem, são teorias que visam explicar a consciência dos estados mentais em termos de estados intencionais e o papel que o conteúdo representacional desses estados desempenha no sistema cognitivo de um organismo (CARRUTHERS, 2000).

Eventos mentais são usualmente concebidos como possuindo funções representacionais, parcialmente porque eles são causados por aquilo que visam representar (GENNARO, 2018, p. 107). Por exemplo, experiências visuais servem a função de representar como o mundo é lá fora (ou ao menos como a luz interage com a superfície dos objetos e com nossos olhos) e são causadas pela interação da luz com nosso sistema visual, nos casos de experiências visuais verídicas. Estados mentais desse tipo são também chamados estados *intencionais*. A intencionalidade de estados mentais é a propriedade de serem sobre a algo, como a percepção de uma árvore é sobre um objeto externo e a sensação de uma dor é, plausivelmente, sobre um estado corporal do sujeito.

Por vezes, estados intencionais são contrastados com estados fenomênicos – os estados tipicamente associados às experiências sensoriais. Estados fenomênicos são estados que possuem algo como se parece tê-los: há algo como é morder uma pimenta e há algo (muito distinto do primeiro caso) com é estar com frio. O que os estados fenomênicos têm em comum, portanto, é certa dimensão qualitativa de que somos capazes de falar apenas aproximativamente, mas cujo conhecimento direto exige ter tido esta experiência, ou ser apto a ter experiências suficientemente semelhantes. Mas, apesar do contraste entre estados fenomênicos e estados intencionais, muitos estados mentais possuem tanto propriedades fenomênicas como intencionais, tais como experiências visuais. Sendo assim, uma motivação inicial para o representacionalismo consiste em notar que estados mentais conscientes frequentemente possuem conteúdos representacionais, como no caso de estados perceptuais (GENNARO, 2018, p. 107).

O representacionalismo sustenta que uma parte relevante de como e/ou porquê estados mentais conscientes se parecem, subjetivamente, de certa maneira para nós está vinculado ao conteúdo representacional que eles possuem. As abordagens representacionalistas (em geral) reconhecem uma distinção entre duas variedades de estados representacionais através dos quais podemos ser conscientes de algo: estados de primeira-ordem e estados de segunda-ordem.

Estados mentais representacionais de primeira-ordem são, paradigmaticamente, estados direcionados a objetos, eventos e estados que residem fora de nossas mentes (i.e., que representam algo no mundo ou no corpo dos indivíduos). Exemplos de estados de primeira-ordem são o meu pensamento sobre a chuva ou a percepção da tonalidade cinzenta do céu. Em contraste, estados representacionais de segunda-ordem são aqueles que são sobre – isto é, que tem como objeto representacional – nossos próprios estados mentais de primeira-ordem. Exemplos mais tipicamente associados aos estados mentais de segunda-ordem são pensamentos introspectivos, como no caso em que eu penso sobre as minhas próprias crenças ou sobre meus

estados perceptuais ocorrentes[12]. Porém, pensamentos introspectivos são, usualmente, conscientes, de modo que eles não poderiam ser evocados nas explicações das teorias de segunda-ordem.

A ênfase no papel explanatório concedido a um ou outro desses tipos de estados intencionais, quando se trata de explicar porque há estados que são intransitivamente conscientes, é um dos fatores centrais na diferenciação entre abordagens representacionalistas de primeira-ordem e de segunda-ordem. Os representacionalistas de primeira-ordem e segunda-ordem disputam sobre o papel que os estados de primeira e segunda-ordem desempenham numa explicação da consciência em sentido intransitivo.

Representacionalistas que adotam uma abordagem de primeira-ordem sustentam que a posse de estados mentais intencionais de primeira-ordem, que cumprem um determinado papel funcional no sistema cognitivo, são condição suficiente para a consciência intransitiva de estado e em particular, a consciência fenomênica. Sob a perspectiva da explicação oferecida por teorias de primeira-ordem, portanto, o que explica que alguns estados mentais sejam fenomênicamente conscientes são fatos a respeito do conteúdo representacional dos estados de primeira-ordem de um indivíduo.

Em contraste, teorias representacionalistas de segunda-ordem sustentam que apenas os fatos a respeito do conteúdo representacional dos estados mentais de primeira-ordem não são suficientes na formulação de uma teoria satisfatória a respeito da consciência fenomênica. Adicionalmente é necessário que apelemos para a postulação de algum tipo de relação com um estado intencional de segunda-ordem – ou seja, estados mentais que representam estados mentais de primeira-ordem (CARRUTHERS, 2000, p.24-25).

2.1.1.1 A teoria PANIC de Michael Tye

A teoria PANIC, articulada por Tye (1995; 2000; 2005), é um dos principais exemplos de teorias pertencentes ao representacionalismo de primeira-ordem[13]. Sua teoria chama-se

[12] Cabe notar, no entanto, que ao sugerir que estados representacionais de segunda-ordem são uma condição necessária para a consciência fenomênica, o representacionalismo de segunda-ordem não está sustentando que a faculdade da introspecção é necessária para a consciência fenomênica. Conforme veremos mais adiante, há uma distinção crucial entre os estados representacionais de segunda-ordem relevantes para a tese da presente abordagem e os estados mentais introspectivos em sentido usual.

[13] Além dele, outros autores que desenvolvem teorias representacionalistas de primeira-ordem são DRETSKE, 1995 e KIRK, 1994; HILL, 2011.

PANIC, em virtude do acrônimo em inglês para as condições necessárias e suficientes para que um estado mental seja fenomênicamente consciente[14].

Na teoria de Tye, os estados mentais conscientes são aqueles que cumprem com as seguintes quatro condições: estados mentais são conscientes, se e somente se, eles têm (1) conteúdo intencional que é (2) abstrato, (3) não conceitual e (4) "disponível" (*Poised*). Ao sugerir que ao menos alguns dos conteúdos representacionais de estados mentais são não conceituais, Tye (1995, p.139) sustenta que há conteúdos intencionais para os quais os sujeitos de experiência não possuem conceitos correspondentes às propriedades representadas. Uma maneira de compreender essa condição é a tese de que o conteúdo dos estados mentais que são fenomênicamente conscientes são determinados de uma maneira que nossos conceitos não são. Em outras palavras, possuímos conceitos para os determináveis do conteúdo intencional de estados mentais fenomênicamente conscientes, mas não possuímos conceitos para os valores determinados desses determináveis. Por exemplo, possuímos na linguagem ordinária o conceito "vermelho", mas não o conceito "vermelho$_{45}$"[15], que indica precisamente a tonalidade de vermelho apresentada pelo sol na tarde do domingo passado (ver, TYE, 1995, p. 139). Sendo assim, o modo como são representados aquilo que é objeto dos estados intencionais fenomênicamente conscientes não são idênticos ao modo como representaríamos esses mesmos objetos caso utilizássemos conceitos da linguagem natural.

Ao sugerir que parte deste conteúdo é *abstrato*, a proposta de Tye é abrir espaço para casos de percepção não verídica ou alucinações, ou seja, casos nos quais o conteúdo representacional não é causalmente ligado a entidades particulares existentes. Dizer que o conteúdo dos estados mentais fenomênicamente conscientes são abstratos pode ser interpretado como afirmando que objetos, eventos ou propriedades particulares não são uma parte constitutiva do conteúdo intencional desses estados. Conforme esclarece Tye (2000, p.62), o que é crucial para como se parecem para nós as experiências que temos é "a representação de propriedades e características gerais".

Por fim, o aspecto mais relevante na explicação para o que torna certos estados mentais conscientes se localiza na caracterização do conteúdo destes estados como disponíveis (Poised). Conforme nota o autor, a condição da "disponibilidade" (*poisedness*) do conteúdo dos estados mentais é, essencialmente, um papel funcional desempenhado pelos estados mentais (TYE, 2000, p.62). O papel funcional desempenhado por conteúdos "disponíveis" é o de estarem de

---

[14] PANIC é um acrônimo para *Poised, abstract, non-conceptual, intentional content* (em uma tradução livre: conteúdos disponíveis, abstratos, não conceituais e intencionais).

[15] O subscrito em "vermelho$_{45}$" serve a função de indicar um tom determinado do determinável "vermelho".

prontidão para causar um impacto à formação de crenças e/ou desejos (TYE, 2000, p.62). Assim, por exemplo, sentir-se faminto é um estado que possui um conteúdo não conceitual que representa um determinado estado do corpo do sujeito e que tem um efeito direto na formação de desejos do sujeito, a saber, o desejo de comer.

Em contraste com abordagens de segunda-ordem sobre a consciência, a teoria PANIC e outras teorias de primeira-ordem não sustentam que devemos ser apropriadamente conscientes de nossos estados mentais para que eles sejam considerados estados mentais conscientes. O que isso significa, em outras palavras, é que estados mentais conscientes não são, necessariamente, estados dos quais um sujeito é tem ou é capaz de ter metarepresentações a respeito. Considere, por exemplo, o caso de um bebê que sente dor devido ao nascimento dos seus primeiros dentes. Plausivelmente, há algo como é para esse bebê sentir essa dor, há uma fenomenologia associada ao evento mental da dor. Porém, é improvável que esse bebê tenha os recursos cognitivos necessários para refletir introspectivamente sobre o evento mental da dor enquanto um dos seus estados mentais ocorrentes. Tendo observações como essa em vista, teorias de primeira-ordem sustentam, como vimos na passagem de Dretske citada anteriormente, que estados mentais conscientes são estados através dos quais somos conscientes de algo (e.g., a dor) e não estados de que precisamos estar conscientes para que eles se pareçam de certa maneira.

2.1.1.1.1 Objeções à teoria PANIC

Uma objeção que pode ser levantada contra a estratégia explicativa da teoria PANIC diz respeito a incapacidade da teoria de explicar, via qualquer um dos critérios elencados pela teoria, o que diferencia estados mentais conscientes de estados mentais inconscientes. Com relação à condição da “disponibilidade”, a seguinte crítica pode ser levantada: estar disponível ou pronto para causar algum efeito sobre a formação de crenças e desejos de um indivíduo é uma propriedade disposicional do conteúdo de um estado mental. Porém, a consciência fenomênica é, intuitivamente, uma propriedade atual dos estados mentais. Portanto, ter uma disposição para causar certos efeitos no sistema cognitivo de um individuo (e.g., produzir crenças ou desejos), não parece o tipo adequado de propriedade que diferencia estados mentais conscientes de estados mentais inconscientes.

Carruthers (2000, p. 170) põe a questão nos seguintes termos: como pode o mero fato de que um estado mental tem a disposição para causar um impacto sobre as crenças e desejos no sistema cognitivo de um indivíduo conferir a esse mesmo estado a propriedade de ter algo como é para o sujeito experienciá-lo, que é característica dos estados mentais fenomênicamente

conscientes? Com efeito, outros estados mentais com conteúdo representacional, nomeadamente, as crenças e desejos de um sujeito, também tem essa mesma disposição. Porém, ainda que possam ser estados (acesso) conscientes, não dizemos que há algo como é crer que 2 + 2 = 4 e, de modo mais geral, não assumimos que crenças tem um caráter fenomênico próprio. Tendo isso em vista, como a teoria PANIC e outras teorias representacionalistas de primeira-ordem poderiam explicar a diferença entre os estados mentais que tem algo como é experienciá-los (i.e., que são fenomênicamente conscientes) e estados que são conscientes, mas que não possuem algo como é experienciá-los (i.e., estado que são conscientes em outro sentido do termo)?

Uma possível reposta consiste em apelar para o tipo particular de conteúdo que os estados mentais fenomênicamente conscientes possuem. Nomeadamente, segundo a teoria PANIC, conteúdo *abstrato* e *não conceitual*. Entretanto, Kriegel (2002) oferece objeções a ideia de que apenas o apelo ao caráter abstrato ou não conceitual dos estados mentais seja suficiente para explicar satisfatoriamente o que distingue estados mentais fenomênicamente conscientes de estados mentais que não são fenomênicamente conscientes.

Segundo Kriegel (2002, p.58), o caráter abstrato do conteúdo dos estados mentais oferece um candidato inadequado para a diferenciação entre estados mentais fenomênicamente conscientes e estados que não são ou não podem ser fenomênicamente conscientes. Tye introduza noção de conteúdo abstrato com o propósito de permitir casos de erros representacionais, como em alucinações ou ilusões, em que o conteúdo representado não corresponde aos particulares no mundo que são alvo da representação. Porém, note que estados mentais que são representacionais, como crenças, também conseguem lidar com casos de erro representacional (e.g., casos nos quais um sujeito crê que P, quando P não é o caso). Além disso, o conteúdo de crenças também é abstrato, na medida em que proposições são entidades abstratas. Sendo assim, conclui Kriegel (2002, p.58) o caráter *abstrato* do conteúdo dos estados mentais não pode explicar a distinção entre os estados mentais que são fenomênicamente consciente e aqueles que não são.

O caráter não conceitual dos estados mentais, por sua vez, parece um candidato mais adequado na diferenciação dos estados mentais fenomênicamente conscientes e estados mentais que não são fenomênicamente conscientes, especialmente no caso das atitudes proposicionais como as crenças. O conteúdo das crenças é obviamente conceitual. Tendo isso em vista, Tye (1995, p. 138) sustenta que sua teoria implica em dizer que crenças não podem ser fenomênicamente conscientes, porque apenas o conteúdo não conceitual pode ser assim classificado, de modo que crenças simplesmente não possuem o tipo apropriado de conteúdo.

Porém, Kriegel (2002, p.59) nota o seguinte ponto a respeito da proposta de conteúdos não conceituais: o fato de que um sujeito pode experienciar, digamos, vermelho$_{17}$ sem possuir o conceito de vermelho$_{17}$ não mostra que o conteúdo dessa percepção deve ser não conceitual. No máximo, isso demonstra que se deve permitir que o conteúdo seja não conceitual para que o sujeito tenha a experiência de vermelho$_{17}$. Pois, seria contraintuitivo negar que um sujeito que tem o conceito de vermelho$_{17}$ é incapaz de experienciar a propriedade designada pelo conceito de vermelho$_{17}$ meramente em virtude de possuir esse conceito. Posto isso, a diferença entre estados mentais fenomênicamente conscientes e os que não são conscientes não pode residir na natureza não conceitual do conteúdo dos estados mentais. Não parece fazer sentido atribuir às limitações cognitivas de que dispomos, como no caso de não termos a capacidade de armazenar conceitos suficientemente refinados para lembrar e empregar em diferentes tempos a noção de vermelho$_{17}$, que dá origem à consciência fenomênica.

Em conclusão, a teoria PANIC parece incapaz de oferecer uma diferenciação entre estados mentais conscientes e inconscientes apenas baseada nos critérios elencados pela teoria como necessários e suficientes para que estados mentais sejam conscientes. Se nem a "disponibilidade", o caráter abstrato ou o caráter não conceitual do conteúdo intencional dos estados mentais são suficientes na diferenciação entre estados mentais conscientes e inconscientes, então devemos apelar para algum elemento externo aos inicialmente previstos pela teoria para explicar essa distinção. Uma alternativa plausível é oferecida pelo princípio da transitividade, sustentado pela abordagem de segunda-ordem.

2.2 PRINCÍPIO DA TRANSITIVIDADE: UMA INTUIÇÃO PRÉ-TEÓRICA ACEITÁVEL?

Vimos até aqui que o princípio da transitividade constitui uma parte fundamental do *explanans* das teorias pertencentes à abordagem de segunda-ordem sobre a consciência. Mais exatamente, sugerimos que a estratégia explicativa comum dessas abordagens consiste em oferecer uma hipótese sobre como o princípio da transitividade é implementado. Ou seja, trata-se de uma explicação que consiste em dizer como se dá a implementação do processo através do qual um sujeito é consciente de um dos seus estados mentais e de que maneira isso implica na consciência intransitiva do estado mental do qual o sujeito está (transitivamente) consciente. Mas, porque assumir, em primeiro lugar, que algo como o princípio da transitividade é um bom ponto de partida para a formulação de uma teoria sobre como estados mentais são fenomênicamente conscientes? Nesta seção, discutimos o que torna o princípio da transitividade um ponto de partida interessante na formulação de uma teoria da consciência.

2.2.1 Plausibilidade pré-teórica

Weisberg (2020) elenca três motivações para a adoção do princípio da transitividade como um ponto de partida para a formulação de uma teoria da consciência fenomênica. A primeira motivação, já mencionada no início desse capítulo, é o fato de que o princípio da transitividade é pré-teoricamente intuitivo, ou seja, a afirmação contida no princípio da transitividade parece verdadeira independentemente de nosso compromisso com teorias de segunda-ordem sobre a consciência. Além disso, o princípio da transitividade parece expressar uma noção tacitamente aceita pelo senso comum sobre a consciência, especialmente se consideramos sua formulação contrapositiva: estados mentais de que não somos apropriadamente conscientes não são estados mentais conscientes. Deste modo, o princípio da transitividade pode ser considerado como a expressão de como o senso comum concebe a noção de consciência fenomênica, ou pelo menos como a expressão de uma das concepções disponíveis no senso comum sobre a consciência.

Sendo assim, se um dos pontos de partida para a teorização sobre a consciência está em ter uma análise aproximada de como o entendemos o fenômeno da consciência e o conceitualizamos em contextos de senso comum[16], então o princípio da transitividade pode ser visto como uma tentativa de oferecer uma análise aproximada de uma das maneiras que o senso comum utiliza a noção de consciência. Por conseguinte, a partir da adoção do princípio da transitividade como ponto de partida teórico, as teorias de segunda-ordem podem formular hipóteses sobre quais são os mecanismos psicológicos que implementam o fenômeno natural que é designado pela análise da noção de consciência contida no princípio da transitividade.

2.2.2 Evidência fenomenológica

Uma segunda razão para assumirmos o princípio da transitividade como um ponto de partida para a teorização sobre a consciência advém de uma observação fenomenológica. Quando refletimos sobre como nos parecem nossos estados mentais fenomênicamente

[16] Podemos compreender, à luz de uma proposta funcionalista empírica, por exemplo, que a compreensão que o princípio da transitividade oferece do conceito de “consciência” utilizado pelo senso comum oferece uma maneira de fixar o referente desse fenômeno natural. Deste modo, o princípio da transitividade serve de pode servir de ponto de partida para a teorização da consciência como uma espécie de fixador do *explanandum* da teoria. Para uma discussão mais completa sobre a proposta funcionalista empírica, ver BRADDON-MITCHELL; JACKSON, 2007, p.84-87; LEWIS, 1972.

conscientes, podemos notar que todos esses estados têm em comum a característica de se apresentarem como envolvendo uma espécie de autoconsciência, tal como os estados mentais conscientes são caracterizados pelo princípio da transitividade. Assim, por exemplo, quando tenho a percepção das palavras na tela do computador, essa percepção é como que acompanhada por um tipo de reconhecimento de mim mesmo como o sujeito que percebe as palavras na tela. De maneira geral, todas as minhas percepções e sensações parecem ter essa característica de serem acompanhadas por um reconhecimento de que as sensações e percepções em questão são *minhas* percepções e sensações. Nesse sentido, Zahavi e Kriegel (2015) sugerem que, a característica dos estados mentais fenomênicamente conscientes necessariamente envolvem um tipo de referência ao sujeito de experiência (o que os autores chamam de *for-me-ness*)[17].

Se admitirmos que tal referência ao *self* é uma parte constituinte de todas as nossas percepções e sensações conscientes, então pode-se argumentar que ao adotarmos o princípio da transitividade como ponto de partida para a teorização da consciência, assumimos um posicionamento inicial que tem a virtude de ser coerente com uma crença independentemente admitida[18].

2.2.3 Utilidade teórica/metodológica na investigação empírica

Por fim, uma terceira razão para adotar o princípio da transitividade como um ponto de partida para a teorização sobre a consciência é o fato de que o assumir pode ser metodologicamente útil na investigação empírica da consciência. Se assumimos que estados mentais conscientes são estados mentais de que um sujeito está apropriadamente consciente de ter, então podemos esperar que os relatos do sujeito possam servir como uma fonte autoritativa sobre se um determinado estado mental é ou não consciente. Particularmente, podemos tomar como evidência de que um estado mental é consciente o fato de que o sujeito é capaz de relatar o modo como se parece ter esse estado verbalmente ou através da realização de tarefas não verbais de discriminação. Nesse sentido, assumir o princípio da transitividade pode ser vantajoso para o design experimental na investigação empírica da consciência.

---

[17] Para esclarecimentos sobre os compromissos envolvidos com a autorreferência ou autoconsciência que alegamos ser presente em todas percepções e sensações, bem como respostas a objeções, ver ZAHAVI; KRIEGEL, 2015. Para argumentos em defesa da proposta de que a auto referencialidade ou autoconsciência é um elemento das experiências, ver KRIEGEL, 2009.

[18] A virtude de uma teoria ou, no caso em questão de um ponto de partida para a formulação de uma teoria de ser compatível com uma crença independentemente admitida é chamada de coerência universal. Para uma elucidação e sumário das virtudes teóricas, ver, KEAS, 2017.

Cabe notar, no entanto, que assumir o princípio da transitividade não precisa implicar na proposta, criticada por Block (2007) de que apenas estados mentais reportáveis são conscientes. O princípio da transitividade é compatível com a ideia de que um sujeito pode perder todas as capacidades de reportar estar em um determinado estado mental sem, por isso, perder a capacidade de estar consciente, de uma maneira apropriada, de um estado mental.

Um segundo aspecto metodológico sobre o princípio da transitividade pode também ser elencado como uma razão para adotá-lo como ponto de partida para teorização da consciência. Quando se trata de fornecer uma teoria da consciência, devemos ter em vista a compatibilização de crenças provindas de diferentes perspectivas sobre esse fenômeno. Isto é, devemos levar em consideração como se compatibilizam nossas crenças e pressupostos filosóficos, de senso comum e científicos sobre a consciência. Nesse sentido, o princípio da transitividade pode exibir uma espécie de virtude teórica, já que ele é compatível (ou pelo menos não é obviamente incompatível) com aspectos de nossa visão de mundo de senso comum, filosófica e científica (ver nota 12). Sendo assim, o princípio da transitividade é um bom candidato para servir como um ponto de partida para a teorização sobre a consciência.[19]

2.2.4 Qualificações ao princípio da transitividade

Vimos anteriormente que o princípio da transitividade constitui um bom ponto de partida para a teorização sobre a consciência. Contudo, ainda que a ideia de que estados mentais conscientes são estados mentais de que um sujeito está apropriadamente consciente seja plausível como uma aproximação inicial sobre o fenômeno a ser investigado por uma teoria da consciência, há objeções obvias que poderiam falsificar esse princípio. Considere, por exemplo, a constatação de que, por vezes, temos estados mentais de que não estamos conscientes, mas cujos efeitos não só são observáveis por outras pessoas, mas que também podem ser inferidos, com base nessas observações e atribuídos a nós. Por exemplo, considere o caso em que eu estou irritado com a revisão de um texto, mas devido à falta de atenção ao meu próprio comportamento, não estou consciente de que estou irritado. Considere agora que um amigo próximo é capaz de inferir, com base no meu comportamento, que estou irritado e, ao notar esse fato, me diz o seguinte: “Você parece irritado hoje. O que acha de sairmos para fazer uma pausa e beber um café?”. Após ser informado por meu amigo, eu posso tomar conhecimento de que

[19] Uma quarta razão em virtude da qual é plausível assumirmos como um ponto de partida para a teorização sobre a consciência é o fato de que essa proposta ostenta uma longa tradição de simpatizantes, como Aristóteles, Locke, Brentano, William James e Sartre (ver, GÜZELDERE, 1995).

estou em um estado mental de irritação, mas devido a própria natureza desse estado, suponha que eu negue honestamente que estou irritado. Nessas circunstâncias, devemos considerar meu estado mental de estar irritado como fenomênicamente consciente? Este é um caso complexo para teorias que admitem o princípio da transitividade. Poder-se-ia pensar que o princípio da transitividade implica em uma resposta positiva à questão. Meu estado mental de estar irritado é fenomênicamente consciente porque eu estou transitivamente consciente desse estado, via o relato do meu amigo de que eu pareço irritado. Porém, esse resultado parece ser contraintuitivo, já que eu mesmo nego estar irritado e, portanto, plausivelmente negaria que me sinto irritado.

Considere ainda um segundo exemplo. Suponha que após ver um vídeo em que eu caio de bicicleta na infância eu infiro que eu senti dor. Nessa situação, uma teoria representacionalista de segunda-ordem deveria assentir que eu estou fenomênicamente consciente da dor que senti na infância, pois ela é alvo de um estado representacional de segunda-ordem. Mas é intuitivamente evidente que, no momento em que vejo o vídeo, eu não estou sentido a dor que senti na infância ao cair de bicicleta.

Em virtude de casos como esses, no qual um sujeito se torna (transitivamente) consciente de um estado mental via um processo que não parece ser adequado para que esse sujeito venha a ter um estado mental consciente (em sentido intransitivo), certas qualificações devem ser introduzidas ao princípio da transitividade.

Há um consenso entre proponentes de teorias de segunda-ordem de que o processo através do qual o sujeito é transitivamente consciente não pode ser mediado por inferências. O que motiva esse posicionamento é o consentimento com a intuição pré-teórica de que parecemos ter um acesso imediato e privilegiado de nossos próprios estados mentais ocorrentes. Isto é, nenhum esforço atencional, inferência, ou observação via algum dos nossos sentidos são necessários para que tenhamos acesso ao que pensamos e sentimos (ROSENTHAL, 2005, p.57). Essa intuição se resume à ideia de que existe uma assimetria entre a maneira como nós sabemos o que se passa em nossa mente, ou seja, como reconhecemos nossos próprios estados mentais conscientes, e a maneira como nós sabemos o que se passa na mente alheia. Quando temos a experiência de uma dor, por exemplo, esse estado se apresenta de uma maneira que não parece necessitar a mediação de nenhum tipo de observação sobre o nosso próprio comportamento ou qualquer tipo de tomada de atenção especial. Em contraste, se buscamos saber se alguém está com dor, é preciso que extrapolemos esse conhecimento a partir de observações sobre o comportamento dessa pessoa. Portanto, se é o caso de que nós parecemos gozar de um acesso direto e privilegiado de nossos próprios estados mentais, então uma

explicação sobre como nossos estados mentais se tornam conscientes deve ser compatível com esse dado fenomenológico.

Entretanto, autores como Rosenthal assumem qualificações sobre a recusa de processos inferenciais como a maneira através da qual estados mentais de primeira-ordem vem ser vinculados com estados mentais segunda-ordem. A qualificação introduzida por Rosenthal (1993, p.157) é a de que, o que importa é não haver uma inferência *consciente* entre os estados de segunda-ordem e os estados de primeira-ordem. Nesse sentido, Rosenthal abre espaço para casos nos quais os estados de primeira-ordem vêm a ser representados por estados de segunda-ordem em virtude de inferenciais inconscientes. Não obstante essa qualificação, não é preciso abrir mão da intuição pré-teórica de que temos acesso direto e privilegiado aos nossos próprios estados mentais. Pois, se um estado mental de segunda-ordem tem como conteúdo representacional um estado de primeira-ordem como produto de uma inferência inconsciente, ainda assim se parecerá para nós que essa vinculação se deu através de um processo imediato, preservando as aparências fenomenológicas (ROSENTHAL, 1993, p. 157).

2.3 O ARGUMENTO MESTRE DO REPRESENTACIONALISMO DE SEGUNDA-ORDEM:

Segundo a proposta das teorias representacionalistas de segunda-ordem sobre a natureza da consciência de estados mentais, um estado mental M é consciente porque o sujeito tem um estado mental M* que representa M de uma maneira apropriada. Assim, analogamente ao modo como ter consciência do ruído da geladeira é uma questão de ter uma percepção sobre esse ruído, ter consciência de um dos nossos próprios estados mentais é uma questão de ter um estado representacional (uma percepção ou um pensamento, por exemplo) cujo objeto representacional é um de nossos próprios estados mentais. Nesse sentido, o que os defensores de teorias representacionalistas de segunda-ordem sustentam é uma explicação da consciência de estados mentais em termos do que chamamos de consciência transitiva de estados mentais (capitulo 1, seção 1.1.4) e da ideia de que ser transitivamente consciente de algo consiste em representar esse algo.

Em amparo a essa proposta, defensores de teorias representacionalistas de segunda-ordem têm sugerido argumentos como o seguinte (ver, LYCAN 2001; ROSENTHAL, 2005, p. 109):

(1) Um estado consciente$_i$ é um estado mental do qual o sujeito está *consciente$_t$ de ter/estar*. (estipulação).

(2) O “de” em “*consciente$_t$ de ter/estar*” é indicativo de intencionalidade; o sujeito é consciente$_t$ de um objeto intencional.

(3) Intencionalidade consiste em representação; um estado tem algo como seu objeto intencional apenas se ele representa esse algo.

(C1) Ter consciência$_t$ de um estado mental consiste em representar esse estado. (2;3)

Portanto,

(C2) Um estado consciente$_i$ é um estado que é representado por algum estado mental do sujeito.

A premissa (1) é uma estipulação do que é ser um estado mental consciente$_i$ baseada no princípio da transitividade. O que a definição sugere é que estados mentais intransitivamente conscientes são estados mentais dos quais os sujeitos são conscientes de estar/ter esses estados. E, se um sujeito é consciente de que tem/está nesses estados, então ele exibe consciência$_t$. Pois, assim como ser consciente de um objeto ou evento no mundo exige um estado mental através do qual se é consciente desse objeto ou evento, ser consciente de um de nossos próprios estados mentais exige ter um estado mental através do qual se é consciente desse estado.

Na medida em que as diferentes versões do representacionalismo de segunda-ordem consistem em implementações do princípio da transitividade, expresso na premissa (1) no argumento acima, essa premissa é o principal alvo das objeções enfrentadas pelo argumento mestre. Consideramos na sequência duas dessas objeções e como defensores do representacionalismo de segunda-ordem podem responde-las.

Uma das críticas que a premissa (1) enfrenta é a de que a definição não é apropriada porque o conceito de consciência transitiva de estado é menos fundamental do que o conceito de consciência intransitiva. Ou seja, o conceito de consciência transitiva é conceitualmente dependente do conceito de consciência intransitiva. Sendo assim, não seria plausível utilizar o conceito de consciência transitiva de estado como explicador do conceito de consciência intransitiva de estado.

Um argumento nessa direção é articulado por Kriegel (2009), que sustenta através de um exame do uso ordinário do termo consciência que o conceito de consciência intransitiva é mais fundamental do que o conceito de consciência transitiva (de estado). O argumento pode ser sintetizado da seguinte maneira: A consciência transitiva (de estado) é caracterizada como a propriedade que um estado mental M tem quando um sujeito S é consciente de algo X através do estado M. Porém, a caracterização acima é ambígua, permitindo duas leituras sobre o estado mental M: ou o estado mental M é consciente ou o estado mental M é inconsciente. Se o estado mental M é consciente, então S é consciente de que tem o estado mental M e de algo X (através

do estado mental M). Se o estado mental M é inconsciente, então S é consciente apenas de X, mas não de que tem o estado mental M. Porém, o uso ordinário do termo consciente não reconhece expressões como “João é inconscientemente consciente da chuva” como uma expressão coerente. Isto é, não há um uso do termo consciente no senso comum que corresponda a ideia de que S é consciente de algo X e o estado mental M, através do qual S é consciente de X é um estado mental inconsciente. Portanto, se desejamos que nosso conceito seja consistente com o uso ordinário da noção de consciência, devemos conceber que os estados mentais aos quais atribuímos a propriedade de ser transitivamente consciente são sempre estados mentais conscientes (em sentido intransitivo). Adotar essa concepção implica em adotar a perspectiva de que consciência transitiva de estado não pode ocorrer na ausência de consciência intransitiva de estado (Ver, KRIEGEL, 2009, p. 30-31). Consequentemente, Kriegel assume a perspectiva de que a consciência transitiva não pode ser o explicador para a consciência intransitiva, visto que a primeira não pode ocorrer na ausência da segunda.

Uma objeção que pode ser levantada contra o argumento de Kriegel diz respeito à relevância do uso ordinário do conceito de consciência para a elaboração de uma explicação cientificamente informada do fenômeno da consciência intransitiva e da sua relação com a consciência transitiva. Podemos argumentar contra Kriegel que, ainda que seja correto afirmar que o uso ordinário da noção de consciência não admite expressões nas quais se faz menção a estados mentais inconscientes através dos quais se é consciente de algo, as limitações exibidas pelo uso ordinário da noção de consciência são irrelevantes como parâmetro para a elaboração de uma explicação substantiva e cientificamente informada da consciência intransitiva de estado. Pois, se a intenção por trás do argumento articulado pelas teorias representacionalistas de segunda-ordem é promover uma explicação para o fenômeno da consciência (intransitiva) de estado, então essa explicação não necessariamente precisa ser compatível com o uso ordinário desses conceitos.

Para usar uma ilustração clássica da história da ciência, podemos comparar o que o argumento mestre promove com relação ao conceito de consciência intransitiva de estado com o que a física do século 19 promoveu com relação ao conceito de calor. Do ponto de vista do senso comum, não há uma relação conceitual entre “calor” e “energia cinética média das moléculas”. Contudo, durante o século 19, o desenvolvimento da termodinâmica permitiu estabelecer que os mecanismos microfísicos responsáveis pelo fenômeno observável que chamamos de calor consistem no movimento das moléculas. Mais precisamente, o desenvolvimento da termodinâmica permitiu estabelecer que o calor nada mais é do que a energia cinética média das moléculas. A partir da descoberta de que o calor consiste no

movimento das moléculas, o conceito de calor é ressignificado. A partir de tal descoberta, o conceito cientificamente informado de “calor” passa a ter uma relação conceitual estreita com o que é expresso pela noção de “energia cinética das moléculas”.

No caso do calor, a motivação para a revisão do conceito diz respeito a descobertas empíricas, mas é plausível pensarmos que há outras motivações que também poderiam promover uma revisão de um conceito. Em particular, é plausível pensar que há boas motivações para a revisão de um conceito se tal revisão promover algum tipo de melhoria epistêmica. Por exemplo, se revisar um conceito é algo que promove uma maneira mais frutífera de explicar o fenômeno denotado pelo conceito em questão, então há boas motivações para efetuar tal revisão (CAPPELEN; PLUNKET, 2020 p.10).

Este parece ser o caso do conceito de “consciência intransitiva”: o que o argumento mestre promove com relação ao conceito de consciência intransitiva de estado é uma tentativa de revisão conceitual motivada pela promoção de uma maneira frutífera de explicar a propriedade que um estado mental tem quando ele é consciente. Ao sustentar que o conceito de consciência transitiva seja compreendido de tal modo que é conceitualmente possível que um sujeito seja consciente de algo X, mas o estado mental através do qual o sujeito é consciente de X é inconsciente, torna-se disponível uma estratégia explicativa potencialmente frutífera. A saber, torna-se possível explicar porque há estados mentais que são conscientes e outros que são inconscientes em termos de relações representacionais entre estados mentais. Ou seja, o que é sugerido é que, se compreendermos a consciência intransitiva como associada à ocorrência de estados mentais (inconscientes) que os representam, então é possível elaborarmos uma explicação da consciência intransitiva em termos de relações representacionais entre estados mentais de um sujeito.

Antecipando esse tipo de contra-argumento, Kriegel (2009, p.30 n13) sustenta que a objeção sobre a irrelevância do uso ordinário do conceito de consciência falha em reconhecer que a caracterização das variedades de consciência é substancialmente fundamentada no uso ordinário do termo consciência. Portanto, se as propriedades características de cada variedade de consciência são definidas em termos do uso ordinário do conceito de consciência, asserções sobre suas interrelações podem e, talvez devam, repousar sobre um exame de aspectos da gramática desse conceito no vocabulário do senso comum.

Essa resposta, no entanto, não é satisfatória. Podemos admitir que os usos ordinários dos conceitos de consciência são um aspecto relevante para o exame de como um conceito é, de fato, utilizado. E podemos admitir que o uso ordinário de um conceito estabelece restrições sobre como o concebemos. Contudo, seria um erro sustentar que o uso ordinário desses

conceitos tem a última palavra no que diz respeito a como os conceitos sobre a consciência *devem* ser concebidos a fim de que se possa elaborar explicações substantivas a partir deles. Pois, isso consistiria em assumir que o senso comum possui a melhor compreensão possível do conceito de consciência intransitiva, o que é disputável (ver, Cappelen, 2020, p.134).

A premissa (1) pode ensejar ainda outro tipo de objeções: objeções sobre a sua circularidade. Os comentários de Alvin Goldman (1993, p.366) sobre a estratégia explicativa adotada por teorias representacionalistas de segunda-ordem ilustram esse tipo de objeção. Segundo Goldman, o que autores como Lycan (1997) e Rosenthal (1986; 1993) propõem como uma caracterização da consciência (de estado) é uma abordagem que busca definir a consciência (intransitiva de estado) com o auto-conhecimento, auto-monitoramento, ou reflexão de segunda-ordem. Deste modo, para que um estado mental M seja consciente, o sujeito que possui M deve ter outro estado mental M' que é consciente de M durante a sua ocorrência. Entretanto, sugere Goldman, há algo insatisfatório com esse tipo de caracterização, a saber, o fato de que a tentativa de caracterizar ou definir consciência emprega, na definição/caracterização o próprio termo definido/caracterizado. Deste modo, a estratégia explicativa adotada por teorias representacionalistas de segunda-ordem estariam embasadas sobre uma definição ou caracterização circular do que é a consciência.

Entretanto, a acusação de circularidade não é plausível. Em nossa formulação do argumento mestre, fizemos o uso de subscritos como uma maneira de evidenciar que, na definição estabelecida na premissa (1), há duas noções distintas de consciência sendo empregadas: consciência transitiva de estado (ou Consciência$_t$) e consciência intransitiva de estado (ou Consciência$_i$). Conforme aponta Rosenthal (2005, p. 109) a estratégia explicativa adotada pelas teorias representacionais de segunda-ordem não é circular porque ela emprega dois conceitos distintos de consciência que podem ser compreendidos de maneira independente. A consciência transitiva de estado pode ser utilizada como uma maneira de falarmos sobre como somos conscientes de objetos, eventos e estados do mundo que estão fora de nossas cabeças. Por exemplo, quando nós percebemos ou pensamos sobre algo que não é um estado mental. Com efeito, "somos capazes de compreender o que é ter um pensamento ou percepção independentemente de compreender o que é para um estado mental ser consciente" (ROSENTHAL, 2005, p. 109). Assim, dada a possibilidade de compreendermos o conceito de consciência$_t$ independentemente de utilizarmos ou compreendermos o uso do conceito de consciência$_i$, não há circularidade envolvida na definição estabelecida por (1).

2.4 O EXPLANANS DA ABORDAGEM REPRSENTACIONALISTA DE SEGUNDA-ORDEM

Uma das questões centrais que uma teoria da consciência deve responder é como estados mentais adquirem a propriedade de ser fenomênicamente conscientes. A resposta das teorias pertencentes à abordagem de segunda-ordem compartilha de um elemento comum: todas elas consistem em *implementações* do princípio da transitividade (GOTTLIEB, 2020, p.1). Mas, o que significa implementar o princípio da transitividade? E como isso está relacionado com a natureza de "segunda-ordem" dessas teorias? Nessa seção, esclarecemos quais de que modo o compromisso com o princípio da transitividade é explorado na estratégia explicativa do representacionalismo de segunda-ordem e como essa explicação pode estar vinculada aos processos metacognitivos.

2.4.1 O princípio da transitividade como *explanans* da abordagem de segunda-ordem

O princípio da transitividade sustenta que estados mentais conscientes são estados mentais de que um sujeito está, de alguma maneira, consciente. Para os defensores da abordagem de segunda-ordem, o que é afirmado pelo princípio da transitividade é um dado pré-teórico intuitivamente verdadeiro (GOTTLIEB, 2020, p. 7, n.9)[20], a partir do qual suas explicações podem ser articuladas.

Há dois usos distintos do termo "consciente" no princípio da transitividade, que correspondem às duas noções distintas de consciência distinguidas no capítulo anterior: consciência *transitiva* e *intransitiva*. Por um lado, podemos atribuir consciência *intransitiva* aos estados mentais. Nesse caso, o predicado "consciente" é unário e tem como argumento um estado mental. O outro uso de consciência presente no princípio da transitividade, e que lhe garante o seu nome, é o de *consciência transitiva*. Há duas variedades de consciência transitiva. Quando dizemos que um sujeito é *transitivamente consciente,* utilizamos um predicado binário, cujos termos são um sujeito e um estado, evento ou objeto de que um sujeito é consciente. Já quando dizemos que um estado mental é *transitivamente consciente,* utilizamos um predicado binário, cujos termos são um estado mental e um estado, evento ou objeto de que um sujeito é consciente em virtude de ter esse estado mental.

[20] A afirmação parece mais intuitiva e próxima do senso comum em sua forma contrapositiva: nenhum estado de que um sujeito não está consciente é um estado mental consciente.

Argumentamos no capítulo anterior que o conceito de consciência transitiva aplicado às criaturas é analiticamente derivado do conceito de consciência transitiva aplicado aos estados mentais (ver, KRIEGEL, 2009 e o capítulo 1, seção 3.4). Resumidamente, sustentamos que o que é para um indivíduo ser transitivamente consciente de P é se relacionar apropriadamente com P, de tal modo que ele seja consciente P. E, uma maneira através da qual um sujeito pode ser consciente de P consiste em ter uma percepção ou pensamento sobre P. Se admitirmos a conclusão desse argumento, então é possível interpretar o princípio da transitividade como sustentando que: um estado mental M é consciente, em sentido intransitivo, apenas se o sujeito é transitivamente consciente de M (via um estado mental M* que o representa). Esta é uma interpretação que oferece uma explicação embrionária do mecanismo em virtude do qual um estado mental adquire a propriedade de ser fenomênicamente consciente. A tarefa de uma teoria de segunda-ordem é expandir e esclarecer essa explicação embrionária.

As explicações sobre a consciência fenomênica oferecidas por teorias de segunda-ordem são implementações do princípio da transitividade. O que isso significa é que as explicações oferecidas por teorias de segunda-ordem consistem em hipóteses sobre quais são (e como operam) os mecanismos psicológicos em virtude dos quais um estado mental é intransitivamente consciente[21]. Vimos acima que a hipótese embrionária oferecida pelo princípio da transitividade consiste em afirmar que a consciência intransitiva de um estado mental é um produto de uma relação representacional entre os estados mentais de um sujeito. Ao oferecer uma implementação do princípio da transitividade, as teorias pertencentes à abordagem de segunda-ordem devem expandir e esclarecer essa proposta. Uma das maneiras de realizar essa tarefa consiste em identificar, por exemplo, quais são os tipos adequados de estados mentais representacionais que compõem o suposto mecanismo (e.g., pensamentos ou percepções).

2.4.2 Pensamentos e percepções de segunda-ordem

Há, fundamentalmente, duas alternativas que podem ser adotadas como resposta à questão sobre qual é o tipo de estado mental de segunda-ordem que é capaz de tornar um estado de primeira-ordem consciente: pensamentos de segunda-ordem ou estado de natureza

[21] Frequentemente, no entanto, as hipóteses são apenas “esboços” de mecanismos, no sentido utilizado por Piccinini e Craver (2011). Por essa razão, é legítimo dizermos também que as explicações oferecidas por teorias da abordagem de segunda-ordem são análises funcionais da consciência.

perceptual de segunda-ordem. Esta divisão dá origem as duas principais versões de teorias representacionalistas de segunda-ordem.

Teorias que sustentam que estados de natureza perceptual são o tipo relevante de estado mental através do qual um sujeito torna-se consciente dos seus estados mentais de primeira-ordem são chamadas de teorias de percepção de segunda-ordem, ou também de *teorias do sentido interno* (LYCAN, 1997). De acordo com essa versão das teorias representacionais de segunda-ordem, seres humanos possuem dois tipos de mecanismos perceptuais. Por um lado, temos mecanismos responsáveis por "monitorar" e construir representações do ambiente (e dos nossos próprios corpos), isto é, os nossos sistemas sensoriais. Por outro lado, possuímos uma espécie de "sentido interno", isto é, mecanismos cuja função é "monitorar" e construir representações sobre os estados perceptuais de primeira-ordem ocorrentes em sua mente (CARRUTHERS, 2000, p. 210). Eis como Lycan sintetiza essa estratégia geral através da postulação de um mecanismo de monitoramento interno:

A teoria do sentido interno diz que a consciência é a operação bem-sucedida de um escâner ou monitor interno que produz representações de segunda-ordem de estados psicológicos de primeira-ordem" (LYCAN, 1997, p. 31, tradução nossa, adaptado).

Assim caracterizada, as teorias de percepção segunda-ordem (ou teorias do sentido interno, no jargão de Lycan) estão comprometidas com, pelo menos, uma suposição empírica. A suposição de que realmente há um escâner ou mecanismo no cérebro de organismos dotados de consciência (de estado) que é responsável por monitorar os estados internos desses organismos e cuja atividade é responsável por fazer com que estados mentais sejam conscientes (cf. LYCAN, 1997, p. 16)

Já as teorias que sustentam que são pensamentos o tipo relevante de estados representacionais de segunda-ordem através dos quais os sujeitos são conscientes de seus estados mentais de primeira-ordem são chamadas de Teorias do Pensamento de Segunda-ordem. Pensamentos são estados tipicamente concebidos como constituídos de conceitos. Portanto, essa versão do representacionalismo de segunda-ordem assume que estados mentais conscientes envolvem sempre o uso de conceitos, ainda que de uma variedade especial a ser especificada pela teoria. Um pensamento de segunda-ordem é um pensamento sobre um estado mental de primeira-ordem (uma crença, um desejo, percepção e assim por diante). Assim, segundo essa versão do representacionalismo de segunda-ordem, o que faz com que minha percepção do ruído da geladeira seja consciente é o fato de que, acompanhando a ocorrência dessa percepção, eu possuo um pensamento assertórico que representa que eu estou tendo uma

percepção do ruído da geladeira. Rosenthal (2005) e Carruthers (2000) são os principais defensores dessa perspectiva.

2.4.3 Teorias de segunda-ordem e metacognição:

Identificar aproximações entre a explicação da abordagem de segunda-ordem sobre a consciência e os mecanismos psicológicos abordados em outras áreas de investigação é uma maneira indireta de oferecer plausibilidade empírica e teórica para a estratégia explicativa da abordagem de segunda-ordem. Tendo esse objetivo em vista, defensores de teorias de segunda-ordem da consciência frequentemente aproximam a sua hipótese sobre os mecanismos psicológicos responsáveis por fazer com que estados mentais seja fenomênicamente conscientes com as hipóteses de mecanismos psicológicos em outras áreas da psicologia. Em especial, há aproximações entre a sugestão de que estados mentais de segunda-ordem são parte fundamental no mecanismo responsável por fazer alguns estados mentais de primeira-ordem serem conscientes e os processos metacognitivos.

Há razões conceituais que justificam a aproximação entre implementações do princípio da transitividade e a metacognição. O princípio da transitividade sustenta que estados mentais conscientes são estados de que um sujeito está, de alguma maneira apropriada, consciente. Vimos que uma hipótese sobre como o princípio da transitividade é implementado consiste em afirmar que estados mentais se tornam fenomênicamente conscientes na medida em que eles são alvo de representações de segunda-ordem. Nesse sentido, o processo postulado como explicador da consciência fenomênica consiste em uma atividade psicológica na qual estados mentais do sujeito representam os estados mentais perceptuais e sensoriais ocorrentes. Podemos rotular esse processo de metarepresentacional, porque os *outputs* desse processo são representações mentais que são sobre outras representações mentais (ver, SCOTT, 2001).

Os processos metacognitivos são, similarmente, caracterizados por serem processos psicológicos que são sobre atividades cognitivas de primeira-ordem e tem metarepresentações como *output*[22]. Exemplos típicos de processos metacognitivos são atividades por meio das quais avaliamos a qualidade ou confiabilidade de nossos próprios processos cognitivos em diversas modalidades, como na memória e na percepção (ROSENTHAL, 2000, p. 203; LAU, 2019; LAU; ROSENTHAL, 2011; MORALES et al., 2018).

[22] Mas, ver PROUST, 2007, que nega que metacognições exijam metarepresentações.

Tendo em vista a natureza de segunda-ordem da metacognição, ou seja, o fato de que são processos mentais sobre outros processos mentais, Rosenthal (2000) aponta para a seguinte aproximação entre a consciência e a metacognição: estados mentais conscientes são estados de que o sujeito está, de alguma maneira apropriada, consciente de ter. Por seu turno, processos metacognitivos são processos através dos quais temos acesso aos nossos estados mentais ocorrentes. Sendo assim, ambos fenômenos envolvem um tipo de acesso de segunda-ordem aos estados mentais que temos ou parecemos estar tendo (ROSENTHAL, 2000, p. 204). Portanto, é plausível buscarmos por aproximações entre os mecanismos que supomos ser responsável pela consciência fenomênica e a metacognição.

Uma segunda razão para associarmos o *explanans* das teorias de segunda-ordem e a metacognição é empiricamente fundamentada (BROWN, et al., 2019). Há evidências que sugerem que a atividade neural no córtex pré-frontal é importante para a percepção consciente (LAU; PASSINGHAM, 2007). Quando a atividade nessa região é manipulada via estimulação magnética transcraniana, há efeitos sob alguns processos metacognitivos explícitos (ver abaixo a distinção explícito/implícito) e sob a percepção consciente (ROUNIS et al., 2010). Por essa razão, é possível sustentar que os processos metacognitivos e os processos responsáveis pela consciência fenomênica compartilham de mecanismos e componentes neurais similares (BROWN et al, 2019, p. 757).

Contudo, nem todos processos metacognitivos parecem ser candidatos apropriados para a formulação de hipóteses sobre a consciência. Processos metacognitivos podem ser divididos em duas categorias: explícitos e implícitos[23]. Processos metacognitivos explícitos são processos cognitivos sobre outros processos cognitivos que um sujeito se engaja conscientemente (BROWN et al, 2019, p. 757). Por exemplo, mediante um pedido de avaliação da sua confiança na identificação de um estímulo visual, um sujeito pode se engajar, conscientemente, em um processo metacognitivo de avaliação da confiabilidade da sua própria percepção e capacidade de identificação de objetos. Processos metacognitivos explícitos não são o tipo adequado de processo para a formulação de hipóteses sobre como estados mentais adquirem consciência fenomênica, pois ele pressupõe o fenômeno que exige explicação. Não obstante, testes envolvendo a metacognição explícita podem desempenhar um papel relevante na avaliação empírica das abordagens de segunda-ordem, conforme sugerido no parágrafo anterior.

[23] Ver também a distinção entre os processos metacognitivos de controle por veículos e por conteúdo, a qual sugere que o veículo de representações mentais pode ter efeitos sob o controle de comportamentos independentemente do reconhecimento explícito do conteúdo dessas representações (DIENES; PERNER, 2002).

Processos metacognitivos implícitos, por outro lado, são processos cognitivos sobre outros processos cognitivos que ocorrem automaticamente (involuntariamente) e de maneira inconsciente (FLEMING, 2021, p.53). Processos deste tipo são, possivelmente, amplamente presentes em nossos sistemas cognitivos (DIENES; PERNER, 2002), desempenhando um papel relevante, por exemplo, no auto monitoramento e regulação do aprendizado desde a infância (BALCOMB; GERKEN, 2006; 2008) e na seleção de estratégias ou procedimentos para realizar uma tarefa (CARY; REDER, 2002).

Ao contrário dos processos metacognitivos explícitos, os processos implícitos parecem ser melhores candidatos à formulação de hipóteses sobre a consciência fenomênica. Se uma teoria articula uma hipótese que postula que processos metacognitivos implícitos são relevantes para que um estado mental adquira a propriedade de ser fenomênicamente consciente, tal explicação não incorre em uma petição de princípio. A realização de processos metacognitivos implícitos não pressupõe que um sujeito deve ter consciência da realização desses processos, nem que ele deve realizá-los voluntariamente. Nesse sentido, as explicações sobre a consciência fenomênica que postulam que algum tipo de processo metacognitivo implícito é relevante para que um estado mental adquira a propriedade da consciência fenomênica podem sustentar que a aquisição dessa propriedade é um produto resultante da interação entre estados e/ou processos mentais inconscientes de um sujeito.

O monitoramento da qualidade ou confiabilidade dos sinais perceptuais disponíveis no sistema cognitivo de um sujeito (ver, LAU, 2019) é um exemplo pontual de um processo metacognitivo de monitoramento implícito que pode ser relevante para a articulação de uma hipótese de segunda-ordem sobre a consciência fenomênica. Na versão articulada por Lau (2019) de uma teoria representacionalista de segunda-ordem, estados mentais fenomênicamente conscientes são o resultado da ocorrência de um estado perceptual de primeira-ordem M (i.e., um estado que representa um objeto, propriedade ou evento do mundo ou do corpo do sujeito) acompanhado de uma representação de segunda-ordem M*, que representa que há um estado perceptual M cujo conteúdo é uma representação confiável do estado atual mundo lá fora (ou do corpo do sujeito). De tal maneira, a ocorrência da representação de segunda-ordem M* faz com que o conteúdo do estado mental M seja consciente porque permitem “a realização de inferências relevantes serem realizadas (automaticamente) e se tornarem disponíveis para o agente (em um nível pessoal, de maneira que a fazer as inferências parecerem justificadas subjetivamente. ” (LAU, 2019, tradução nossa).

Apesar de haverem aproximações entre processos metacognitivos e a consciência fenomênica realizados pelas teorias de segunda-ordem, a maneira como as teorias pertencentes à abordagem

de segunda-ordem enxergam a relação entre os processos metacognitivos e a consciência fenomênica exige qualificações. Processos metacognitivos e a consciência fenomênica são fenômenos distintos (ROSENTHAL, 2000, p.204). Nem todos os processos metacognitivos são relevantes para a consciência fenomênica, por exemplo, o caso dos processos explícitos. Além disso, há características significativamente diferentes no tipo de acesso de segunda-ordem aos estados mentais exibidos pela consciência de estados mentais e por alguns processos metacognitivos.

Segundo Rosenthal (2000, p.204-205), há uma diferença no modo como metacognições podem acessar os estados e processos mentais de primeira-ordem e a forma do acesso dos estados mentais de segunda-ordem em virtude do qual somos (transitivamente) conscientes de estados mentais de primeira-ordem. Quando se trata de metacognições, são possíveis casos como juízos de metamemória nos quais julgamos saber ou ter retido uma determinada informação, mas somos incapazes de relatar exatamente qual é essa informação. Por vezes, esta informação pode parecer estar "na ponta da língua", isto é, sentimos estar muito próximos de conseguir relatar que informação é essa[24]. Em outros casos podemos sentir que sabemos essa informação, mas não julgamos ser capazes de relatar exatamente qual é o conteúdo dessa informação[25]. Segundo Rosenthal (2000, p. 204), casos dessa natureza envolvem estarmos conscientes de que sabemos algo, sem sermos conscientes do que é que sabemos (a informação que não conseguimos lembrar ou relatar). Nesse sentido, estamos conscientes de tais estados apenas na medida em que eles são estados que retêm uma informação relevante para uma determinada questão, mas não com relação à informação que eles realmente carregam. Sob essas condições, porém, não estamos *realmente* conscientes dos estados em questão (como estaríamos se soubéssemos e tivéssemos acesso adequado para relatar a informação contida por esses estados mentais). Em outras palavras, para que um estado mental conte como um caso de estado mental (intransitivamente) consciente, é o segundo tipo de acesso de segunda-ordem sobre o estado mental que devemos possuir.

Para Rosenthal (2000, p. 207), notar essa diferença revela implicações cruciais sobre a sua versão do representacionalismo de segunda-ordem, a teoria dos pensamentos de segunda-ordem. Se apenas os estados mentais de que temos consciência do conteúdo que eles carregam contam como estados (intransitivamente) conscientes (ver, ROSENTHAL, 2000, p.205), então a maneira através da qual somos conscientes dos estados mentais (n.b. que atribuímos consciência intransitiva) deve ser sensível para a informação que eles carregam

[24] Em inglês, esse fenômeno é conhecido como *tip-of-the-tongue states* (TOTs).
[25] Em inglês, estes fenômenos são conhecidos como *feeling of knowing* (FOF).

(ROSENTHAL, 2000 p. 206). Nós podemos representar algo, fundamentalmente de duas maneiras, via percepções ou via pensamentos. Representações perceptuais dos estados mentais de primeira-ordem parecem ser o tipo inadequado de representação para capturar a diferença entre estar consciente de um estado mental com relação ao conteúdo que ele carrega ou com relação às questões que ele potencialmente pode responder (como nos juízos de metamemória). Sendo assim, ter pensamentos que representam conceitualmente o estado mental parecem ser a única outra alternativa viável. Tais pensamentos representam de maneira conceitual os estados mentais de primeira-ordem. Mais precisamente, os pensamentos de segunda-ordem representam, conceitualmente, que o indivíduo tem um estado de determinado tipo, que é caracterizado em termos de uma atitude tida frente ao conteúdo sensorial ou conceitual particular (ROSENTHAL, 2000, p. 206).

2.5 MOTIVAÇÕES EMPÍRICAS PARA A ADOÇÃO DO REPRESENTACIONALISMO DE SEGUNDA-ORDEM

Vimos que o representacionalismo de segunda-ordem sustenta que representações de segunda-ordem são uma condição, no mínimo, necessária para que estados mentais sejam conscientes. Esta é uma tese que, se interpretada à luz de teses empíricas adicionais, implica em predições testáveis a respeito dos mecanismos responsáveis pela consciência fenomênica. Para uma teoria da consciência, a capacidade de gerar predições testáveis é uma importante virtude teórica – a virtude diacrónica da fecundidade – que pode pesar a seu favor quando se trata de escolher entre teorias rivais (KEAS, 2017). Se, além disso, as predições oferecidas por uma teoria são compatíveis com as evidências empíricas disponíveis, essa teoria desfruta da virtude teórica da acurácia evidencial (KEAS, 2017). Neste capítulo argumentamos que o representacionalismo de segunda-ordem desfruta de ambas as virtudes teóricas, fecundidade e acurácia evidencial, pois (1) oferece previsões testáveis (2) que tem recebido suporte empírico. Por essa razão, esta abordagem é uma forte candidata a uma teoria não só empiricamente plausível, mas verdadeira.

Na presente seção identificamos quais são as predições empíricas associadas à tese central do representacionalismo de segunda-ordem e diferenciamo-las de repercussões empíricas de abordagens rivais similares (subseções 1 e 2). Na sequência, revisamos criticamente algumas das evidências provindas de estudos neurocognitivos que corroboram

com as previsões empíricas oferecidas pelo representacionalismo de segunda-ordem e contrariam as previsões das abordagens rivais (subseções 2 e 3).

2.5.1 Implicações empíricas do representacionalismo de segunda-ordem

Algumas das predições empíricas do representacionalismo de segunda-ordem têm sido alvo de investigações nas neurociências cognitivas e tem encontrado suporte evidencial substancial (LAU; ROSENTHAL, 2011; BROWN et al, 2019; LAU; BROWN, 2019). Esta seção identifica quais são estas predições e como elas se diferenciam de previsões empíricas de abordagens rivais, como as teorias de primeira-ordem em geral e a teoria da área de trabalho global em particular.

A abordagem representacionalista de segunda-ordem sustenta o princípio da transitividade, segundo o qual um estado mental é consciente apenas se o sujeito é, de alguma maneira apropriada, consciente de estar nesse estado mental. Representacionalistas de segunda-ordem em geral sustentam que a maneira através da qual somos conscientes de nossos estados mentais consiste em ter uma representação (de segunda-ordem) de que nós estamos tendo um estado mental de primeira-ordem. Nesse sentido, a estratégia explicativa da abordagem representacionalista de segunda-ordem consiste oferecer hipóteses sobre a implementação do princípio da transitividade em termos de representações de segunda-ordem: para essa abordagem o mecanismo psicológico responsável por implementar a consciência é algum tipo de processo meta-representacional.

Se alguma versão da abordagem representacionalista de segunda-ordem é correta, então nossa experiência consciente não depende unicamente de nossos estados mentais de primeira-ordem. Isto é, não depende unicamente daqueles estados que são sobre objetos ou propriedades no mundo, mas também, ou principalmente, de estados mentais que são sobre nossos próprios estados mentais de primeira-ordem.

Sustentar o representacionalismo de segunda-ordem implica em sustentar que mudanças nas nossas representações (ou processos representacionais) de segunda-ordem são condição suficiente para implicar em mudanças na nossa experiência subjetiva (mesmo que nenhuma mudança em nossos estados mentais de primeira-ordem ocorra) (LAU; ROSENTHAL, 2011, p.385). É a partir desta consequência da abordagem representacionalista de segunda-ordem que é possível entretermos maneiras de investigá-la empiricamente. As investigações empíricas da consciência podem corroborar e garantir credibilidade empírica ao representacionalismo de

segunda-ordem, caso as evidências geradas sejam compatíveis com a ideia de que mudanças na nossa experiência subjetiva dependem, de alguma maneira, dos processos representacionais de segunda-ordem.

O vínculo entre a abordagem representacionalista de segunda-ordem com os dados empíricos coletados na investigação em neurociências cognitivas depende de hipóteses empíricas externas ao representacionalismo de segunda-ordem. Ambas as hipóteses dizem respeito aos componentes anatômicos do cérebro, cuja atividade se assume que está correlacionada com processos ou estados representacionais de primeira ou segunda-ordem. Os componentes anatômicos do cérebro cuja atividade se correlaciona de alguma maneira com a presença de estados mentais conscientes são chamados de *correlatos neurais da consciência* (CHALMERS, 2000). A noção relevante de um correlato neural da consciência empregada neste capítulo diz respeito a correlação entre (padrões) de ativação do cérebro que se correlacionam com experiências particulares (e.g, ver uma nuvem, ouvir um acorde). Nesse sentido, a noção de um correlato neural da consciência empregada neste capítulo é distinta do que poderia ser chamado de um *correlato neural do nível de consciência*, que busca identificar padrões de ativação do cérebro que se correlacionam com estados mais gerais como vigília, sono ou coma ao invés de experiências particulares (CHALMERS, 2000; ver, REES et al, 2002, p.260). Esse recorte sobre como compreender os correlatos neurais da consciência é motivado pela ideia de que a noção relevante para a investigação da consciência e formulação de uma teoria sobre esse fenômeno é a variedade de consciência de estado (em contraste com a noção de consciência de criatura).

A primeira hipótese empírica adicional que devemos assumir para que o representacionalismo de segunda-ordem seja empiricamente investigável diz respeito às atividades neuronais de que se acredita depender as representações perceptuais de primeira-ordem. Comumente assume-se que as representações perceptuais de primeira-ordem dependem da atividade neuronal em regiões sensoriais primárias no cérebro de humanos e outros primatas.

Por exemplo, no caso da visão, um dos sentidos mais estudados e compreendidos em primatas e humanos, a estrutura encefálica primariamente associada aos estímulos visuais é o córtex estriado, também conhecido como córtex visual primário (ou área de Brodmann 17, ou V1) que se localiza no lobo occipital do encéfalo de primatas. O córtex visual primário é assim chamado por que é a primeira estrutura encefálica a receber aferências dos olhos via núcleo geniculado lateral (NGL) (BEAR et al, 2017, p. 212; p.341). Os neurônios do córtex estriado são sensíveis a diversas propriedades visuais, como orientação, direção de movimento, cor e contraste (TONG, 2003, p.220). Uma das principais e inaugurais evidências que vinculam a

atividade no córtex estriado à percepção visual consciente diz respeito a observação de lesões cerebrais pontuais. Por exemplo, a partir da observação de que lesões pontuais localizadas em determinadas partes do córtex estriado eram responsáveis por escotomas ou casos de cegueira cortical, foi possível estabelecer que há correspondências entre as regiões do campo visual e a atividade realizada por regiões do córtex estriado. Mais exatamente, notou-se que o córtex estriado está de tal modo correlacionado com o campo visual, que a superfície bidimensional da retina pode ser mapeada sobre as superfícies bidimensionais do córtex estriado (TONG, 2003, p. 220; BEAR, 2017, p.342). Tais observações garantem uma forte justificação empírica para a hipótese de que a atividade nas áreas sensoriais primárias (ao menos no caso da visão) são diretamente responsáveis pela experiência consciente de estímulos sensoriais.

A segunda hipótese diz respeito às atividades neuronais de que se acredita dependerem as representações de segunda-ordem que promovem a consciência. Supõem-se que as representações de segunda-ordem dependem da atividade neural principalmente nos lobos pré-frontal e parietal em humanos e primatas – em especial algumas regiões do córtex pré-frontal dorsolateral (Brodmann 46 e 9) e córtex frontopolar (Brodmann 10) (LAU; ROSENTHAL, 2011, p. 365-366). Tais regiões corticais são usualmente associadas a processos cognitivos complexos como as funções executivas (e.g, atenção e memória de trabalho) e processos metacognitivos. Por exemplo, comentando uma série de estudos sobre a memória de trabalho (e.g, HASSELMO; STERN, 2006; CURTIS; DESPOSITO, 2003), Block (2007) afirma que a região do córtex pré-frontal dorsolateral, uma das regiões que o representacionalismo de segunda-ordem assume ser relevante para a produção de estados mentais conscientes, não armazena sinais sensoriais, mas é uma das principais envolvidas na manutenção de representações em centros sensoriais, sensório-motores e espaciais na parte de trás da cabeça [...] isto é, a área frontal mantém e está vinculada às representações sensoriais na parte de trás da cabeça que representam, por exemplo, cores, formas e movimento (BLOCK, 2007, p.496, tradução nossa, adaptado).

Concomitante a ideia de que estas regiões corticais estão implicadas na realização das funções executivas, o representacionalismo de segunda-ordem assume que tais regiões desempenham um papel necessário para a realização da consciência de estados mentais[26]. Mas,

[26] Entretanto, cabe salientar novamente que o representacionalismo de segunda-ordem não está necessariamente comprometido com a especificação de tais componentes anatômicos como associados a realização da consciência fenomênica em estados mentais. Estas são hipóteses empíricas independentes da hipótese filosófica proposta por essa abordagem. Não obstante, tais hipóteses empíricas são comumente admitidas por seus defensores (KRIEGEL, 2009; LAU, 2010; ROSENTHAL; LAU, 2011; BROWN et al, 2019) e intérpretes (e.g, BLOCK, 2009). Em conformidade com esta leitura, assumiremos neste capítulo que a abordagem representacionalista de segunda-ordem adota estas hipóteses.

diferentes versões do representacionalismo postulam diferentes processos psicológicos através dos quais essas regiões estão implicadas na produção de estados conscientes – por exemplo, Rosenthal (2005) sustenta que quando um pensamento de segunda-ordem representa um estado de primeira-ordem, este último se torna um estado mental consciente, já Lau (2008) sustenta que a consciência é o resultado de um processo de segunda-ordem que arbitra sobre a confiabilidade estatística de um sinal perceptual.

Uma vez admitidas essas hipóteses que correlacionam estados e processos psicológicos com a atividade neuronal em determinadas estruturas ou regiões anatômicas do cérebro humano e de primatas, torna-se possível diferenciar repercussões previstas por abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica. Se as teorias de segunda-ordem são corretas, então a atividade nos córtices pré-frontal e parietal é, de alguma maneira, relevante para termos nossas experiências. De tal modo, lesões nessas regiões corticais ou alterações provocadas artificialmente, por exemplo, através de instrumentos de estimulação magnética transcraniana (EMT), devem ter algum efeito sobre as nossas experiências. Se, ao contrário, teorias de primeira-ordem são corretas, então a atividade nas regiões pré-frontais e parietais do córtex humano não é determinante para as nossas experiências e as lesões ou as alterações provocadas nessas regiões não implicariam em diferenças nas experiências.

Nos casos típicos, as teorias de primeira-ordem sustentam que a experiência perceptual é determinada pela atividade nas áreas sensoriais primárias, independentemente de representações de segunda-ordem (LAU; ROSENTHAL, 2011, p. 367). Não obstante, algumas teorias de primeira-ordem (especialmente a Teoria do Espaço de Trabalho Global) admitem que processos realizados para além das áreas corticais sensoriais primárias podem desempenhar uma função relevante na disponibilização da informação fenomênicamente consciente a outros módulos cognitivos do indivíduo, ou até mesmo que a própria disponibilização da informação das áreas sensoriais primárias para outros módulos cognitivos do indivíduo realiza o que chamamos de consciência fenomênica. Esses são casos de aparentemente sobreposição entre os compromissos e previsões do representacionalismo de segunda-ordem e algumas teorias de primeira-ordem. Porém, veremos na seção seguinte que essa sobreposição não é completa. Mesmo nesses casos, há importantes descontinuidades entre as previsões empíricas do representacionalismo de segunda-ordem e teorias de primeira-ordem. Na seção seguinte, nos voltamos a discussão de um desses casos.

2.5.2 Diferenças entre o representacionalismo de segunda-ordem e a teoria da área do trabalho global

As abordagens representacionalistas de segunda-ordem não estão sozinhas ao sustentar que nossas experiências perceptuais são, de alguma maneira, determinadas pela atividade nos lobos pré-frontal e parietal do córtex cerebral. A Teoria do Espaço de Trabalho Global[27] (doravante, TETG) também assume que mudanças na atividade relevante nos lobos pré-frontal e parietal são suficientes para mudanças na experiência perceptual. Não obstante, há diferenças significativas entre as predições realizadas por essa teoria e o representacionalismo de segunda-ordem. Na sequência apresentamos em que medida a TETG e o representacionalismo de segunda-ordem se aproximam com relação as suas previsões empíricas e de que maneira elas se diferem em sua compreensão de que a atividade em regiões pré-frontais e parietais do córtex estão associadas a consciência fenomênica.

A Teoria do Espaço de Trabalho Global é caracterizada pela tese de que a consciência (fenomênica) consiste no processo de disponibilização global e estabilização dos sinais sensoriais no sistema cognitivo do indivíduo (BROWN et. al, 2019, p.757). Mais detalhadamente, Baars (2005, p. 46) caracteriza a TETG como uma teoria que sustenta uma arquitetura cognitiva na qual a consciência desempenha uma *função*: estabilizar e disponibilizar informações globalmente para diversos módulos cognitivos.

O cérebro é um sistema físico que implementa diversos processadores de informação especializados, que operam em paralelo e de maneira largamente independente. Devido este fato, é útil possuir alguma maneira de estabelecer canais de comunicação entre os processadores paralelos, de modo que eles possam cooperar entre si na resolução de tarefas complexas. Para a TETG, a função da consciência é servir como um canal de comunicação central através do qual a informação de alguns processadores se torna disponível para outros processadores de informação. Nas palavras de Baars: “A teoria sugere uma capacidade de memória fugaz que permite o acesso entre funções cerebrais que de outra forma são separadas”. (BAARS, 2005, p. 46, tradução nossa). O que faz dessa disponibilidade algo útil para o sistema é o fato de que, uma informação que de outro modo estaria isolada a um único processador (e.g, um módulo de processamento de informações sensoriais) se torna disponível para o uso por parte de outros processadores, possibilitando assim o controle e coordenação entre tais processadores via troca de informações (e.g., controle e planejamento de ações com base nos estímulos ambientais).

---

[27] Originalmente, em inglês, *Global Workspace Theory* ou GWT.

Assim como o representacionalismo de segunda-ordem, a teoria do espaço de trabalho global assume que a atividade nas regiões sensoriais primárias não é suficiente para gerar estados conscientes em sentido fenomênico, além disso, são necessários outros processos cognitivos. A TETG (assim como o representacionalismo de segunda-ordem) assume que a atividade neural em estruturas das regiões corticais do lobo pré-frontal e parietal são relevantes para a consciência fenomênica (DEHAENE; CHANGEUX, 2011, p. 213). De tal modo, ambas as abordagens assumem que nossas experiências perceptuais são, de alguma maneira, determinadas pela atividade nos lobos pré-frontal e parietal do córtex cerebral.

Entretanto, a natureza psicológica dessa atividade neural é diferentemente concebida por cada teoria. Em contraste com o representacionalismo de segunda-ordem, a teoria do espaço de trabalho global não assume que a disponibilização e estabilização dos sinais sensoriais envolvem representações de segunda-ordem ou outros processos metacognitivos. Tal fato localiza esta abordagem entre as teorias de primeira-ordem, já que a disponibilização e estabilização dos sinais sensoriais são, fundamentalmente, propagações do mesmo sinal antes do que novas representações geradas pelo sistema cognitivo do organismo (BROWN et al, 2019, p. 757).

Além disso, os processos cognitivos, no caso da teoria do espaço de trabalho global consistem na disponibilização e estabilização da informação proveniente das áreas sensoriais primárias para outras regiões do córtex, que são responsáveis pelo exercício de coordenação e controle do comportamento (LAU; ROSENTHAL, 2011, p. 366). Deste modo, há um segundo aspecto sobre o qual divergem as previsões empíricas da teoria do espaço de trabalho global e as previsões do representacionalismo de segunda-ordem. A TETG, diferentemente do representacionalismo de segunda-ordem, prevê que a atividade neural associada às experiências está associada ao desempenho de funções cognitivas e comportamentais. É relevante expandir brevemente esse ponto, observando como cada abordagem teórica compreende a função associada à consciência fenomênica.

2.6.2.1 A função da consciência no Representacionalismo de Segunda-ordem

O representacionalismo de segunda-ordem é neutro com relação ao impacto que a atividade neuronal responsável por produzir a consciência fenomênica tem sobre o comportamento e o desempenho em tarefas comportamentais. Por exemplo, Rosenthal (2008, p. 831) alega que a consciência não desempenha nenhuma função significativa em processos

psicológicos.[28] Em outras palavras, as representações de segunda-ordem de que depende a consciência fenomênica não impactam no desempenho de tarefas ou nas funções comportamentais. A razão pela qual a abordagem representacionalista de segunda-ordem se posiciona dessa maneira frente a questão sobre a função da consciência está vinculada a ideia de que todos os tipos de estados mentais (ou a maior parte deles) admitem variedades conscientes e inconscientes (CARRUTHERS, 2000; ROSENTHAL, 2008). Na ausência de representações de segunda-ordem os estados mentais de primeira-ordem são estados mentais inconscientes, mas nem por isso eles são incapazes de desempenhar funções psicológicas, como controle cognitivo e de ações.

Há evidências de que somos capazes de realizar ações intencionais, deliberações racionais complexas e resolução de problemas de maneira inconsciente (DIJKSTERHUIS, 2004). Por exemplo, os experimentos de Libet (1985) e Haggard (1999) corroboram com a hipótese de que, quando um sujeito decide conscientemente realizar uma ação, o evento neural que inicia a ação ocorre significativamente antes da volição consciente (ver, ROSENTHAL, 2008, p.833). Evidências como essa justificam a proposta do representacionalismo de segunda-ordem de que a consciência de (pelo menos alguns tipos de) estados mentais não é um aspecto necessário para que eles sejam capazes de desempenhar diversas funções psicológicas. Assim, para essa abordagem, as funções psicológicas, como o desempenho em tarefas perceptuais, comportamentais e cognitivas são dependentes, principalmente, das propriedades pertencentes aos estados mentais de primeira-ordem independentemente da sua relação com as representações de segunda-ordem que os fazem ser estados mentais conscientes (LAU; ROSENTHAL; 2011, p.366). Na medida em que o representacionalismo de segunda-ordem sustenta que desempenho em tarefas perceptuais e cognitivas é uma capacidade atribuível aos estados mentais independentemente de eles serem conscientes ou inconscientes, ele está comprometido coma previsão empírica de que o desempenho em tais tarefas não é uma função específica da consciência e, portanto, não deveria ser (significativamente) prejudicado na ausência de consciência desses estados mentais (ROSENTHAL, 2008)[29].

A Teoria da área de trabalho global, por outro lado, compreende de maneira distinta a relação entre consciência e o desempenho de tarefas cognitivas, perceptuais e comportamentais.

---

[28] A discussão em Rosenthal (2008) é principalmente direcionada a consciência de estado intencionais como pensamentos, desejos e intenções e apenas tangencialmente sobre percepções.

[29] Cabe notar que, pode ser muito difícil a realização natural ou artificial de uma situação na qual apenas os processos responsáveis pela consciência dos estados mentais são alterados. Por esta razão, não seria surpreendente que a maioria dos casos empíricos envolvendo anomalias na consciência são acompanhadas também algum tipo de déficit no desempenho de tarefas, como é o caso da *visão-cega*.

Este aspecto da teoria da área de trabalho global se torna evidente, por exemplo, na maneira como Sun e Franklin (2008, P.162-163) descrevem a TATG defendida por Bernard Baars:

> Em sua teoria da área de trabalho global, Baars (1988) postula que a cognição humana é implementada por uma multitude de processadores de propósitos especiais, relativamente pequenos, que quase sempre são inconscientes. A comunicação entre eles é rara e se dá entre canais de comunicação de banda estreita. Coalizões desses processos são agrupadas na área de trabalho global (e, deste modo, são conscientes). Esta área de trabalho de capacidade limitada transmite as mensagens da coalizão de processadores para todos os processadores inconscientes com o propósito de recrutar outros processadores para se juntar a resolução do evento inédito ou do problema ocorrente. Assim, *a consciência, nesta teoria, nos permite lidar com situações inéditas ou problemáticas que não poderiam ser lidadas eficientemente ou de qualquer maneira por processos habituais inconscientes* (SUN; FRANKLIN, 2008, p. 162-163, tradução e ênfase nossas, adaptado).

O que essa descrição da teoria sugere é que ela supõe haver uma conexão fundamental entre a consciência de estados mentais e a capacidade do sistema cognitivo de mobilizar a informação fornecida pelos estados mentais na realização de diversas tarefas cognitivas complexas. Dessa maneira, uma previsão empírica sustentada pela teoria é de que há uma vinculação, indissociável, entre três elementos: a atividade neural necessária para a realização da consciência; o desempenho em tarefas perceptuais e cognitivas complexas e; a presença de estados mentais conscientes (LAU; ROSENTHAL, 2011, p. 366).

2.5.3 Correlatos neurais da consciência e a relação entre (relatos de) experiência subjetiva e o desempenho cognitivo em estudos neurocognitivos sobre a consciência

Diversos estudos de neuroimagem e de estimulação local da atividade neural identificaram uma associação entre a atividade no lobo pré-frontal e parietal e experiências de estímulos sensoriais e intenções (LAU; ROSENTHAL, 2011, p. 367; DEHAENE; CHANGEUX, 2011). As evidências de que a atividade em áreas do lobo pré-frontal e parietal está correlacionada com a experiência de estímulos sensoriais e intenções (bem como com a sua reportabilidade) garantem suporte às previsões empíricas do representacionalismo de segunda-ordem e da TETG, e contraria as previsões empíricas de outras teorias de primeira-ordem da consciência.

Em um estudo utilizando palavras como estímulos visuais e técnicas de mascaramento visual, Dehaene et al. (2001) (apud DEHAENE; CHANGEUX, 2011, p. 203) mostraram aos participantes do experimento (1) palavras visíveis ou (2) palavras seguidas de mascaramento visual, de modo que não sejam percebidas/relatadas. Em ambos os casos, se observou através da análise de imagens de fMRI, que havia atividade nas áreas responsáveis pelo processamento

visual de palavras em regiões occipitais e temporais do córtex. Porém, apenas na condição das palavras visíveis (i.e., palavras que podiam ser experienciadas e relatadas) se observou atividade distribuída adicional de regiões no córtex cingulado e em áreas do córtex parietal e pré-frontal. Além disso, observou-se um acréscimo na atividade das regiões occipitais e temporais, responsáveis pelo que os autores sugerem ser "processadores inconscientes" de palavras.

Em um estudo sobre a percepção da fala utilizando técnicas de neuroimagem (fMRI), Hasson et al. (2007) observaram, que o conteúdo do que é perceptualmente escutado não é dependente apenas da atividade em regiões sensoriais primárias do córtex auditivo, mas também da atividade em regiões frontais e parietais responsáveis por processamentos de informações multimodais (i.e., processamento de informações de diferentes modalidades sensoriais)[30] (DEHAENE; CHANGEUX, 2011, p.207). Utilizando-se do efeito McGurk, que consiste em uma tripla dissociação entre a silaba percebida (ouve-se "ta"), a sílaba que é pronunciada no áudio de origem ("pa") e a que é vista no vídeo de origem ("ka"), Hasson et al. (2007) puderam observar que o córtice auditivo primário é responsável pelo processamento do estímulo auditivo objetivo (i.e, as propriedades sensoriais), enquanto o processamento da experiência subjetiva (por exemplo, a sílaba "ta" resultante da ilusão perceptual) está correlacionado com a atividade em estruturas localizadas nas regiões pré-frontais e parietais do córtex.

Já no âmbito das ações intencionais, Desmurget et al. (2009) observaram que a experiência de que se pretende realizar uma ação se origina da atividade envolvendo os córtices pré-frontal e parietal, em especial o giro angular (DEHAENE; CHANGEUX, 2011, p.207). Utilizando técnicas de estimulação elétrica direta (durante cirurgias de remoção de tumores cerebrais em humanos sob anestesia local), Desmurget et al. observaram que a estimulação do córtex motor primário levava a produção de movimentos de que os sujeitos não tinham consciência de ter realizado. Por seu turno a estimulação de algumas regiões do giro angular e do giro supramarginal (áreas de Brodmann 39 e 40) resultam em relatos de sensações de intenção de movimento (e.g, "sinto desejo de lamber meus lábios") ou mesmo a ilusão de movimento (e.g. "eu movi minha boca, o que eu disse?", apesar de nenhum movimento ter sido realizado. Ver, DESMURGET et al, 2009, p. 812). Segundo concluem os autores do estudo:

Nosso estudo sugere que a intenção motora e sua consciência são consequências emergentes do aumento da atividade parietal anterior a execução de movimentos. A sensação

[30] Especificamente, o giro frontal inferior e a parte anterior do lóbulo parietal inferior (HASSON, 2007, p.1117-1118).

subjetiva (e potencialmente ilusória) que nós estamos executando um movimento não surge do próprio movimento, mas é gerado pela intenção consciente anterior e suas consequências previstas (DESMURGET, 2009, p. 813, tradução nossa)

As evidências produzidas pelos estudos sintetizados acima são compatíveis com as previsões empíricas do representacionalismo de segunda-ordem e da teoria da área de trabalho global e, portanto, oferecem suporte empírico às duas abordagens. Mas, apesar do conjunto de evidências acima oferecer suporte empírico tanto para o representacionalismo de segunda-ordem como para a TATG, há outros conjuntos de evidências que suportam mais plenamente o representacionalismo de segunda-ordem do que a TATG em virtude dos compromissos de cada teoria com a função desempenhada pela consciência.

Conforme apresentamos anteriormente, a TATG sugere uma conexão inextrincável entre a consciência e o desempenho em tarefas cognitivas complexas. Portanto, se houverem evidências de que a manipulação da atividade nos supostos correlatos neurais da consciência (nas regiões pré-frontal e parietal) provocam mudanças na experiência, mas não provocam mudanças no desempenho de tarefas cognitivas, então essas evidências corroboram às previsões do representacionalismo de segunda-ordem e contrariam as previsões da TATG sobre a função desempenha pela consciência. Em outras palavras, evidências que revelem uma dissociação entre desempenho cognitivo e experiência no que diz respeito a atividade neuronal dos correlatos da consciência pesam contra a TATG e a favor da ROS.

Um estudo realizado por Lau e Passingham (2006) utilizou de procedimentos de mascaramento visual para realizar duas condições experimentais em que os sujeitos apresentam o mesmo nível de desempenho na discriminação de estímulos visuais, mas relatam diferenças na experiência subjetiva. A tarefa realizada pelos participantes consistia na discriminação forçada de um estimulo visual, se um diamante ou um quadrado em duas condições distintas, a apresentação do estímulo visual seguido de um mascaramento visual após 30ms ou 100ms. Na sequência, uma avaliação da experiência subjetiva era realizada, na qual os sujeitos deveriam indicar se haviam discriminado o estimulo visual ou apenas adivinhado forçadamente (LAU; PASSINGHAM, 2006, p.18763). Além disso, o estudo contou com medições da atividade cerebral através de fMRI, cujo propósito fora localizar possíveis correlatos neurais da consciência visual.

Há dois pontos relevantes sobre os resultados do estudo de Lau e Passingham que contribuem para o suporte empírico do representacionalismo de segunda-ordem e contraposição das teorias de primeira-ordem. O primeiro ponto diz respeito às avaliações da experiência

subjetiva nas duas condições experimentais e é relevante para contrastar a adequação do representacionalismo de segunda-ordem e da TATG aos resultados do experimento.

Vimos anteriormente que o representacionalismo de segunda-ordem prevê que desempenho em tarefas cognitivas pode estar dissociado da experiencia subjetiva dos sujeitos. Em contraste, a TATG prevê que a presença de experiência subjetiva está vinculada ao desempenho em tarefas cognitivas, de modo que poderia se esperar que um déficit na capacidade de experienciar subjetivamente o estímulo fosse acompanhado de um déficit no desempenho de discriminar esse estímulo, o que não foi o caso observado. Porém, os resultados do experimento de Passingham e Lau demonstraram que em ambas as condições experimentais os sujeitos apresentam os mesmos desempenhos na tarefa de discriminar o estimulo visual, mas há diferenças na avaliação da experiência subjetiva relatada pelos participantes. Na condição em que a distância temporal entre o estímulo visual e o mascaramento visual eram menores, os participantes relataram mais frequentemente estarem apenas adivinhando forçadamente a identificação do estímulo. Por outro lado, quando a distância temporal entre o estímulo e o mascaramento visual eram maiores, os participantes relataram mais frequentemente estarem conscientes de terem experienciado o estimulo visual alvo.

O fato de que há uma diferença no relato sobre a experiência subjetiva na ausência de diferenças significativas no desempenho da tarefa contraria as previsões da TATG e corrobora com as previsões do representacionalismo de segunda-ordem. Não obstante, devemos notar que a observação desse resultado não precisa implicar em uma evidência decisiva contra a TATG.

A TATG prevê que alterações na consciência refletem em alteração no desempenho cognitivo somente se a tarefa cognitiva em questão demanda, para a sua execução competente, a cooperação de múltiplos processadores inconscientes. Em outras palavras, as diferenças ou a manutenção do desempenho em uma tarefa são relevantes para avaliar questões referentes à consciência apenas se a tarefa é tal que demanda as funções desempenhadas pela consciência, tal como a cooperação entre diversos sistemas inconscientes de processamento de informação. Mas, poderia ser alegado que a tarefa de discriminação visual do estudo de Lau e Passingham (2006) não é o tipo de tarefa que demanda um alto nível de esforço cognitivo dos sujeitos. Isto é, é possível que a tarefa de discriminação visual do estudo não exija a cooperação de diversos processadores inconscientes. Se este for o caso, então a realização competente da tarefa não exigirá a função que a TATG atribui a consciência, a saber, disponibilizar globalmente informações para que elas sejam processadas, cooperativamente, por diversos processadores inconscientes. Mas, se não é o caso de que a tarefa de discriminação visual do estudo de Lau e Passingham exige a função desempenhada pela consciência (segundo a proposta da TATG),

então a constância de desempenho em conjunto com a alteração na experiência subjetiva relatada não conta como evidência contrária a alegação da TATG sobre a função da consciência. Isto é, assim como o representacionalismo de segunda-ordem, a TATG pode alegar que o desempenho na tarefa cognitiva em questão, devido a sua simplicidade, é algo que pode ser realizado através de processos inconscientes, pois não exige a coalizão de diversos "processadores".

Embora a observação de que a alteração nos relatos da experiência subjetiva seja acompanhada de uma constância no desempenho da tarefa cognitiva não constitua uma evidência decisiva contra a TATG, essa observação ainda favorece o representacionalismo de segunda-ordem em comparação com outras teorias de primeira-ordem. Considere, por exemplo, o caso da teoria PANIC de Michael Tye (1995). Em sua teoria, estados mentais conscientes são estados cujo conteúdo intencional abstrato e não conceitual está "disponível/pronto" para contribuir de alguma maneira aos processos cognitivos do organismo. Mas, com a exceção da capacidade de relatar a respeito da própria experiência subjetiva, a presença de consciência não demonstra ser relevante para a realização da tarefa de discriminação do estímulo visual, visto que o desempenho na tarefa de discriminação visual das figuras se manteve equivalente em ambas as condições experimentais. Nesse sentido, se pode argumentar que a disponibilidade do conteúdo intencional do estado perceptual em questão para a realização de tarefas de discriminação visual não é algo relacionado à função da consciência fenomênica *per se*. Isto é, parcialmente contra a proposta da teoria PANIC, a disponibilidade do conteúdo intencional para o controle do comportamento e/ou processos de formação de crenças não é um elemento fundamental para a *consciência fenomênica*. Pois, como demonstra o fato de que os mesmos níveis de desempenho foram obtidos nas condições experimentais na qual os sujeitos relatavam ter experienciado subjetivamente o estímulo e na qual os sujeitos não relatavam ter tido a experiência subjetiva, a disponibilidade do conteúdo perceptual para o controle do comportamento na tarefa de discriminação dos estímulos não está vinculada à experiência subjetiva.

O segundo aspecto relevante dos resultados de Lau e Passingham (2006) para a avaliação do representacionalismo de segunda-ordem em comparação com as teorias de primeira-ordem diz respeito aos achados de neuroimagem de possíveis correlatos neurais da consciência. Segundo a interpretação padrão das teorias de primeira-ordem, a consciência fenomênica é determinada apenas pela atividade em áreas sensoriais primárias (LAU; ROSENTHAL, 2011, p.367). Contudo, os resultados do estudo de neuroimagem oferecem evidências que corroboram a hipótese sobre a correlação entre a atividade em áreas corticais

superiores, isto é, no córtex pré-frontal e parietal e a experiência subjetiva, tal como sustenta o representacionalismo de segunda-ordem.

Na condição experimental na qual a distância temporal entre o estímulo e o mascaramento visual era maior, observou-se um nível consideravelmente maior de relatos de que a identificação do estímulo era subjetivamente experienciado, em comparação com a condição experimental na qual a distância temporal era menor (LAU; PASSINGHAM, 2006, p.18765). Junto desse resultado comportamental, observou-se, via métodos estatísticos aplicados às imagens de fMRI, que há uma correlação entre uma maior atividade no córtex pré-frontal dorsolateral esquerdo (áreas de Brodmann, 46 e 9/46) e a frequência de relatos de que o estímulo visual é experienciado subjetivamente: "Nós encontramos uma correlação indicando que quanto maior era atividade no córtex pré-frontal dorsolateral esquerdo, maiores eram as porcentagens dos testes que os sujeitos declararam que eles estavam conscientes da identidade do estímulo" (LAU; PASSINGHAM, 2006, p. 18765).

A constância no desempenho da tarefa cognitiva e a diferença nos relatos a respeito da experiência subjetiva, em conjunto com os achados de neuroimagem corroboram com as previsões do representacionalismo de segunda-ordem, mas não com as previsões das abordagens de primeira-ordem, incluindo a TATG. Algumas versões da TATG (e.g, Baars, 2005) e a teoria PANIC de Tye (1995) vinculam a consciência fenomênica com funções psicológicas de controle comportamental, atenção, memória, motivação entre outras (PRINZ, 2012, p.29). Mas, as evidências resultantes do experimento de Lau e Passingham permitem sustentar que a experiência subjetiva do estímulo não depende, necessariamente, dos mesmos mecanismos que regulam a performance em tarefas de discriminação visual. Assim, dado que há evidências que são incompatíveis com as previsões empíricas de ambos tipos de teorias de primeira-ordem e que são compatíveis com o representacionalismo de segunda-ordem, podemos sustentar que, considerado o experimento em questão, o caso para a acurácia evidencial da representacionalismo de segunda-ordem é comparativamente melhor do que o das suas rivais de primeira-ordem.

Para fortalecer esta conclusão podemos considerar um segundo experimento, realizado por Rounis et al. (2010). Utilizando estimulação magnética transcraniana para inibir a atividade do córtex pré-frontal dorsolateral de ambos os hemisférios cerebrais, Rounis e colegas submeteram sujeitos a um teste de discriminação de um estímulo visual semelhante ao descrito anteriormente (envolvendo a identificação forçada de um estímulo dentre duas possíveis escolhas). O experimento contou com duas seções separadas do teste de discriminação de um estimulo visual, uma seção sem a estimulação magnética transcraniana (TMS) e outra após a

aplicação de estimulação ao córtex pré-frontal dorsolateral em ambos os hemisférios cerebrais. Assim como no experimento de Lau e Passingham, os sujeitos no experimento de Rounis et al (2010, p. 173) não exibiram diferenças significativas no desempenho da tarefa de discriminação visual entre as seções com e sem estimulação magnética. Contudo, após a estimulação do córtex pré-frontal dorsolateral, que resulta numa inibição de sua atividade, os sujeitos reportaram alterações na experiência subjetiva dos estímulos (LAU; ROSENTHAL, 2011, p. 367). Em particular, os autores do experimento sustentam que a inibição do córtex pré-frontal dorsolateral reduziu a eficácia dos sujeitos em distinguir entre os julgamentos corretos e incorretos (isto é, os sujeitos do experimento julgaram mais frequentemente não estarem confiantes da sua classificação do estímulo apesar de estarem corretos) devido a um efeito de redução da visibilidade do estímulo nas tentativas corretas (ROUNIS et al, 2010, p.173). Nesse sentido, o efeito observado se assemelha ao da visão-cega, no qual pacientes negam ter experiências subjetivas do estímulo apesar de serem capazes de discriminar visualmente os estímulos em condições controladas (ROUNIS et al, 2010, p.173). Assim como no experimento anterior, os resultados de Rounis et al. (2010) são mais compatíveis com as previsões empíricas do representacionalismo de segunda-ordem do que com os da TATG ou outras teorias de primeira-ordem (como a teoria PANIC), já que esse segundo grupo de teorias sustenta que a atividade no córtex pré-frontal reflete não apenas a consciência fenomênica, mas a capacidade para a realização de processos cognitivos superiores (LAU; ROSENTHAL, 2011, p. 367).

3 SERIA O DEBATE ENTRE ABORDAGENS DE PRIMEIRA-ORDEM E DE SEGUNDA-ORDEM DA CONSCIÊNCIA FENOMÊNICA UMA NEGOCIAÇÃO METALINGUÍSTICA?

3.1 INTRODUÇÃO

O que faz com que a minha percepção do céu azul ou do canto dos pássaros sejam percepções conscientes? E o que distingue estados mentais fenomênicamente conscientes de estados mentais que não são conscientes nesse sentido? Formular hipóteses que respondam à essas questões é uma das tarefas centrais para quaisquer teorias da consciência. Contudo, entre abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica, há amplas divergências sobre enfrentar essas questões. Neste capítulo, localizamos a origem dessas divergências em uma disputa em torno da adequada caracterização do conceito de consciência fenomênica. Mais precisamente, sustentamos que o debate entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica pode ser interpretado como uma negociação metalinguística a respeito de como caracterizar o *explanandum* de uma teoria da consciência fenomênica. A interpretação metalinguística do debate entre abordagens de primeira e segunda-ordem sobre a consciência tem importantes repercussões. Na medida em que o desacordo entre as abordagens possa ser interpretado como sendo de natureza metalinguística, não há incompatibilidade em supormos que ambas as abordagens possam ser simultaneamente verdadeiras no que diz respeito ao *explanandum* adotado por cada uma delas. Pois, em uma negociação metalinguística, a disputa não se concentra no conteúdo semântico literalmente expresso pelos debatedores, de modo que não é impossível que os conteúdos expressos por ambas abordagens sejam mutuamente verdadeiros. Tal resultado pode ser explorado por uma tentativa de compatibilização das abordagens, caso seja possível situar distintos papéis explanatórios para cada abordagem.

A primeira parte deste capítulo introduz a proposta metodológica das negociações metalinguísticas. O objetivo desta primeira parte do capítulo consiste em identificar e caracterizar quatro condições que, se satisfeitas por um debate, permitem que ele seja interpretado como uma negociação metalinguística. São elas: (1) a condição de disputa, que exige que a suposta negociação metalinguística seja um troca linguística na qual parece se expressar uma disputa; (2) a condição de desacordo, que exige que realmente haja uma desacordo entre os participantes do debate; (3) a condição de equivocação, que exige evidências de que os debatedores compreendem ou utilizam os termos em disputa de maneira distinta e;

(4) a condição da normatividade, que exige evidências de que o desacordo é de natureza normativa e não (apenas) de natureza descritiva (BELLERI, 2020; SANTOS, 2019; PLUNKET; SUNDELL, 2013; PLUNKET; SUNDELL, 2019).

Após a caracterização das quatro condições, a segunda parte deste capítulo argumenta que o debate entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica pode ser frutiferamente interpretado como uma negociação metalinguística. Para tanto, apresentamos evidências de que o debate atende às quatro condições elencadas anteriormente. Dificilmente poder-se-ia contestar que o debate entre abordagens de primeira e segunda-ordem atendem às condições da disputa e do desacordo, pois há uma literatura considerável dedicada a articulação de contra-argumentos e respostas de objeções entre teorias de primeira e segunda-ordem sobre a consciência (e.g, BLOCK, 2011; BROWN; LAU; LeDOUX, 2019; GENNARO, 2004). Contudo, há diferentes maneiras de caracterizar a natureza do desacordo entre estas abordagens. Na seção 3, diferenciamos duas caracterizações do desacordo. Segundo a *caracterização factual*, o desacordo é apenas sobre os fatos que melhor descrevem as condições que um estado mental deve atender para ter a propriedade da consciência fenomênica. Segundo a *caracterização metalinguística*, o desacordo é sobre como compreender o termo **"**consciência fenomênica" que figura no lugar do *explanandum* de uma teoria da consciência. Após sugerir que diferentes condições devem ser atendidas pelo debate para que adotemos uma ou outra caracterização, alegamos que se houverem evidências para sustentar a condição da equivocação, então há maior plausibilidade em adotarmos a caracterização metalinguística.

Na seção 4 buscamos demonstrar que o debate entre abordagens de primeira-ordem e segunda-ordem sobre a consciência fenomênica atende à condição (3). Isto é, que há evidências de que o termo **"**consciência fenomênica" é utilizado diferentemente por cada um dos debatedores. A fim de cumprir com esta etapa do argumento, concedemos inicialmente a intuição de que ambos os disputantes adotam uma caracterização comum de consciência fenomênica, chamada de *concepção nageliana.* Resumidamente, a caracterização nageliana da consciência fenomênica diz que para um estado mental M e um sujeito S, M é consciente, se e somente se, há algo como é para S ter/estar em M (NAGEL, 1974, p. 436; FARRELL, 2017, p.3).

Há diversas passagens que podem ser utilizadas como evidências textual de que ambas as abordagens assumem a concepção nageliana como *explanandum* de suas teorias. Contudo, argumentamos que, devido uma ambiguidade contida nesta caracterização da consciência fenomênica, especificamente na expressão **"**algo como é", há espaço para a defesa de diferentes

interpretações da concepção nageliana e, portanto, espaço para a satisfação da condição da equivocação.

Para caracterizar as diferentes interpretações da concepção nageliana sustentadas por abordagens de primeira e de segunda-ordem sobre a consciência, nos valemos de uma distinção conceitual realizada por Carruthers (2000) entre duas propriedades de estados mentais às quais a expressão **"**algo como é", na concepção nageliana, pode se referir. Por um lado, a expressão pode ser utilizada como se referindo a propriedade de um estado mental de representar o mundo (ou o corpo do próprio organismo) como sendo qualitativamente de determinada maneira. Esta propriedade é chamada de *subjetividade mundana*. Por outro lado, a expressão pode ser utilizada como se referindo a propriedade de um estado mental de estar subjetivamente disponível para um sujeito. Esta propriedade é chamada de *subjetividade experiencial*.

Tendo em vista esta distinção, argumentamos que, se ao atribuir consciência fenomênica a um estado mental as abordagens de primeira-ordem e de segunda-ordem assumem diferentes noções de *subjetividade* como sendo relevantes na aplicação do conceito, então cada abordagem aplica diferentes parâmetros de uso ao conceito consciência fenomênica, o que é uma forte evidência a favor da tese de que o debate satisfaz a condição da equivocação.

Após estabelecermos que o debate satisfaz a condição da equivocação, cabe evidenciar se o debate também satisfaz a condição da normatividade. Na seção 5 alegamos que é plausível interpretar o debate como uma disputa (implícita) sobre como o conceito de consciência fenomênica *deveria ser* utilizado. Esta interpretação é justificada pelo fato de que ambas as abordagens buscam, através das propostas específicas de cada abordagem sobre como estados mentais adquirem consciência fenomênica, advogar por uma compreensão específica (dentre as duas interpretações candidatas elucidadas na seção 4) da concepção nageliana da consciência. Sendo assim, parte do que está em jogo nessa discussão é estabelecer qual é o *explanandum* que deveria ser o foco de uma teoria da consciência.

O capítulo encerra, na seção 6, considerando a possibilidade de uma conciliação da disputa entre abordagens de primeira e segunda-ordem sobre a consciência. Segundo esta perspectiva conciliatória, o fato de que as abordagens de primeira e segunda-ordem adotam diferentes *explananda* permitiria sustentar que aspectos de ambas as abordagens podem ser simultaneamente verdadeiros para os seus respectivos alvos de investigação, de modo que há espaço para uma divisão de tarefas entre elas. Segundo esta perspectiva a consciência fenomênica consiste na mútua instanciação das propriedades da *subjetividade mundana* e *subjetividade experiencial*, de modo que as intuições de ambas as abordagens podem ser complementares na formulação de uma teoria da consciência (LAU; BROWN, 2019).

3.2 NEGOCIAÇÕES METALINGUÍSTICAS: O QUE SÃO? E COMO IDENTIFICÁ-LAS?

3.2.1 O que são negociações metalinguísticas

Negociações metalinguísticas são disputas sobre conceitos. Em tais disputas, os participantes do debate estão engajados em uma negociação a respeito do significado e uso que um determinado conceito *deve* ter em um contexto específico. Para introduzirmos essa noção, considere como exemplo o caso de Callie e Oscar (PLUNKET; SUNDELL, 2013, p.15). Ambos são cozinheiros, e estão cozinhando molho chili juntos. Após ambos provarem o molho, Oscar afirma (a) e Callie afirma (b):

(a) Esse molho está picante!

(b) Não, esse molho não está picante.

Sobre o que Oscar e Callie discordam nessa situação? Parece haver duas alternativas. O desacordo entre Oscar e Callie pode ser interpretado como um desacordo sobre o *fato* de o molho estar picante. Nesse caso, ambos Oscar e Callie concordam sobre o que é para um molho chilli estar picante, mas discordam sobre o sabor do molho. Isto é, uma condição para que o desacordo entre ambos diga respeito exclusivamente aos fatos a respeito da picância do molho é que ambos concordem, previamente, sobre como utilizar o conceito de picância no contexto em questão. Se Oscar e Callie tem diferentes conceitos em mente sobre a picância, então a disputa não seria apenas sobre os fatos a respeito do molho.

Dizer que Oscar e Callie estão em uma disputa sobre os fatos a respeito da picância do molho pode ser pouco natural, porque é improvável que Oscar e Callie estejam discordando sobre a picância do molho com base em uma concepção previamente estabelecida sobre algum **"**limiar de picância" contextualmente relevante para avaliar o sabor do molho. Assim como os termos "alto" ou "pesado", o termo "picante" é contextualmente sensível e pode variar suas condições de aplicação dependendo dos particulares aos quais eles se aplicam. Portanto, parece pouco provável que apenas o sabor do molho seja matéria de disputa entre Oscar e Callie, já que a própria noção de picância é contextualmente variável.

Uma segunda interpretação é a de que o desacordo entre Callie e Oscar não é sobre fatos, mas sobre o conceito de picância. Neste caso, ambos concordam sobre o sabor que o molho tem, mas discordam sobre como empregar o conceito de picância. Ou seja, sua disputa reside

na questão sobre onde fixar o limiar para que algo seja considerado picante (PLUNKETT; SUNDELL, 2013, p.15). É este segundo tipo de disputa, no qual os parâmetros para o uso de um conceito estão em questão, que se caracteriza como uma negociação metalinguística.

Negociações metalinguísticas são disputas normativas em torno do significado e uso de um conceito em um determinado contexto (PLUNKETT; SUNDELL, 2019, p.1; 2013, p.3). Para que um debate seja considerado como uma disputa, tudo o que é necessário é que o debate em questão *pareça*, para os debatedores ou para um observador externo, como uma troca linguística na qual se expressa um desacordo entre os debatedores (PLUNKETT, 2015, p. 835). A noção de desacordo neste contexto é bastante ampla. Algo se configura como um desacordo se há uma espécie de incompatibilidade entre os conteúdos aceitados pelas diferentes partes da disputa. Por exemplo, há um desacordo se o sujeito A defende que *p* e o sujeito B defende *q*, e a sustentação de *p* demanda de um sujeito que ele admita crenças conflitantes com a sustentação de *q* (PLUNKETT, 2015, p. 835; PLUNKETT; SUNDELL, 2013, p. 11).

Negociações metalinguísticas são disputas de natureza *normativa* porque os disputantes estão engajados em estabelecer qual significado e uso um conceito *deve* ter em um determinado contexto (BELLERI, 2020, p.2). O caso de Callie e Oscar ilustra este aspecto na medida em que a disputa entre os dois diz respeito ao significado e aplicação que o termo **"**picante" deve ter no contexto em questão. Enquanto para Oscar **"**picante" é um conceito que se aplica ao molho, para Callie esse conceito não é aplicável ao molho. Portanto, há um desacordo entre os dois sobre quais parâmetros para a compreensão e uso do conceito **"**picante" devem ser adotados neste contexto.

Ainda que negociações metalinguísticas digam respeito a desacordos sobre parâmetros de compreensão e uso de um conceito, não é necessário que essa disputa se desdobre explicitamente nesses termos. Uma negociação metalinguística pode, e frequentemente, ocorre tacitamente através do uso do conceito disputado em declarações. Novamente, o exemplo acima pode auxiliar a elucidar esse ponto. O desacordo entre Callie e Oscar pode não parecer dizer respeito aos parâmetros de compreensão e uso do conceito de picância, pois na superfície do debate não há uma tematização explícita sobre parâmetros de uso do conceito **"**picante". Isto é, Callie e Oscar não utilizam em nenhum momento sentenças como **"**o que eu entendo por picante é..." ou "Algo é picante se e somente se...". Por esta razão, suas divergências sobre o conceito podem nunca ficar evidentes na superfície do debate. Contudo, o uso do termo **"**picante" no diálogo entre Callie e Oscar pode *expressar implicitamente* concepções distintas a respeito do que o termo significa e de como ele deve ser utilizado. Ao fazer uso do termo **"**picante" em

suas respectivas declarações, Callie e Oscar subscrevem de maneira pragmática a diferentes usos do termo picante. Em casos como esse, há um uso *metalinguístico* do termo.

Usos *metalinguísticos* de um termo consistem em usos (e não apenas menções) de um termo com o propósito de comunicar informações sobre como esse termo *deve ser usado no contexto em questão* (PLUNKET; SUNDELL, 2013, p.3). Assim, quando Oscar afirma que o molho está picante, sua declaração expressa implicitamente que o significado do termo picante é tal que a sua atribuição ao molho chili é legítima. Em contraste, o uso de Callie expressa uma concepção de "picante" que não considera como legítima sua atribuição ao molho chili.

Notamos acima que o uso do termo "picante" é governado por diferentes critérios de aplicação para Oscar e Callie em (a) e (b). Mas, se o termo "picante" assume diferentes significados para Callie e Oscar, poder-se-ia pensar que não há um desacordo genuíno entre eles, já que eles sequer estariam falando sobre a mesma coisa. Isto é, poder-se-ia pensar que o caso de Oscar e Callie é equivalente ao seguinte (adaptado de PLUNKETT, 2015):

Oscar*: Há um banco em frente na frente da praça.

Callie*: Não há um banco na frente da praça.

À primeira vista, o caso acima aparenta ser um caso de desacordo. Trata-se de um desacordo sobre a (in)existência de um banco em frente à praça. Porém, suponha que descobrimos posteriormente que Oscar* utiliza o termo "banco" para se referir ao objeto feito para sentar e descansar, e que de fato há um banco nesse sentido do termo na frente da praça. Já Callie* utiliza o termo "banco" para se referir ao estabelecimento no qual se realiza diversas operações financeiras, e não há um banco nesse sentido do termo na frente da praça. Na situação descrita acima, portanto, não há um desacordo genuíno entre Oscar* e Callie*, pois não há uma incompatibilidade entre os conteúdos sustentados por ambos. Isto é, a proposição que Oscar* defende não implica em contradições ou outros tipos de incompatibilidade com a proposição que Callie* sustenta.

Entretanto, o caso de Oscar e Callie não é similar ao caso de Oscar* e Callie*. No caso de Oscar* e Callie*, uma vez que ambos reconhecem que o uso do termo "banco" não está sendo utilizado com o mesmo significado, a aparência de que há uma disputa é derrotada. Mas, no caso de Oscar e Callie, o reconhecimento de que "picante" não possui um significado compartilhado entre ambos não elimina a aparência de que há uma disputa.

Uma razão pela qual tal reconhecimento não elimina a aparência de que há uma disputa é que a troca linguística entre Oscar e Callie não é de natureza descritiva. Eles não estão

buscando, através das suas declarações, dizer o que "picante" significa atualmente (ou o que "picante" significa no idioleto falado por cada um dos participantes nessa troca linguística). A natureza da troca linguística de Oscar e Callie é *normativa*. O que está em jogo no debate entre Oscar e Callie é uma negociação sobre o que o termo "picante" deve significar e, consequentemente, o que podem ser atribuições legítimas de picância[31]. Sendo assim, reconhecer que há dois parâmetros distintos sobre como conceber e utilizar o termo "picante" no debate não elimina o desacordo sobre qual dos dois parâmetros deve ser adotado. E, conforme sugerem os autores do exemplo, a disputa sobre parâmetros rivais de uso do termo "picante" é o tipo de coisa sobre a qual dois falantes podem discordar e sobre a qual vale a pena se engajar em uma disputa, pois tais parâmetros desempenham um importante papel, por exemplo, nas tomadas de decisão coletivas na cozinha (PLUNKETT; SUNDELL, 2013, p.15).

Enfim, com base nas elucidações realizadas a partir do exemplo discutido, podemos sumarizar o que são negociações metalinguísticas da seguinte maneira: uma negociação metalinguística é (1) um caso genuíno de desacordo no qual (2) está (implicitamente) em disputa o significado e uso que um conceito *deve ter* e (3) o uso (metalinguístico) desses conceitos nas expressões dos disputantes *comunica informações sobre como melhor utilizar esse conceito*.

3.2.2 Condições para negociações metalinguísticas e justificação

Tendo em vista a caracterização acima, abordemos agora a questão sobre como identificar que um determinado debate se configura como uma negociação metalinguística. Há quatro condições que podemos identificar para que um determinado debate possa ser *interpretável como* uma negociação metalinguística (BELLERI, 2020 p.4; PLUNKETT, 2015, p. 847; SANTOS, 2019, p. 176)[32]:

1. Condição da disputa: Há uma troca linguística que pode ser descrita como uma disputa ocorrente.

---

[31] Note que as proposições expressas pelas declarações de Callie e Oscar podem ambas ser verdadeiras, mas apenas em virtude do fato de que cada uma delas é orientada por diferentes parâmetros de uso do termo "picante".

[32] A qualificação de que o debate que satisfaz as condições elencadas abaixo é *interpretável* como uma negociação metalinguística é relevante, principalmente, pela seguinte razão. Em muitos casos, os disputantes não reconhecem que o seu desacordo é linguístico, compreendendo antes que sua disputa é sobre fatos. Porém, o fato de que os disputantes tem uma auto concepção do seu debate segundo o qual ele não é sobre termos não impede que um terceiro interprete, plausível e frutiferamente, que o debate se desdobra como uma espécie de negociação metalinguística.

2. Condição do desacordo: Um desacordo está realmente sendo expresso neste debate.

3. Condição da equivocação: Há evidências de que os disputantes neste debate atribuem diferentes significados para a mesma expressão chave.

4. Condição da normatividade: Há evidências de que o suposto desacordo metalinguístico é *normativo* (isto é, diz respeito à questão sobre como *devemos* usar a expressão disputada).

As condições (1) e (2) introduzem os conceitos de *disputa* e *desacordo* como características que um debate deve apresentar para que se possa cogitar que ele é uma negociação metalinguística. Vejamos algumas considerações a respeito destes conceitos. Um debate pode ser considerado uma *disputa* quando ele é uma troca linguística que parece expressar um desacordo entre os participantes. Conforme esclarecem Plunkett e Sundell (2019), para que um debate seja uma disputa, nem sempre é necessário que ele se pareça para os seus participantes como expressando um desacordo. Debates que são disputas podem ser trocar linguísticas que parecem expressar um desacordo apenas para observadores externos ou teóricos que investigam a troca linguística em questão. Em outras palavras, o fato de que um debate expressa um desacordo não precisa ser evidente para os participantes.

Um desacordo é um *estado* de conflito racional entre os participantes de um debate. Por exemplo, se os membros do debate expressam conteúdos conflitantes em suas declarações de tal modo que, sustentar uma das declarações inviabilizaria sustentar racionalmente alguma das declarações do outro membro do debate. Canonicamente, tal conflito pode se dá entre o conteúdo proposicional literalmente expresso pelas declarações de cada participante, mas (não-canonicamente) um desacordo pode ser realizado por outras facetas do conteúdo sustentado por cada participante. Por exemplo, as informações comunicadas via implicaturas ou pragmaticamente pelos participantes do debate podem ser conflitantes, apesar do conteúdo semântico literalmente expresso por cada um deles ser compatível (ver, PLUNKETT; SUNDELL, 2013, p.12-13). A qualificação de que um desacordo é um *estado* serve como um contraste a ideia de que um descordo é uma *atividade* na qual os participantes do debate devem se engajar. Nesse sentido, um desacordo pode ocorrer entre participantes de uma troca linguística sem que eles se engajem explicitamente em um desacordo ou que eles estejam cientes de que o conteúdo que cada um deles sustenta provoca um estado de desacordo (Ver, PLUNKETT; SUNDELL, 2013, p.10; 2019, p.4).

Nem todas as disputas envolvem estados genuínos de desacordo (como é o caso de Callie* e Oscar*) e nem todos desacordos envolvem trocas linguísticas (por exemplo, duas

ações podem estar em desacordo se a realização de uma impede a realização da outra). Mas, para sustentarmos a tese de que um debate é ou pode ser interpretado como uma negociação metalinguística, devem haver evidências de que o debate envolve algum tipo de troca linguística (i.e., escrita ou falada) e que tal debate parece expressar um desacordo. É importante notar, porém, que não é preciso que o desacordo expresso seja evidente para os próprios debatedores.

As condições (3) e (4) são menos intuitivas do que as duas anteriores, e encontrar evidências de que um debate satisfaz tais condições pode ser uma tarefa difícil. Com relação a condição (3), o que conta como evidência de que os debatedores que participam de uma determinada disputa atribuem diferentes significados para um determinado termo? Um conjunto relevante de evidências consiste no reconhecimento de que, dadas as mesmas condições, os disputantes exibem sistematicamente diferentes disposições para aplicar o termo disputado (PLUNKETT, 2015, p.847; BELLERI, 2020, p. 4). Pois, se não há diferenças nas condições factuais em que os participantes do debate se baseiam para aplicar um conceito C, então as diferentes disposições entre os participantes do debate para aplicar C são o resultado da adoção de diferentes concepções do significado de C. Novamente, o caso de Oscar e Callie ilustra o ponto em questão. Ambos têm acesso e compartilham das mesmas informações factuais sobre o molho chili, por exemplo, o seu sabor ou quanto tempero já foi usado no seu preparo. Porém, apesar de compartilharem destas informações, há um desacordo sobre a atribuição do termo **"**picante" ao molho. Mas, como não há uma diferença nas condições factuais que servem de base para a aplicação do termo **"**picante", então o que pode estar por trás das diferenças na disposição entre Oscar e Callie de utilizar o conceito de picância pode ser uma diferença nas concepções que cada um deles adotado sobre o significado desse conceito. Por exemplo, pode ser o caso de que Callie utilize como parâmetro um limiar mais alto na escala de Scoville[33] quando se trata de atribuir picância a um molho chilli do que Oscar.

Com relação a condição (4), uma das principais maneiras de verificar que um debate envolve uma discussão normativa sobre o significado consiste em verificar que o debate não se encerra uma vez que os membros do debate reconhecem que estão empregando conceitos distintos. Isto é, quando se reconhece que a origem da disputa é uma divergência conceitual e não sobre os fatos, a disputa entre os debatedores pode continuar, de maneira frutífera, através da questão sobre como deveríamos estar utilizando o conceito em torno do qual há divergências sobre o seu significado ou aplicação (PLUNKETT; SUNDELL, 2013, p.15). Disputas metalinguísticas normativas concernem questões práticas sobre como um conceito deve ser

[33] A escala Scoville é utilizada para mensurar a pungência das pimentas, baseada na concentração de capsaicina que elas possuem.

usado em um determinado contexto e/ou questões avaliativas a respeito das virtudes ou defeitos das diversas opções disponíveis para o uso e compreensão de um conceito para um determinado contexto (PLUNKETT; SUNDELL, 2019, p. 6).

Nem toda disputa que é reconhecida como metalinguística é normativa, sendo algumas delas de natureza descritiva. Disputas metalinguísticas descritivas são aquelas que se concentram em como um determinado conceito está *atualmente* sendo utilizado em um determinado contexto de enunciação (PLUNKETT; SUNDELL, 2019, p.6). Assim, por exemplo, uma disputa pode satisfazer os critérios (1)-(3) e falhar em satisfazer (4) porque não envolve uma disputa sobre o que um conceito *deveria* significar em um determinado contexto, mas apenas sobre o qual é o significado atual deste conceito em um determinado contexto de enunciação. Em casos nos quais a disputa não se concentra em questões sobre a normativas sobre o significado de um termo chave, o debate não será considerado uma negociação metalinguística.

Sustentar que uma determinada disputa satisfaz as condições (3) ou (4) pode ser uma tarefa complicada. Em algumas situações, os membros do debate podem reconhecer que há uma disputa conceitual em jogo, mas ainda assim situar o ponto mais relevante de suas divergências como tendo origem em um desacordo no plano dos fatos. O caso de Oscar e Callie poderia se desdobrar de uma maneira que ilustra esse ponto: suponha que Callie e Oscar reconheçam que usam o conceito de picância de maneiras distintas. Porém, ao invés de enxergar a raiz da sua disputa apenas no uso de conceitos distintos, eles sugerem que a raiz da disputa se encontra nas distintas percepções que cada um deles têm a respeito da concentração de capsaicina (a substância que garante picância às pimentas).[34] Nesta situação, a disputa sobre se o molho chilli está ou não picante é entendida com um sintoma das diferentes disposições perceptuais de cada cozinheiro, e apenas contingentemente sobre os conceitos que cada um deles têm sobre a picância. Pois, o conceito de picância que cada um deles utiliza é regulado pelas diferentes disposições perceptuais que cada um deles possui sobre o sabor do molho. Sendo assim, aos olhos dos disputantes, o desacordo não é metalinguístico, ainda que possa ser normativo (já que

[34] Note que a disputa, tal como compreendida por Oscar e Callie nesta situação não diz respeito a concentração da substância capsaicina. Tal disputa poderia ser resolvida, por exemplo, através de um teste cromatográfico capaz de medir a concentração dessa substância. Antes, Callie e Oscar interpretam sua disputa como envolvendo discrepâncias na percepção subjetiva da concentração, que pode variar de acordo com vários fatores, como a constituição fisiológica dos indivíduos e o nível de fadiga sensorial apresentado pelo indivíduo no momento da degustação. Neste caso, é difícil estabelecer uma maneira de testar empiricamente a discrepância entre as percepções de Oscar e Callie. E, mesmo que seja possível estabelece-la, realizar tal teste não inviabilizaria a possibilidade de que além da diferença perceptual, ambos utilizam de parâmetros conceituais distintos para a atribuição de picância ao molho chilli.

a disputa pode então se tratar sobre qual disposição perceptual deveria ser adotada como norma para o uso comum da noção de picância).

O que este possível desdobramento do caso de Callie e Oscar revela é uma dificuldade geral que pode ser enfrentada ao buscarmos estabelecer que uma determinada disputa *é* uma negociação metalinguística: conceber uma disputa como metalinguística pode ser uma interpretação que contraria as intuições dos próprios debatedores a respeito de como enxergar a natureza do seu debate.

Há outros casos em que é possível que os participantes em um debate não reconhecem que o ponto em disputa é de natureza metalinguística. Os participantes de um debate podem reconhecer que utilizam um determinado conceito de maneiras distintas, mas supor que essa divergência é decorrente de crenças conflitantes que regulam o uso desse conceito (BELLERI, 2020, p.4). No exemplo de Callie e Oscar, essa situação consistiria em um cenário no qual Callie e Oscar reconhecem que utilizam de maneira distinta o termo **"**picante", mas no qual insistem que essa divergência remete a crenças conflitante sobre como avaliar a picância de molhos chilli.

Os participantes de um debate podem também conceder que sua disputa é metalinguística, mas que não é normativa, mas descritiva. Em tais situações, os participantes do debate concebem sua disputa como tendo por objetivo estabelecer o significado ou os parâmetros de uso que, de fato, o conceito em disputa possui no contexto de enunciação relevante. Assim, por exemplo, Oscar e Callie poderiam converter o seu debate sobre a picância do molho em um debate sobre como tal noção é usualmente entendida no contexto de enunciação da sua cozinha ou no contexto de enunciação mais geral da culinária. Nesta situação, o debate não seria considerado uma negociação metalinguística.

Tendo em vista as situações mencionadas acima, o que um defensor da interpretação metalinguística poderia fazer em prol da sua causa? É teoricamente custoso negar as intuições dos debatedores e insistir que a disputa em questão é, de fato, metalinguística (BELLERI, 2020, p.5). Pois, em tais casos, o defensor da interpretação metalinguística deverá oferecer razões para atribuir crenças falsas aos debatedores. Ou seja, o defensor de tal interpretação deverá explicar porque são falsas as interpretações dos debatedores sobre as suas próprias disputas e o que poderia tê-los levado a formar tais crenças equivocadas de que suas disputas consistem em divergências sobre fatos e não sobre conceitos.

Felizmente, há uma estratégia alternativa mais modesta e teoricamente menos custosa: ao invés de sustentarmos que um debate *é uma negociação metalinguística*, podemos sustentar que um debate *pode ser interpretado como uma negociação metalinguística* (BELLERI, 2020,

p.5). Para sustentar esse posicionamento, tudo o que precisamos é mostrar que um debate pode ser interpretado satisfatoriamente como atendendo as condições (1)-(4). Note que, neste caso, não é preciso demonstrar que, de fato, o debate satisfaz as condições especificadas em (1)-(4). Tudo o que é necessário é sustentar que, avaliadas às evidências de que dispomos sobre o debate em consideração, não há evidências disponíveis que derrotam essa possibilidade interpretativa. Desta maneira, o ônus de provar que um debate não é interpretável como uma negociação metalinguística se volta contra o oponente da interpretação da negociação metalinguística (BELLERI, 2020, p.6).

No presente capítulo sustentaremos essa estratégia interpretativa mais modesta. Nas próximas seções deste capítulo sustentaremos que é possível interpretar o debate entre defensores de teorias de primeira-ordem e de segunda-ordem sobre a consciência fenomênica como envolvendo uma negociação metalinguística. Nesse sentido, na seção 3 examinaremos o debate entre essas abordagens em busca de evidências de que ele pode ser interpretado como satisfazendo às condições (1) e (2). Ou seja, que o debate em questão pode ser interpretado como (1) uma troca linguística que pode ser descrita como uma disputa e que (2) expressa implícita ou explicitamente um desacordo. Com esse propósito, veremos inicialmente como interlocutores deste debate caracterizam a relação entre as abordagens de primeira e segunda-ordem. O resultado dessa etapa é uma caracterização inicial da disputa como um desacordo a respeito das condições de existência de estados mentais fenomênicamente conscientes que gira em torno do status do princípio da transitividade. Embora essa seja uma interpretação natural da relação entre as abordagens de primeira e segunda-ordem sobre a consciência fenomênica, buscaremos argumentar que ela não é a única interpretação possível. Uma maneira alternativa de interpretar o desacordo expresso no debate consiste em localizar a origem do desacordo entre abordagens de primeira e segunda-ordem no significado que a noção de consciência fenomênica deveria assumir no contexto da elaboração de uma teoria sobre esse fenômeno. Nesse sentido, o desacordo sobre as condições de existência de estados mentais fenomênicamente conscientes é um reflexo do desacordo mais fundamental a respeito do conceito de consciência fenomênica, que opera como *explanandum* de uma teoria.

Nas seções 4 e 5 examinamos se, assumida a interpretação de que o desacordo é sobre o conceito de consciência fenomênica, há evidências satisfatórias que permitam sustentam que o debate atende as condições (3) e (4), respectivamente. Isto é, se no suposto desacordo metalinguístico em questão (3) há evidências de que há a atribuição de diferentes significados para a mesma expressão chave e (4) se há evidências de que o suposto desacordo metalinguístico é *normativo* (isto é, diz respeito à questão sobre como *devemos* usar a expressão

disputada). Na seção 6 concluímos a discussão sintetizando algumas das consequências e possíveis desdobramentos de nossa hipótese de que é possível interpretar o debate entre abordagens de primeira-ordem e segunda-ordem sobre a consciência fenomênica como sendo uma negociação metalinguística.

3.3 O DESACORDO ENTRE ABORDAGENS DE PRIMEIRA E SEGUNDA-ORDEM DA CONSCIÊNCIA FENOMÊNICA: DUAS INTERPRETAÇÕES

Para que possamos sustentar que um determinado debate pode ser interpretado como uma negociação metalinguística, devem haver evidências de que ele atende as quatro condições apresentadas na seção anterior. Nesta seção examinamos se há evidências de que o debate entre abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica atendem as condições (1) e (2). Ou seja, se há evidências de que o debate entre abordagens de primeira e segunda-ordem sobre a consciência fenomênica são (1) uma troca linguística que pode ser descrita como uma disputa, que (2) realmente expressa, implícita ou explicitamente, um desacordo.

Vimos que uma disputa é uma troca linguística que *parece* expressar um desacordo (PLUNKETT; SUNDELL, 2019, p.4). E, um desacordo é, grosso modo, definido como um *estado* no qual há algum tipo de conflito racional entre o conteúdo sustentado pelos participantes da disputa (PLUNKETT; SUNDELL, 2019, p.5). Nesse sentido, podemos nos voltar a literatura publicada a respeito da relação entre abordagens de primeira-ordem e de segunda-ordem sobre a consciência em busca de evidências de que há entre essas abordagens uma troca linguística que parece expressar um desacordo.

A literatura que trata sobre a consciência fenomênica é repleta de exemplos que apresentam o debate entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência como uma disputa (ver, CARRUTHERS, 2000; BLOCK, 2011; ROSENTHAL, 2012; GOTTLIEB, 2016; LAU; BROWN, 2019). Eis um par de exemplos de caracterizações do debate entre teorias de primeira-ordem e segunda-ordem que enquadram o debate como orientado em torno de um conflito sobre uma das teses centrais a sustentação das abordagens de segunda-ordem, o status do princípio da transitividade:

> [O princípio da transitividade] ocupa um lugar curioso na literatura sobre a consciência. Por um lado, ele é um divisor de águas entre as duas maiores famílias teóricas: teóricos de segunda-ordem, que o aceitam e teóricos de primeira-ordem, que o rejeitam. [...] Teorias de primeira-ordem são melhor descritas negativamente: uma teoria T é uma teoria de primeira-ordem da consciência se, de acordo com T, representações de segunda-ordem

> de um estado mental não são necessárias para que um estado mental seja consciente. [...] Por outro lado, muitos teóricos de segunda-ordem acreditam que o [princípio da transitividade] é *obviamente* verdadeiro (GOTTLIEB, 2016, p. 353, tradução e colchetes nossos, grifos do autor).
>
> Uma divisão central entre teorias filosóficas da consciência fenomênica é aquela entre abordagens de primeira-ordem e de segunda-ordem. [...] a divisão entre teorias de primeira-ordem e teorias de segunda-ordem corresponde, grosso modo, a aceitação do que David Rosenthal chama de princípio da transitividade, que sustenta que um estado perceptual ser *fenomênicamente* consciente consiste na consciência que um sujeito tem de si mesmo como estando nesse estado (de alguma maneira apropriada) ou, inversamente, que um estado perceptual de que um sujeito não está de modo algum consciente de estar não pode ser fenomênicamente consciente (LAU; BROWN, 2019, p. 171)

Há três elementos comuns entre as passagens sobre a caracterização do debate. Em primeiro lugar, ambas as passagens apresentam as teorias de primeira-ordem e de segunda-ordem sobre a consciência como abordagens que rivalizam entre si no âmbito da articulação de uma explicação sobre a consciência fenomênica. Em segundo lugar, ambas as passagens caracterizam a relação de rivalidade entre estas abordagens teóricas como se concentrado sobre um desacordo sobre as condições necessárias e/ou suficientes para que um estado mental seja considerado consciente em sentido fenomênico. Por fim, em terceiro lugar, as passagens convergem na identificação da aceitação ou rejeição do princípio da transitividade como uma condição necessária para que estados mentais sejam fenomênicamente conscientes como estando no centro desse desacordo. Enquanto defensores de uma abordagem de primeira-ordem rejeitam a ideia de que o princípio da transitividade é uma condição necessária para a consciência fenomênica de estados mentais, defensores de abordagens de segunda-ordem assumem que o princípio da transitividade é uma condição necessária para a consciência fenomênica de estados mentais.

Sendo assim, a impressão inicial que podemos formar sobre o debate é o de que ele claramente atende às condições (1) e (2). Há uma troca linguística que pode ser caracterizada com uma disputa[35] e ela parece ser, fundamentalmente, centrada sobre as condições necessárias para que um estado mental seja consciente em sentido fenomênico. Em outras palavras, há um desacordo sobre o status do princípio da transitividade como uma condição necessária para que estados mentais sejam fenomênicamente conscientes.

Contudo, caso o desacordo entre as abordagens de primeira e segunda-ordem sobre a consciência seja somente sobre o status do princípio da transitividade enquanto uma condição

[35] Esse ponto é também corroborado pelo volume da literatura que se concentra em argumentar favoravelmente a uma das abordagens através da articulação de objeções à abordagem rival ou através da resposta às objeções da abordagem rival. Alguns exemplos são: BLOCK (2011); BYRNE (1997); CARRUTHERS (2000); MANSON (2002); WEISBERG (2011).

necessária para consciência fenomênica de estados mentais, o debate se concentraria apenas sobre questões a respeito da metafísica da consciência e não questões a respeito do significado ou uso desse conceito. Porém, uma leitura mais cuidadosa revela que as coisas não são tão simples quanto podem parecer à primeira vista.

Ambas as passagens apresentarem indícios de que há um desacordo explícito neste debate que concerne a aceitação ou rejeição do *princípio da transitividade* como uma condição necessária para que um estado mental exiba consciência fenomênica. Entretanto, não há só uma interpretação possível do desacordo que parece ser expresso nesse debate. Além da interpretação segundo a qual o desacordo é explicitamente sobre as condições em virtude das quais um estado mental é fenomênicamente consciente, há uma interpretação alternativa segundo a qual o descordo é, implicitamente, sobre o significado e uso do termo "consciência fenomênica". Vejamos a seguir o que caracteriza cada uma destas interpretações e como podemos justificá-las.

Há diferentes maneiras através das quais se pode interpretar a natureza da disputa ocorrente no debate entre abordagens de primeira-ordem e segunda-ordem sobre a consciência. Podemos diferenciar duas interpretações possíveis sobre o que está em jogo no desacordo entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência, as nomearemos, respectivamente, de interpretação factual e interpretação metalinguística.

3.3.1 Interpretação factual

Segundo a *interpretação factual*, o que está em jogo no debate é um conflito sobre qual é o conjunto de fatos – se aqueles elencados pela abordagem de primeira-ordem ou a abordagem de segunda-ordem – que oferece a descrição mais acurada ou verdadeira das condições necessárias para que um estado mental exiba consciência fenomênica. Segundo esta interpretação, o desacordo ocorrente neste debate consiste na sustentação de maneiras incompatíveis de elencar as condições para que um estado mental exiba consciência fenomênica. Tais descrições incompatíveis consideram que diferentes condições factuais devem ser atendidas para a atribuição de consciência fenomênica. Enquanto defensores de abordagens de segunda-ordem sustentam que um sujeito estar transitivamente consciente[36] de um estado mental M é uma condição necessária para que M exiba consciência fenomênica,

[36] No capítulo 1, seção 2.4, definimos a consciência transitiva de estado em termos de um estado mental M em virtude do qual um sujeito se torna consciente de algo P.

defensores de abordagens de primeira-ordem rejeitam essa condição (ver, GOTTLIEB, 2016, p.353; 2018, p. 320).

Note que a interpretação factual do desacordo pressupõe que ambas as abordagens envolvidas no debate empregam uma noção comum de consciência fenomênica. Ou seja, pressupõe-se que ambas as abordagens usam o mesmo conceito de "consciência fenomênica" para identificar o alvo de explicação de suas respectivas teorias, de modo que suas propostas de critérios necessários para a consciência fenomênica são *explanans* incompatíveis sobre o mesmo *explanandum*. Essa é uma condição fundamental para que a interpretação factual se sustente porque, do contrário, haveria espaço para duvidar da legitimidade do desacordo entre as abordagens de primeira e de segunda-ordem (ver, CHALERMS, 2011).

Caso o alvo da explicação de cada abordagem fosse distinto, seria possível sustentar que a rejeição do princípio da transitividade por parte das teorias de primeira-ordem é compatível com a admissão desse princípio por parte das teorias de segunda-ordem. Para ver com isso poderia ocorrer, considere o caso em que os teóricos de primeira-ordem estão discutindo as condições necessárias para a existência da consciência fenomênica$_1$, enquanto teóricos de segunda-ordem estão discutindo as condições necessárias para a existência de consciência fenomênica$_2$[37]. Nessa situação, as abordagens não empregam conceitos idênticos para denotar o alvo da explicação das suas respectivas teorias. Mas, se os conceitos não são idênticos, então é logicamente possível que as condições necessárias para a atribuição desses conceitos também sejam distintas. Consequentemente, seria racionalmente possível rejeitar que o princípio da transitividade é uma condição necessária para consciência fenomênica$_1$ e admitir que o princípio da transitividade é uma condição necessária para consciência fenomênica$_2$.

Esta hipótese ganha em plausibilidade caso se argumente que os conceitos empregados por cada abordagem denotam diferentes fenômenos na realidade, já que então podem haver diferentes condições de existência para esses fenômenos. Se os conceitos de consciência fenomênica$_1$ e consciência fenomênica$_2$ denotarem diferentes fenômenos na realidade, tal caso será similar ao exemplo de Oscar* e Callie* (ver seção 2.1). Oscar* e Callie* utilizam o termo homônimo "banco" em suas declarações, mas não utilizam o mesmo conceito e nem buscam denotar a mesma entidade na realidade. Enquanto as declarações de Oscar* são sobre um objeto feito para sentar e descansar, as declarações de Callie* são sobre uma instituição financeira. Nesse sentido, as condições necessárias para a verdade da proposição "há um banco na minha frente" são extremamente distintas se proferidas por Oscar* ou por Callie*. Os conceitos de

[37] Utilizamos aqui o que Chalmers (2011, p.532) chama de gambito do subscrito.

“banco” adotados por Oscar* e Callie* são orientados por diferentes parâmetros de uso. Entre tais parâmetros, constam a satisfação determinados fatos na realidade. A declaração de Oscar* de que “há um banco na minha frente” pode ser verdadeira quando há na realidade um objeto para sentar e descansar na sua frente (e falsa quando não há); e a declaração de Callie* de que “há um banco na minha frente” pode ser verdadeira quando há uma instituição financeira na sua frente (e falsa quando não há).

Similarmente, quando simpatizantes de abordagens de primeira-ordem utilizam o conceito de consciência fenomênica em suas declarações sobre o tema, pode ser verdadeiro que a ocorrência do fenômeno em questão não envolve a satisfação do princípio da transitividade. E quando um teórico da abordagem de segunda-ordem utiliza o conceito de consciência fenomênica em suas declarações sobre o tema, pode ser verdadeiro que a ocorrência do fenômeno em questão exige a satisfação do princípio da transitividade. O que tornaria essa situação possível é uma situação em que diferentes porções da realidade são representadas pelos conceitos utilizados pelos teóricos simpatizantes das abordagens de primeira-ordem e de segunda-ordem sobre a consciência fenomênica. Sendo assim, teorias de primeira e de segunda-ordem sobre a consciência fenomênica estariam, na realidade, buscando explicar diferentes fenômenos, de modo que o conteúdo expresso por cada uma das teorias não incidiria em um desacordo legítimo a respeito dos fatos.

Não é nosso propósito aqui oferecer evidências para sustentar que diferentes porções da realidade são, de fato, representadas pelos conceitos de consciência fenomênica das abordagens de primeira e de segunda-ordem. Nossa proposta é ligeiramente menos audaciosa: sugerimos apenas que essa é uma hipótese possível. Assim, ao levantá-la, buscamos notar que cabe ao defensor da interpretação factual demonstrar que tal hipótese é falsa. Na seção 4 deste capitulo, retomaremos essa hipótese, aprofundando como um dos defensores de uma abordagem de segunda-ordem articula essa ideia.

3.3.2 Interpretação metalinguística

Uma segunda interpretação possível sobre o desacordo entre as abordagens de primeira e segunda-ordem sobre a consciência consiste em interpretá-lo como remetendo a uma disputa sobre termos, e mais exatamente, uma disputa de *metalinguística normativa*. Segundo a *interpretação metalinguística*, o desacordo entre as abordagens é terminológico e implícito – o desacordo não é abertamente sobre o conteúdo semântico expresso nas declarações realizadas no debate (ver, PLUNKET; SUNDELL, 2019, p.5). Para essa interpretação do desacordo, ele

consiste na sustentação de concepções conflitantes sobre como deveríamos utilizar o conceito de consciência fenomênica ao elaborar uma teoria explicativa sobre esse tema. De tal modo, o desacordo é centrado em como deveríamos conceber o *explanandum* de uma teoria da consciência fenomênica e sua expressão no debate se dá pelo uso desse conceito nas declarações de cada teoria.

O uso do conceito de consciência fenomênica por parte de cada abordagem é metalinguístico porque este conceito é utilizado para comunicar perspectivas a respeito de como o próprio conceito deve ser compreendido e usado (ver, PLUNKETT; SUNDELL, 2019, p.1). E ele é normativo na medida em que, o uso do conceito em suas declarações expressa como o conceito deve ser entendido e usado no contexto de uma explicação da consciência (em contraste com o modo como o conceito é atualmente usado nesse contexto). Isto é, ao utilizá-lo em suas declarações, cada abordagem está oferecendo, implicitamente, informações sobre como deveríamos conceber e empregar o conceito de consciência fenomênica no contexto de uma teoria explicativa desse fenômeno mental. Nesse sentido, enquanto as declarações envolvendo o conceito de consciência fenomênica utilizado por defensores de uma abordagem de primeira-ordem são tais que a satisfação do princípio da transitividade não consta entre os parâmetros que regulam a aplicação desse conceito, defensores de abordagens de segunda-ordem utilizam o conceito de consciência fenomênica tendo como parâmetro a necessária satisfação do princípio da transitividade por parte do estado mental em questão.

Contra a interpretação metalinguística do desacordo poder-se-ia objetar que ela não enquadra corretamente a disputa que é expressa em boa parte da literatura relevante sobre o debate entre abordagens de primeira e segunda-ordem. Segundo esta objeção, a interpretação metalinguística se equivoca ao sugerir que o desacordo é sobre como deveríamos usar ou entender o conceito de consciência fenomênica quando, na verdade, o desacordo é sobre se o princípio da transitividade é ou não uma das condições necessárias para que um estado mental exiba a propriedade que é denotada pelo conceito de consciência fenomênica. Em resposta a essa objeção, podemos observar o seguinte: pessoas que discordam sobre as condições para que algo seja classificado como exibindo a propriedade X podem discordar dessas condições porque elas representam de maneiras distintas o que é X. Em outras palavras, as divergências nos níveis dos fatos podem ser originárias de divergências conceituais iniciais (ver, CHALMERS, 2011, p. 525n8; p. 528). Assim, no caso do debate a respeito da consciência fenomênica, a divergência sobre qual é o conceito que deve ocupar o lugar do *explanandum* de uma teoria da consciência não é incompatível com a existência de divergências sobre o status do princípio da transitividade como uma condição necessária para a consciência fenomênica. Com efeito, pode-

se alegar que tal divergência tem origem no fato de que diferentes conceitos são adotados como *explanandum* por cada uma das abordagens.

O caminho inverso também é possível – uma divergência sobre quais fatos são relevantes para a explicação do que é o fenômeno a ser explicado podem conduzir à compreensão e uso distintos do conceito de "consciência fenomênica". Ou seja, o desacordo sobre um termo pode também ser um produto de desacordos no nível dos fatos. Em ambos os casos, porém, há conjuntamente às divergências sobre os fatos, uma divergência sobre o conceito de consciência fenomênica. E, na medida em que tal divergência possa ser interpretada como dizendo respeito a como esse conceito *deveria* ser usado, será possível interpretar o debate como envolvendo uma negociação metalinguística.

Ambas as interpretações do desacordo expostas acima são compatíveis com a evidência textual de que os participantes no debate entre as teorias de primeira-ordem e de segunda-ordem da consciência compreendem seu desacordo como dizendo respeito a admissão ou rejeição do princípio da transitividade. Porém, diferentes condições têm que ser atendidas para que cada uma das interpretações sejam plausíveis sobre o debate. A interpretação factual presume que há uma noção de consciência fenomênica que é comum às duas abordagens e exige que tal pressuposição prove ser verdadeira para que se sustenta a ideia de que o desacordo é (exclusivamente) a respeito da admissão/rejeição de um ou outro conjunto de fatos como o conjunto de fatos que melhor descreve as condições necessárias para a consciência fenomênica. Em contraste, a interpretação metalinguística sugere que há diferenças na concepção e uso do conceito de consciência fenomênica entre as abordagens, e que tal diferença tem como reflexo as divergências factuais que aparecem mais evidentemente na superfície do debate. Ou, alternativamente, que divergências factuais acabam por refletir em divergências conceituais. Seja como for, para que possamos defender a plausibilidade da interpretação metalinguística do desacordo, devemos mostrar que há evidências de que o debate satisfaz também às condições (3) da equivocação e (4) da normatividade do significado. Isto é, que existem evidências de que há a atribuição de diferentes significados para alguma expressão chave e de que a disputa em torno do significado dessa expressão não é meramente descritiva. Argumentar que o debate satisfaz essas condições são o objetivo das próximas seções.

3.4 UMA DISPUTA SOBRE TERMOS: A CONCEPÇÃO NAGELIANA E SUA AMBIGUIDADE

Esta seção trata sobre a possibilidade de interpretarmos a disputa entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência como sendo sobre termos. Argumentaremos que o debate entre as abordagens de primeira-ordem e de segunda-ordem sobre a consciência satisfaz a condição da equivocação porque cada abordagem compreende diferentemente o significado do termo "consciência fenomênica". Particularmente, cada abordagem assume que o termo "consciência fenomênica" denota diferentes propriedades de estados mentais. Consequentemente, elas discordam sobre os parâmetros de uso desse conceito e o que é necessário de uma teoria da consciência para explicá-lo.

A presente seção inicia concedendo que há uma concepção de consciência que aparenta ser compartilhada por ambas as abordagens: a *concepção nageliana*. Apresentar o que é a concepção nageliana e evidenciar que ambas as abordagens estão comprometidas com o seu uso são os objetivos da subseção 4.1. Na subseção 4.2 argumentaremos que a concepção nageliana é ambígua: a expressão "algo como é", que desempenha um papel central na designação das condições necessárias e suficientes que um estado mental deve atender para ter consciência fenomênica pode denotar duas diferentes propriedades de um estado mental. De tal modo, há espaço para divergências sobre o significado do termo "consciência fenomênica", porque são possíveis diferentes leituras da concepção nageliana e dos seus parâmetros de uso. A subseção 4.3 introduz dois conceitos utilizados por Carruthers (2000) para caracterizar e distinguir as diferentes propriedades que podem ser denotadas pela expressão "algo como é". De posse dessa distinção conceitual, exemplificamos na subseção 4.4, através da discussão sobre o tema da consciência animal, como pode ser frutífero ler uma das divergências entre abordagens de primeira-ordem e de segunda-ordem da consciência como tendo origem na adoção de diferentes interpretações da concepção nageliana, conforme articuladas na seção anterior.

3.4.1 A recepção da concepção nageliana da consciência fenomênica entre abordagens de primeira-ordem e de segunda-ordem

A concepção nageliana da consciência fenomênica remete a seguinte famosa passagem do artigo "Como é ser um morcego" de Thomas Nagel (1974):

> [...] o fato de que um organismo tem experiência consciente significa, basicamente, que há algo como é ser esse organismo. Podem haver outras implicações sobre a forma da experiência; pode até haver implicações adicionais sobre o comportamento do organismo (porém, eu duvido disso). Mas, fundamentalmente um organismo tem estados mentais conscientes se e somente se há algo como é ser esse organismo – algo como é para o organismo. (NAGEL, 1974, p. 436, tradução nossa, adaptado).

A expressão **“**algo como é*”* ocupa um lugar central nesta passagem. Nagel oferece, através dela, uma caracterização da consciência fenomênica em termos das condições necessárias e suficientes para que um estado mental seja consciente (em sentido fenomênico). A partir da caracterização oferecida por Nagel nessa passagem, tornou-se comum a associação entre estados mentais fenomênicamente conscientes e estados que possuem a elusiva propriedade de ter **“**algo como é” vivenciá-los.

A popularidade da concepção nageliana na literatura filosófica contemporânea é inegável[38]. Frequentemente encontrarmos menções diretas à expressão **“**algo como é” ou expressões equivalentes no contexto de introdução e definição da noção de consciência fenomênica (e.g., KIM, 2010, p. 15; BRADDON-MITCHELL; JACKSON, 2006, p.129). A situação não é diferente no que concerne os principais participantes do debate entre as abordagens de primeira-ordem e de segunda-ordem da consciência. Em sua maioria, encontramos alguma menção à expressão **“**algo como é” neste debate, especialmente no contexto da introdução ou caracterização inicial do que é a consciência fenomênica. Assim, por exemplo, Ned Block (1995), um dos principais defensores de abordagens de primeira-ordem, introduz o conceito de consciência fenomênica (P-consciência na passagem abaixo) através de uma variante da expressão nageliana **“**algo como é”:

> Como deveríamos apontar para a P-consciência? Uma maneira é através de sinônimos aproximados. Como eu havia dito, P-consciência é experiência. Propriedades da P-consciência são propriedades experienciais, ou seja, um estado é P-consciente se ele tem propriedades experienciais. A totalidade de propriedades experienciais de um estado são “como é ter” esse estado (BLOCK, 1995, p. 230, tradução nossa).

Outro defensor da abordagem de primeira-ordem sobre a consciência, Michael Tye (1995), também utiliza a expressão **“**algo como é” para caracterizar a consciência fenomênica de uma maneira similar:

> À medida que as qualidades fenomênicas ou sentidas da experiência mudam, as próprias experiências mudam, mas se o aspecto fenomenal ou “como é” desaparecer por inteiro, então não há experiência alguma. [...] um estado mental é um estado fenomênicamente consciente apenas se há algum “sentir” subjetivo para o estado, alguma qualidade experiencial distintiva (TYE, 1995, p. 3, tradução nossa).

Defensores de abordagens de primeira-ordem da consciência não são os únicos a mencionar e adotar alguma versão da concepção nageliana no contexto de introdução ou

[38] Ao comentar sobre a expressão “algo como é” [what *it’s like*] de Nagel, Güzeldere (1995, p. 36, tradução nossa) diz que essa noção “[...] tem sido tão influente que ela parece ser onipresente em diversos, e distintos (ainda que relacionados), problemas relativos à consciência”. Esta constatação leva o autor a chamar essa expressão de um “coringa” da discussão sobre a consciência, na medida em que ela é evocada para tratar problemas tão diversos como: a realidade metafísica dos *qualia*, a existência de pontos de vista subjetivos ontologicamente irredutíveis e demonstração da falsidade do *fisicalismo*, enquanto uma doutrina metafísica (via argumento do conhecimento).

caracterização da noção de consciência fenomênica. Encontramos a mesma tendência também entre os defensores de abordagens de segunda-ordem. Carruthers (2000) introduz a noção de consciência fenomênica através de uma menção direta à concepção nageliana da consciência:

> A mais óbvia e chamativa (e a mais famosa) forma de consciência de estado é a consciência *fenomênica*. Esta é a propriedade que os estados mentais possuem quando há *algo como é* tê-los (a famosa frase de Nagel, 1974). Dito de outro modo, estados fenomênicamente conscientes tem *sensações* subjetivas distintivas; e alguns diriam que eles possuem *qualia*. (2000, p.13, tradução nossa, itálicos do autor).

Semelhantemente, Rosenthal (2000; 2005) faz referência a Nagel e utiliza uma variante da expressão **"**algo como é" ao falar sobre o que é relevante para caracterizar a consciência fenomênica:

> Quando alguém não tem acesso consciente a um estado, literalmente não há nada como é para esse alguém estar nesse estado. Sem acesso a um estado, uma pessoa não tem uma perspectiva de primeira pessoa de tal estado e, portanto, não há nada como é estar nesse estado. Como insiste Thomas Nagel, o que importa para a consciência é que exista algo como é "para o organismo" [NAGEL, 1979, p.166] (ROSENTHAL, 2000, p.275, tradução nossa, colchetes do autor).

Como demonstram as passagens acima, a concepção nageliana da consciência e, consequentemente, a expressão **"**algo como é" se tornaram recorrentes no vocabulário da filosofia da mente contemporânea. Apesar das diferenças internas entre as teorias sustentadas por cada um dos autores apresentados acima, todos estão dispostos a identificar ou caracterizar inicialmente o seu objeto de investigação em termos da concepção nageliana da consciência. Grosso modo, todos autores estão de acordo com a ideia de que o alvo das explicações das teorias da consciência (fenomênica) são os estados mentais que possuem **"**algo como é" tê-los.

À primeira vista, portanto, a suposição da caracterização factual do desacordo, segundo a qual ambas as abordagens assumem uma concepção comum da consciência fenomênica parece justificada. Isto é, aparentemente, ambas as abordagens concordam sobre o *explanandum* que deve ser adotado por uma teoria da consciência, a saber, estados mentais que possuem **"**algo como é" tê-los. É possível formular mais concisamente a caracterização oferecida por Nagel da seguinte maneira:

> **Concepção Nageliana**: para um estado mental M e um sujeito S, M é um estado fenomênicamente consciente, se e somente se, há algo como é para S ter/estar em M (ver FARRELL, 2017, p.3).

Na definição da concepção nageliana, apresentada acima, a noção de ter **"**algo como é" e de ser um estado fenomênicamente consciente estão fortemente vinculadas. Identificamos o

que são estados mentais fenomênicamente conscientes através da característica comum entre eles de possuir **“**algo como é” tê-los. E, conforme as passagens acima sugerem, tanto defensores de abordagens de primeira-ordem, como defensores de abordagens de segunda-ordem da consciência estão de acordo com a proposta de que ter **“**algo como é” é a característica peculiar dos estados mentais conscientes em sentido fenomênico.

Mas, a concepção nageliana não oferece, de fato, uma explicação a respeito do que é a consciência. O propósito da concepção nageliana não é elucidar *como* estados mentais adquirem a propriedade de ser fenomenicamente conscientes, nem dizer *porque* os estados mentais conscientes têm as características que lhes são atribuíveis. Seu propósito parece ser, antes, oferecer uma maneira de identificarmos quais estados mentais são conscientes, em sentido fenomênico, através de uma característica que lhes é peculiar. O que a concepção nageliana oferece é, assim, um *explanandum* sobre o qual as teorias da consciência devem concentrar seus esforços explicativos.

Tendo em vista que a concepção nageliana oferece uma maneira de descrever o fenômeno alvo de investigação para teorias da consciência, podemos inferir a partir das evidências textuais apresentadas anteriormente que as abordagens de primeira-ordem e de segunda-ordem sobre a consciência compartilham do mesmo *explanandum.* O que teorias da consciência buscam explicar, seja elas de primeira-ordem ou de segunda-ordem, são os estados que tem **“**algo como é” tê-los.

Se é o caso de que abordagens de primeira-ordem e de segunda-ordem sobre a consciência convergem em adotar a concepção nageliana como *explanandum* de uma teoria da consciência, a caracterização metalinguística perde significativamente a sua plausibilidade. Pois, poder-se-ia argumentar que, se ambas as abordagens adotam a concepção nageliana da consciência fenomênica como *explanandum*, é difícil compreender como o desacordo poderia ser caracterizado como sendo uma disputa terminológica sobre o *explanandum* de uma teoria da consciência.

Contudo, é questionável que a concepção nageliana seja unanimemente interpretada da mesma maneira. Conforme veremos a seguir, há indícios de que o ponto central desta concepção, a expressão **“**algo como é”, é ambígua entre duas interpretações possíveis, cada qual sugerindo diferentes propriedades que servem de alvos de explicação para teorias da consciência. Dada a possibilidade de que a concepção nageliana seja ambígua, há espaço para que abordagens que adotam a concepção nageliana na definição do *explanandum* de suas teorias da consciência discordem sobre o alvo de suas explicações.

3.4.2 Problemas com a concepção nageliana da consciência: “algo como é” como fonte de ambiguidade

Os exemplos citados na seção anterior demonstraram que há boas evidências de que alguns dos principais defensores das abordagens de primeira e de segunda-ordem adotam a concepção nageliana da consciência como aquilo sobre o que trata uma teoria da consciência (i.e., como *explanandum* de uma teoria da consciência). Entretanto, apenas sustentar que ambas as abordagens compartilham da concepção nageliana como *explanandum* de suas teorias não é suficiente para desbancar a plausibilidade da caracterização metalinguística do desacordo. Ainda que ambas as abordagens adotem a concepção nageliana, há razões para suspeitarmos que cada abordagem compreende diferentemente o significado dessa concepção.

A expressão **“**algo como é” é central para a caracterização do que é para um estado mental ter consciência fenomênica segundo a concepção nageliana. Nesta seção, argumentaremos que essa expressão é uma fonte de ambiguidade porque ela pode ter duas interpretações distintas, cada qual atribuindo diferentes propriedades a um estado mental. Por um lado, dizer de um estado mental que ele satisfaz a condição exigida pela concepção nageliana de haver **“**algo como é”, pode significar que um estado *mental aparenta ser de um determinado modo para um sujeito de experiência*. Sob outra interpretação, o que a sentença pode significar é que o estado mental *apresenta o mundo como aparentando ser de uma determinada maneira*. Em vista destas duas interpretações possíveis, há espaço para um desacordo implícito sobre o significado da concepção nageliana no debate entre defensores de abordagens de primeira-ordem e de segunda-ordem sobre a consciência.

3.4.2.1 Ambiguidade e condições de verdade de uma proposição

Grosso modo, termos ou expressões são ambíguas quando podem ter mais de um significado possível em uma sentença. Há diferentes fontes e tipos de ambiguidade que podem ser distinguidos, por exemplo: ambiguidade lexical, sintática e pragmática (ver, SENNET, 2021). Uma fonte de ambiguidade que nos interessará especialmente nesta seção é um subtipo da ambiguidade lexical. Na ambiguidade lexical, um termo assume dois ou mais significados. Uma das razões para que um único termo tenha dois ou mais significados é a sua capacidade de denotar diferentes tipos objetos, propriedades ou relações no mundo. Quando um termo pode se referir a diferentes tipos de objetos, propriedades ou relações no mundo e, em virtude desse

fato, assumir diferentes interpretações legítimas do seu significado, diremos que este termo é referencialmente ambíguo[39].

A ambiguidade possui consequências relevantes para a interpretação e para a avaliação da verdade e falsidade das sentenças. Quando há termos ambíguos em uma sentença, a sentença como um todo poderá expressar diferentes proposições, dependendo do significado que vinculamos ao termo. O caso de Callie* e Oscar* sobre o termo "banco" exemplifica uma situação na qual diferentes proposições são expressas por uma sentença contendo um termo referencialmente ambíguo. O termo "banco" é ambíguo entre dois tipos distintos de objetos no mundo, as instituições financeiras e os objetos feitos para sentar e descansar. Por essa razão, a sentença "Há um banco na frente da praça" pode significar tanto que há um objeto feito para sentar e descansar na frente da praça, ou que há uma instituição financeira na frente da praça. Quando Oscar* utiliza o termo "banco" na sentença, sua intenção é denotar um objeto feito para sentar e descansar. E quando Callie* usa esse mesmo termo, ela tem a intenção de denotar é uma instituição financeira. Nesse caso, quando Oscar* declara a sentença "Há um banco na frente da praça", a proposição expressa por ele é distinta daquela expressa quando Callie* declara essa sentença.

Como há duas proposições que podem ser expressas pela mesma sentença, há também duas condições sob as quais a verdade ou falsidade dessa sentença pode ser avaliada. No caso de Oscar* e a declaração de que "Há um banco na frente da praça", as condições exigidas para a verdade da declaração são a existência de um objeto feito para sentar e descansar localizado na parte da frente da praça. Já no caso de Callie*, a declaração "Há um banco na frente da praça" tem a intenção de denotar uma instituição financeira, portanto, é em virtude da presença de uma instituição financeira na frente da praça que sua declaração poderá ser considerada verdadeira (ou falsa).

Um aspecto importante a ser notado a respeito das condições de verdade sob as quais as proposições que uma sentença contendo termos referencialmente ambíguos expressa é que nada impede que as condições para a veritação das duas (ou mais) proposições expressas pela sentença ambígua possam ser simultaneamente satisfeitas. Isto é, a frente da praça pode ser de tal maneira que lá existe um objeto feito para sentar e descansar, como também uma instituição financeira. De modo mais geral, portanto, as condições de verdade das proposições expressas por sentenças que apresentam esse tipo de ambiguidade podem vir a ser simultaneamente

[39] O modo como utilizamos a noção de ambiguidade referencial é distinto de como esta noção é utilizada na linguística. Diferentemente da ambiguidade referencial utilizada no contexto da linguística, não estamos sugerindo que a expressão "algo como é"

satisfeitas pelo mundo. Assim, nada impede que ao declarar a sentença “Há um banco na frente da praça”, Oscar* e Callie* estejam dizendo algo verdadeiro, ainda que distinto em significado. Com efeito, é devido a essa possibilidade que sugerimos anteriormente que o caso de Oscar* e Callie* não constitui um caso legítimo de desacordo, já que não há uma contradição explícita na sustentação simultânea da afirmação de que “Há um banco na frente da praça” e na negação de que “Há um banco na frente da praça”, desde que “banco” signifique diferentes objetos em cada caso.

3.4.2.2 O caso do motorista distraído e a ambiguidade da expressão “algo como é”

Nesta seção sustentaremos que a expressão “algo como é”, quando utilizada em contextos de atribuição de consciência, é ambígua de maneira similar ao caso de Oscar* e Callie*. Especificamente, sustentaremos que quando defensores de abordagens de segunda-ordem utilizam a expressão “algo como é”, eles utilizam esta expressão como denotando uma propriedade distinta dos estados mentais do que aquela visada pelos defensores de abordagens de primeira-ordem. Para mostrarmos que a expressão *pode* assumir diferentes interpretações, recorreremos à discussão de Lycan (1996) sobre o *caso do motorista distraído*. Este caso permite observarmos que a sentença “Há algo como é para o motorista distraído perceber algo” pode expressar diferentes proposições em virtude das diferentes propriedades que a expressão “algo como é” pode denotar.

O caso do motorista distraído consiste no seguinte exercício imaginativo: considere um motorista que está dirigindo distraidamente enquanto pensa em seus planos para as férias de verão. Repentinamente, o motorista se dá por conta de que esteve dirigindo por um longo período sem ter consciência das suas ações ou do que se passava na estrada. Felizmente, nenhum incidente ocorrera, na verdade, o motorista desviou habilmente dos buracos na pista e parou em um semáforo quando a luz vermelha estava acesa.

O caso do motorista distraído é um experimento de pensamento a respeito de uma situação corriqueira: a realização de uma tarefa que exige habilidades de discriminação perceptual e coordenação adequada do comportamento, mas que é realizada, aparentemente, sem que o sujeito esteja consciente da discriminação dos estímulos ou da coordenação do seu comportamento realizada a partir de tais discriminações.

Do ponto de vista da concepção nageliana da consciência, este é um caso intrigante que pode evidenciar duas possíveis interpretações da expressão **“**algo como é”[40] (LYCAN, 1996, p. 76). Por hipótese, o motorista distraído não está consciente de suas percepções. Portanto, se assumimos como ponto de partida a concepção nageliana da consciência, não poderíamos afirmar que **“**Há algo como é para o motorista distraído perceber a luz do semáforo”. Com efeito, se considerarmos literalmente a proposta de que o motorista não está consciente de seus próprios estados mentais, o caso do motorista distraído deverá ser interpretado como um caso no qual o princípio da transitividade não é satisfeito. Portanto, trata-se de uma situação na qual defensores de uma abordagem representacionalista de segunda-ordem não estariam dispostos atribuir a propriedade da consciência fenomênica aos estados mentais do motorista. Segundo defensores dessa abordagem, durante o lapso atencional do motorista, as suas percepções devem ser todas consideradas como ocorrendo de maneira inconsciente, mas elas se tornam conscientes (ou poderiam se tornar conscientes) a partir do momento em que há um estado mental que representa tais percepções (CARRUTHERS, 2019, p.90).

Contudo, há um sentido no qual poderíamos estar dispostos a afirmar que **“**Há algo como é para o motorista distraído perceber a luz do semáforo” sem abandonarmos a concepção nageliana. Eis uma maneira de argumentar a favor desta interpretação alternativa: se motorista parou em um sinal vermelho, presumivelmente ele percebeu a luz do semáforo como sendo vermelha, antes do que verde. A percepção de que a luz era vermelha antes do que verde é a única razão pela qual ele poderia ter parado no semáforo. Portanto, em certo sentido, podemos dizer que há algo como é para o motorista distraído perceber a luz do semáforo, ainda que tal fenomenologia não tenha sido explicitamente vivenciada pelo sujeito, devido ao seu lapso atencional (ver, LYCAN,1997, p. 76). Com efeito, Tye (2017, p.14-15) parece assumir exatamente esta interpretação a respeito do caso do motorista distraído. Segundo sua interpretação do caso, as representações perceptuais do motorista distraído são inerentemente conscientes, na medida em que se tratam de experiências do motorista. Porém, o que falta às experiências do motorista distraído é um tipo de reconhecimento de segunda-ordem sobre a ocorrência dessas experiências: **“**As experiências visuais do motorista distraído são inconscientes no sentido de que ele não está consciente *de* suas experiências. Ele também não está consciente *de que* elas estão ocorrendo” (TYE, 2019, p.14, tradução nossa, itálicos do autor). Nesse sentido, Tye sugere realizarmos uma distinção entre dois sentidos que a noção de consciência pode assumir. Um estado mental pode ser consciente no sentido de que esse estado

[40] No original, “*what it’s like”*.

mental é uma experiência, o que simplesmente consiste em um estado mental que possui uma determinada fenomenologia, quer ela seja reconhecida ou não pelo sujeito. Em um segundo sentido, a noção de consciência é utilizada para denotar casos nos quais um sujeito possui uma experiência que é alvo representacional de um pensamento consciente (e.g., quando refletimos, introspectivamente, sobre nossas experiências ocorrentes). Esse segundo sentido é nomeado por Tye como um tipo de consciência de segunda-ordem. Além de conceitualmente distinta do primeiro sentido de consciência, que é um tipo de processo psicológico de primeira-ordem, a consciência de segunda-ordem não é essencial para a ocorrência da consciência de primeira-ordem: **"**experiências são estados conscientes de primeira-ordem sobre os quais estados conscientes de segunda-ordem podem ou não estar dirigidos" (TYE, 2017, p. 15).

Constatamos, assim, uma situação na qual a sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo" pode assumir diferentes interpretações, particularmente porque a expressão **"**algo como é" contida na sentença pode expressar diferentes significados. Por um lado, o caso supõe que o motorista distraído não é consciente. Portanto, que ele não satisfaz a condição exigida pela concepção nageliana de **"**ter algo como é" para ele ter seus estados mentais ocorrentes. Sendo assim, a sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo" deve ser considerada falsa, quando avaliada sob o contexto do experimento de pensamento. Por outro lado, um segundo olhar para o caso mostra-nos que há um sentido em que poderíamos estar dispostos a dizer que **"**Há algo como é para o motorista distraído perceber a luz do semáforo". Pois, conforme a descrição do caso sugere, o motorista realizou habilmente tarefas que exigem o exercício de capacidades de discriminação perceptual. Portanto, o motorista foi capaz de perceber determinados objetos em seu percurso *como sendo de tal-e-tal maneira*. Em outras palavras, tais objetos percebidos pelo motorista se pareciam de alguma maneira para ele, o que, sob uma interpretação alternativa, satisfaria a condição exigida pela concepção nageliana para a atribuição de consciência.

Diante da possibilidade de assumirmos posicionamentos divergentes sobre como se aplica a concepção nageliana da consciência ao caso do motorista distraído, Lycan (1996, p.77) sustenta o diagnóstico de que a expressão **"**algo como é" é ambígua entre duas interpretações possíveis. Segundo o diagnóstico de Lycan, essa expressão pode se referir a dois tipos distintos de propriedades[41], o que ocasiona a possibilidade de diferentes posicionamentos frente a sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo".

[41] Lycan (1996, p.76) identifica estas propriedades como sendo, por um lado, os chamados *qualia* e, por outro lado a propriedade de um estado mental de ser introspectivamente percebido por um sujeito.

Eis uma primeira aproximação de como entender os dois tipos de propriedades que a expressão algo como é pode denotar, e que implicam em duas diferentes condições de verdade para a sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo". Poderíamos dizer que a sentença é verdadeira, se compreendermos que **"**algo como é" se refere a propriedade de um estado mental de representar o mundo como sendo de determinada maneira qualitativa. Sob essa interpretação, o caso do motorista distraído consistiria em ter uma percepção que representa a luz do semáforo como tendo determinadas propriedades qualitativas (i.e., a cor vermelha), mas cujo conteúdo dessa representação não é (necessariamente) um conteúdo de que o motorista está consciente durante o seu lapso atencional. De tal modo, dizer que **"**Há algo como é para o motorista distraído perceber a luz do semáforo" consiste apenas em afirmar que o estado perceptual visual do motorista distraído é tal que ele representa a luz do semáforo como tendo determinadas propriedades qualitativas (i.e., o vermelho fenomenal) em virtude do qual a luz do semáforo se parece vermelha – ou alternativamente, se pareceria vermelha, caso o motorista tivesse seus recursos atencionais propriamente direcionados.

Por outro lado, a sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo" pode ser avaliada como falsa, se compreendermos que a propriedade a qual a expressão **"**ter algo como é" se refere é a propriedade de um estado mental de estar subjetivamente disponível para um sujeito. Sob essa interpretação, a disponibilidade subjetiva de um estado mental pode ser compreendida simplesmente como a capacidade de um sujeito de reconhecer, não inferencialmente, a ocorrência desse estado mental. Segundo a descrição do caso, o motorista distraído não reconhece que percebeu a luz do semáforo como vermelha (pelo menos não até o momento posterior à sua percepção). Portanto, durante a realização da discriminação perceptual que desencadeou sua ação de parar no semáforo, a percepção de que a luz do semáforo estava vermelha não era algo subjetivamente disponível para o motorista[42]. Nesta segunda interpretação da sentença **"**Há algo como é para o motorista distraído perceber a luz do semáforo", afirmá-la consistiria em dizer que o estado perceptual visual do motorista distraído tem a propriedade de ser subjetivamente disponível para o motorista, o que é falso no caso em questão.

Se o diagnóstico de Lycan está correto, a expressão **"**ter algo como é" exibe o mesmo tipo de ambiguidade que o termo **"**banco" exibe no exemplo de Oscar* e Callie* e, plausivelmente, compartilha das mesmas características listadas acima. A sentença afirmativa **"**Há algo como é para o motorista distraído perceber a luz do semáforo" pode ter diferentes

[42] Veremos mais detalhadamente no que consiste a disponibilidade subjetiva de um estado mental na seção seguinte.

valores semânticos por razões similares que a sentença **"**Há um banco na frente da praça" pode obter diferentes valores semânticos quando proferidas por Oscar* ou por Callie*. Similarmente ao termo **"**banco" que pode denotar diferentes objetos no mundo, a expressão **"**Há algo como é" pode denotar diferentes propriedades dos estados mentais: a propriedade de apresentar o mundo como sendo de determinada maneira e a propriedade de estar subjetivamente disponível para um organismo. Sendo assim, a mesma sentença admite duas interpretações plausíveis, que expressam diferentes proposições e exigem ser avaliadas sob diferentes condições de verdade.

As condições que um estado mental deve atender para que se possa atribuir uma das duas propriedades distinguidas não são as mesmas. Não há razões para crermos que os processos psicológicos em virtude dos quais um sujeito pode ser capaz de reconhecer, não inferencialmente, que um estado mental M está ocorrendo são os mesmos que aqueles em virtude dos quais o estado M representa o ambiente como tendo determinadas propriedades qualitativas. Com efeito, casos empíricos de estados mentais que representam determinadas parcelas do mundo, como a percepção visual de pacientes que sofrem de *visão-cega*, servem como evidência de que a capacidade de um sujeito de reconhecer a ocorrência de um estado mental não está necessariamente correlacionada a capacidades dos estados mentais do sujeito de representar o mundo como tendo determinadas propriedades[43]. Sendo assim, a expressão **"**algo como é" quando utilizada em casos de atribuição de consciência pode ser interpretada e avaliada sob diferentes condições de verdade, a depender da propriedade que o falante pretende que essa expressão denote na sentença.

3.4.3 Dois sentidos para "algo como é": subjetividade mundana e subjetividade experiencial

Através da discussão do caso do motorista distraído vimos que a concepção nageliana pode assumir diferentes significados de acordo com a propriedade dos estados mentais que interpretamos ser denotada pela expressão **"**algo como é". Nesta seção, aprofundaremos nossa discussão sobre como são caracterizadas e distinguidas as propriedades de um estado mental de representar o mundo como sendo de determinada maneira e a propriedade de ser subjetivamente disponível para um sujeito. Para tanto, nos valemos de uma distinção conceitual de Carruthers

[43] Nos casos de visão-cega, pacientes com lesões na área V1 do lobo occipital relatam ser incapazes de enxergar estímulos em determinadas áreas do seu campo visual. Não obstante, esses mesmos pacientes demonstram um bom desempenho em tarefas discriminação de propriedades visuais, nas quais os pacientes são forçados a escolher entre algumas opções (e.g., escolha entre a orientação de uma linha como vertical ou horizontal) (CARRUTHERS, 2000, p.155).

(2000), um dos principais defensores do representacionalismo de segunda-ordem. Um importante objetivo dessa subseção consiste em reforçar que essas propriedades, nomeadas respectivamente de *subjetividade mundana* e *subjetividade experiencial*, possuem condições de instanciação distintas. Se essas propriedades têm diferentes condições de instanciação, há consequências sobre a compreensão e uso da concepção nageliana. Pois, ao atribuirmos consciência fenomênica a um estado mental, as condições que exigimos que tal estado cumpra variam de acordo com o tipo de subjetividade que se considera ser relevante para a concepção nageliana da consciência.

Carruthers (2000) utiliza o par de conceitos *subjetividade mundana* e *subjetividade experiencial*[44] para caracterizar e diferenciar as duas propriedades que a expressão **"**algo como é" pode se referir (LURZ, 2002). Um dos principais objetivos na introdução desta distinção conceitual é oferecer um esclarecimento do *explanandum* de sua teoria da consciência, a qual se concentra em explicar como estados mentais vem a ter *subjetividade experiencial*[45]. Nesta seção, destacaremos dois aspectos desta distinção conceitual: como Carruthers caracteriza cada um dos tipos de subjetividade e quais condições são exigidas para que possamos atribuir subjetividade mundana ou experiencial a um estado mental.

A distinção conceitual entre a subjetividade mundana e subjetividade experiencial reflete uma diferença entre propriedades de estados mentais (2000, p. 128). Para o autor, há uma diferença conceitual e real entre dois tipos de subjetividade que podem ser atribuídas aos estados mentais de um organismo. A *subjetividade mundana* diz respeito a como o mundo se apresenta para um organismo através dos seus estados perceptuais e sensoriais. A *subjetividade experiencial* diz respeito a como é para o organismo experimentar seus próprios estados mentais (CARRUTHERS, 2000, p. 128).

A propriedade designada pelo nome de subjetividade mundana se origina na relação entre o mundo externo e um organismo percipiente, ela é uma propriedade de um estado mental em virtude da qual *um organismo representa o mundo como tendo determinadas qualidades*[46].

---

[44] No original: *worldly subjectivity* e *experiential subjectivity*, respectivamente.

[45] Um segundo objetivo associado à distinção conceitual entre a subjetividade mundana e a subjetividade experiencial consiste em oferecer uma base conceitual para a formulação de objeções às abordagens de primeira-ordem. Especificamente, através desses conceitos, Carruthers promove argumentos contrários às abordagens (de primeira-ordem) que sustentam uma compreensão distinta do que é a consciência fenomênica e como se deveria explicá-la. Estes objetivos estão interligados: ao especificar o *explanandum* da sua teoria da consciência como distinto do *explanandum* de teorias de primeira-ordem, Carruthers busca argumentar que a sua teoria é explanatoriamente superior às rivais de primeira-ordem, porque a noção de consciência fenomênica adotada pelas principais abordagens representacionalistas de primeira-ordem é inadequada para oferecer uma distinção clara entre estados mentais conscientes e inconscientes (ver, CARRUTHERS 2000, p.168-179).

[46] Da maneira como a noção de subjetividade mundana é caracterizada aqui, seguindo a caracterização de Carruthers, atribuir subjetividade mundana a um estado mental carrega consigo um compromisso com a

Uma maneira de evidenciar o que é a *subjetividade mundana*, consiste em compararmos a maneira como indivíduos de diferentes espécies atribuem, via seus estados perceptuais, propriedades qualitativas aos seus ambientes. Considere, por exemplo, a diferença entre humanos e gatos domésticos com relação a percepção visual. Diferentes informações perceptuais estão disponíveis para indivíduos de cada espécie formar suas respectivas representações visuais do ambiente. Os gatos possuem um sistema visual dicromático, o que os impede de discriminar determinados tipos de cores no ambiente. Particularmente, gatos parecem ser incapazes de discriminar alguns tons de vermelho e verde, que são percebidos por humanos com visão normal[47] (CLARK; CLARK, 2016). Nesse sentido, as propriedades cromáticas que a percepção visual dos gatos é capaz de atribuir ao ambiente são distintas das propriedades cromáticas que a visão humana é capaz de atribuir ao ambiente, de modo que uma mesma cena seria visualmente representada de maneira distintas por indivíduos de cada espécie.

As diferenças em como indivíduos de cada espécie representam visualmente uma mesma cena se devem às diferentes constituições fisiológicas de seus aparatos sensoriais e às capacidades discriminatórias que os indivíduos de cada espécie possuem[48]. Assim, no sentido em que Carruthers utiliza a noção de subjetividade mundana, as diferenças representacionais destacadas acima são diferenças em como os estados mentais dos indivíduos de cada espécie desfrutam da propriedade da subjetividade em sentido *mundano*. Resumidamente, a subjetividade mundana é a maneira como o mundo (ou o corpo do próprio sujeito) é representado por um estado mental do sujeito (ver, CARRUTHERS, 2000, p. 147).

Por seu turno, a propriedade designada pela noção de *subjetividade experiencial* é originária da relação que um organismo tem com os seus próprios estados mentais sensoriais e perceptuais. Um estado mental dotado de subjetividade experiencial é um estado que *se parece de alguma maneira para o organismo* (CARRUTHERS, 2000, p. 128). Para Carruthers (2005,

---

abordagem segundo a qual o conteúdo representacional de um estado mental é responsável pelo seu caráter fenomenal. Contudo, esse compromisso pode ser dispensado através de uma caracterização mais neutra da subjetividade mundana, segundo a qual um estado mental M tem subjetividade mundana se M atribui, via um processo P, propriedades qualitativas ao ambiente do organismo. Sob essa caracterização mais neutra, P não é necessariamente um processo representacional, abrindo espaço para a compatibilização com teorias *fenomenalistas*, como em Block (2003).

[47] A visão dos gatos é, nesse sentido, similar à de humanos com deuteranopia, um tipo de daltonismo que afeta a discriminação de tons verdes (CLARK; CLARK, 2016).

[48] Horácio de Sá Pereira (2014, p. 64) identifica como *internismo fenomenal* a tese de que o caráter fenomenal da experiência, o que se aproxima do que chamamos aqui de "ponto de vista", sobrevém localmente do extrato biológico do indivíduo. A caracterização realizada acima da subjetividade mundana é compatível com a adoção do internismo fenomenal, na medida em que a constituição biológica de cada indivíduo (ou espécie a qual pertence o indivíduo) é considerada relevante para as diferentes propriedades que podem ser representadas na experiência como pertencentes ao mundo. Em outras palavras, as distinções físicas entre as espécies são relevantes para diferenças nas propriedades que são representadas na experiência como pertencentes ao ambiente.

p. 43), a subjetividade experiencial está relacionada a capacidade dos organismos de possuir informações sobre seus próprios estados mentais perceptuais (e/ou sensoriais), de maneira não inferencial, e de ser capaz de fazer discriminações entre os estados mentais com base nessas informações.

A subjetividade experiencial de um estado mental não adiciona nenhuma qualidade sensorial ao estado mental, como é o caso da subjetividade mundana (ver, CARRUTHERS, 2000, p. 212; 237). Mas, segundo a proposta de Carruthers, é apenas em virtude de um estado mental ter subjetividade experiencial que ele se parece de uma determinada maneira para um sujeito. Isto é, na ausência de subjetividade experiencial, as qualidades sensoriais que um estado mental atribui para uma determinada parcela do mundo (ou do corpo do sujeito) não são experimentadas pelo sujeito.

A subjetividade experiencial, na medida em que concerne a disponibilidade ou o acesso que um sujeito tem dos seus estados mentais ocorrentes, está estreitamente vinculada ao princípio da transitividade. O princípio da transitividade, conforme vimos anteriormente, sustenta que estados mentais conscientes são estados de que um sujeito está apropriadamente consciente. Similarmente, um estado mental que desfruta de subjetividade experiencial é um estado que se parece de alguma maneira para o organismo. Sendo assim, é possível sustentar que estados mentais que desfrutam da subjetividade experiencial são estados que, de alguma maneira apropriada, implementam o princípio da transitividade.

Com efeito, Carruthers (semelhantemente a outros defensores do representacionalismo de segunda-ordem, ver ROSENTHAL, 2005; KRIEGEL, 2009) sustenta que não há outra maneira de um organismo ter as capacidades relacionadas à subjetividade experiencial, senão através da implementação de representações de segunda-ordem. Carruthers (2000, p.185; 188) considera a propriedade de um estado mental de **“**se parecer de alguma maneira para um organismo” como uma propriedade metafisicamente constituída pelo adequado funcionamento de uma faculdade de representações de segunda-ordem, que viabiliza informações sobre os estados mentais ocorrentes para um módulo de teoria da mente, entre outros módulos cognitivos.

O papel desempenhado pela faculdade de representações de segunda-ordem é o de alterar o perfil funcional de um estado mental de primeira-ordem. Considere um estado mental de primeira-ordem M que representa uma propriedade qualitativa Q (e.g., vermelho). Na medida em que o estado mental M se torna disponível para uma faculdade de pensamento de segunda-ordem, ele adquire um conteúdo intencional de segunda-ordem **“**experiência ocorrente

de Q”[49]. De tal modo, além de veicular a informação de que uma parcela do mundo tem a propriedade qualitativa Q, quando M é alvo de uma representação de segunda-ordem, ele veicula a informação de que uma sensação/percepção de Q está ocorrendo (ver, CARRUTHERS, 2005, p. 135; 2000, p. 242). Nas palavras do autor:

> Cada experiência do mundo-corpo se torna, ao mesmo tempo, uma representação de que tal experiência está ocorrendo; cada experiência com um conteúdo, digamos, *vermelho*$_a$ é, ao mesmo tempo, um evento com o conteúdo [...] *experiência de vermelho*$_a$ (CARRUTHERS, 2000, p.242, tradução nossa, itálicos do autor)[50].

No que diz respeito a teoria representacionalista de primeira-ordem articulada por Tye (1995), o reconhecimento, por parte de um organismo, de que uma experiência está ocorrendo também é explicada em termos de processos de segunda-ordem. Tye (1995, p. 115), ao discutir o que permite que alguns organismos reconheçam a ocorrência dos seus próprios estados mentais adota uma versão da teoria do pensamento de segunda-ordem. Segundo a sua proposta, o que permite que sejamos capazes de reconhecer que determinadas experiências estão ocorrendo é a presença de pensamentos de segunda-ordem capazes de aplicar conceitos fenomênicos ao conteúdo de nossos estados experienciais. Os conceitos em questão são ou indexicais (i.e., conceitos capazes de apontar ou **“**destacar” aspectos de uma experiência ocorrente) ou conceitos predicativos (i.e., conceitos por meio dos quais somos capazes de discriminar uma experiência de outras com base no seu caráter fenomênico) (TYE, 1995, p.167-169). Assim, a situação do motorista distraído é explicada por Tye (1995, p.115) como um caso no qual o motorista é incapaz de reconhecer que tem determinadas experiências ocorrentes porque, devido ao seu lapso atencional, ele não as conceitualiza sob os conceitos fenomênicos necessários.

Há dois pontos cruciais a serem notados a respeito da distinção conceitual entre a subjetividade mundana e a experiencial. O primeiro ponto é que, na caracterização de Carruthers, a instanciação da subjetividade mundana é independente da instanciação da subjetividade experiencial (mas, é disputável se o contrário também é metafisicamente

---

[49] Na abordagem de Carruthers (2000) a aquisição do conteúdo de segunda-ordem se deve à hipótese adicional assumida pelo autor de alguma versão de uma teoria do conteúdo mental baseada em sistemas de consumo (e.g, MILIKAN, 1984). Segundo estas teorias do conteúdo mental, o conteúdo de um estado mental M é determinado pelas funções realizadas pelos sistemas que utilizam M como *input*. Nesse sentido, M adquire um novo tipo de conteúdo ao ser disponibilizado para faculdades de representações de segunda-ordem responsáveis por produzir juízos (ou percepções) de que o organismo está atualmente em um determinado estado perceptual/sensorial (ver, CARRUTHERS, 2000, p.244-246).

[50] O subscrito em “*vermelho*$_a$” e “*experiência de vermelho*$_a$” é um recurso utilizado por Carruthers para indicar que o conteúdo em questão é analógico ou não conceitual, em contraste com conteúdos digitais ou conceituais. Para mais detalhes sobre as distinções analógico-digital e não conceitual e conceitual, ver CARRUTHERS, 2000, p. 129-133.

possível)[51]. Quando um estado tem apenas subjetividade mundana, um sujeito representa o mundo como tendo determinadas propriedades qualitativas, porém, o conteúdo de tal representação pode não *se parece de alguma maneira para o sujeito*. Isto ocorre porque o sujeito não é capaz de ter informações, obtidas de maneira não inferencial, sobre a ocorrência dos seus estados mentais, ou porque os estados mentais em questão não estão, por alguma razão, disponíveis para a faculdade de representação de segunda-ordem.

A independência das condições de instanciação da subjetividade mundana relativamente as condições de instanciação da subjetividade experiencial permite que essas noções sejam exploradas na articulação de uma explicação sobre o que distingue estados mentais de que um sujeito tem consciência e os estados mentais de que um sujeito não tem consciência. Isto nos leva ao segundo ponto crucial a respeito desta distinção conceitual: o papel que as noções de subjetividade mundana e subjetividade experiencial podem desempenhar na articulação de uma teoria da consciência.

Em consonância com outras abordagens representacionalistas de segunda-ordem (ver, ROSENTHAL, 2005; LYCAN, 1996; KRIEGEL, 2009), Carruthers (2000, p. 129, n.7) assume que um estado mental que possui apenas *subjetividade mundana* não é um estado consciente (em sentido fenomênico). Pois, há diversos exemplos de estados mentais que representam o ambiente como sendo de determinada maneira (i.e., possuem subjetividade mundana), entre os quais os casos de visão-cega, mas que não são intuitivamente considerados como estados mentais conscientes de um sujeito. Não sendo a *subjetividade mundana* a propriedade que caracteriza estados mentais conscientes, Carruthers sustenta que o que há em comum entre todos os estados mentais conscientes (em sentido fenomênico) é a *instanciação conjunta da subjetividade mundana* e *experiencial*[52]. Por esta razão, Carruthers considera que um dos

[51] Segundo a compreensão de Carruthers (2000) dessas propriedades, a subjetividade mundana é uma parte constitutiva da subjetividade experiencial. Portanto, a subjetividade experiencial não poderia ocorrer na ausência da subjetividade mundana. Outros autores como Rosenthal (2005) discordam desse posicionamento, sustentado que a subjetividade experiencial não é constituída pela subjetividade mundana. De tal modo, seriam possíveis representações de segunda-ordem que veiculam, erroneamente, a ocorrência de um estado perceptual ou sensorial, quando na verdade nenhum estado de primeira-ordem está ocorrendo.

[52] Este posicionamento não é unanimemente admitido entre representacionalistas de segunda-ordem. Rosenthal (ROSENTHAL, 2002; ROSENTHAL; LAU, 2011) nega que a subjetividade mundana seja essencial para a atribuição de consciência fenomênica para um estado mental. Não obstante, sua versão do representacionalismo de segunda-ordem sobre a consciência endossa a tese de que a propriedade relevante para que um estado mental tenha consciência fenomênica é a *subjetividade experiencial* (chamada de fenomenalidade espessa em ROSENTHAL, 2002). Negar que a subjetividade mundana seja essencial para estados mentais conscientes implica em admitir a existência de estados de segunda-ordem "vazios". Isto é, estados mentais de segunda-ordem que representam estados mentais de primeira-ordem que não existem ou que são, por alguma razão, desprovidos de conteúdo representacional. Tais casos dão origem a chamada objeção da representação radical. Ver, GOTTLIEB, 2020; BLOCK, 2011; FARRELL, 2017; WEISBERG, 2011.

principais objetivos de uma teoria da consciência (fenomênica) é oferecer uma explicação do que é, e como é realizada, a *subjetividade experiencial* em estados mentais.

Uma consequência deste posicionamento da abordagem representacionalista de segunda-ordem para a questão da ambiguidade da concepção nageliana da consciência é a interpretação de que a subjetividade experiencial é (ou deveria ser) a propriedade denotada pela expressão "ter algo como é". Nesse sentido, uma chave de leitura plausível para analisar as divergências entre abordagens de primeira-ordem e segunda-ordem da consciência a respeito da atribuição de consciência fenomênica em casos problemáticos consiste em compreender que as diferentes abordagens assumem que o conceito denota a instanciação de distintas propriedades de estados mentais. Ou seja, quando abordagens de primeira-ordem sustentam que um estado mental M é consciente, o que eles sustentam é que M instancia a propriedade da subjetividade mundana, mas não necessariamente a subjetividade experiencial. E quando abordagens de segunda-ordem sustentam que M é consciente, é a instanciação conjunta da subjetividade mundana e experiencial que é atribuída ao estado mental M.

3.5 A CONDIÇÃO DA NORMATIVIDADE DO SIGNIFICADO NO DEBATE ENTRE ABORDAGENS DE PRIMEIRA-ORDEM E SEGUNDA-ORDEM DA CONSCIÊNCIA

Sustentamos anteriormente que o debate entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência fenomênica satisfaz três das quatro condições exigidas para que um debate seja considerado uma negociação metalinguística.

O debate satisfaz a (1) condição da disputa, na medida em que se apresenta no formato de uma troca linguística na qual um desacordo parece ser expresso. Uma importante evidência neste sentido é o fato de que as abordagens de primeira-ordem e segunda-ordem são, frequentemente, apresentadas como abordagens rivais sobre como explicar a propriedade dos estados mentais de serem fenomênicamente conscientes. Por seu turno, sugerimos que a condição (2) do desacordo é satisfeita porque é possível interpretar que as divergências no nível dos fatos (metafísicos) a respeito da consciência entre as abordagens pode ser produto de uma divergência mais fundamental sobre o significado e uso do conceito de consciência fenomênica. Ou, alternativamente, que as divergências sobre os fatos (metafísicos) sobre a consciência que cada abordagem assume podem implicar em divergências sobre o significado e uso do conceito de consciência fenomênica.

A justificação para a interpretação metalinguística do desacordo está vinculada às evidências de que o debate satisfaz a (3) condição da equivocação. Na seção 4, argumentamos que o debate satisfaz a condição da equivocação porque há evidências de que, apesar de ambas as abordagens assumirem a concepção nageliana como *explanandum*, tal concepção é ambígua entre duas possíveis compreensões. E, crucialmente, cada uma das abordagens sustenta diferentes compreensões sobre o que a concepção nageliana exige de um estado para que ele seja fenomênicamente consciente. Para abordagens de segunda-ordem, a propriedade com a qual a consciência fenomênica é identificada na concepção nageliana é a *subjetividade experiencial*. Ou seja, a propriedade de um estado mental de estar subjetivamente disponível para um sujeito. Esta propriedade é, por seu turno, explicada pelo representacionalismo de segunda-ordem em termos de mecanismos psicológicos metarepresentacionais.

Já as abordagens de primeira-ordem, que negam que a disponibilidade de um estado mental para o sujeito seja uma condição necessária para que estados mentais sejam conscientes, sustentam que é a propriedade da *subjetividade mundana* que deve ser identificada com a propriedade relevante para que estados mentais tenham consciência fenomênica. A subjetividade mundana consiste na propriedade de um estado mental de apresentar o mundo como tendo determinadas qualidades (propriedades que não são necessariamente disponíveis para os processos cognitivos do sujeito). Nesse sentido, explicar essa propriedade dispensa o recurso a processos metarepresentacionais.

Nesta seção, exploramos as evidências que corroboram a tese de que o debate satisfaz a condição (4) da normatividade do significado. Um debate satisfaz essa condição se há evidências de que o debate envolve questões práticas sobre como um conceito deve ser usado em um determinado contexto e/ou questões avaliativas a respeito das virtudes ou defeitos das diversas opções disponíveis para o uso e compreensão de um conceito para um determinado contexto. Nesse sentido, argumentamos que o uso da concepção nageliana por parte das abordagens de primeira-ordem e de segunda-ordem consiste em uma tentativa de advogar, implicitamente, como deveríamos utilizar e compreender o significado do conceito de consciência fenomênica.

3.5.1 Argumento a partir da ambiguidade à semântica normativa do conceito de consciência

Um argumento a favor da tese de que o debate entre abordagens de primeira-ordem e segunda-ordem da consciência fenomênica envolve questões sobre semântica normativa pode

ser articulado a partir da ambiguidade apresentada pela concepção nageliana. Notamos na seção 4.1 que a concepção nageliana da consciência não oferece uma explicação sobre a consciência fenomênica, mas uma maneira de caracterizar os estados mentais conscientes através da identificação de uma característica comum entre eles, a saber, ter "algo como é" tê-los. Contudo, na seção 4.2 vimos razões para crer que essa concepção é ambígua, em particular que ter "algo como é" pode assumir diferentes interpretações que apontam para diferentes propriedades de estados mentais.

Ora, constatar que há uma ambiguidade na expressão que deveria desempenhar o papel de caracterizar o objeto de investigação de teorias da implica em um problema às teorias que adotam essa concepção como *explanandum*. Se admitirmos que a concepção nageliana é referencialmente ambígua, no sentido que utilizamos o termo na seção 4.2.1, isto implica em reconhecermos que essa expressão, de fato, não está cumprindo adequadamente o seu papel de identificar o objeto de investigação das teorias da consciência.

Diante desta situação uma teoria poderia adotar algum dos seguintes posicionamentos: (1) Negar a realidade desta ambiguidade ou ignorá-la enquanto algo que não constitui um problema para a adoção da concepção nageliana. (2) Admitir que há uma ambiguidade na expressão "algo como é" e buscar por uma maneira de solucioná-la sem abandonar a concepção nageliana. Ou (3) eliminar a expressão "algo como é" e, por conseguinte, a concepção nageliana em virtude do reconhecimento de tal ambiguidade.

Negar que, de fato, há uma ambiguidade na expressão "algo como é" ou optar por ignorá-la nos parece uma resposta inviável ao problema. O fato de que há uma literatura dedicada a esclarecer o significado da expressão "algo como é" e que oferece sugestões divergentes sobre o seu significado é prova suficiente de que a expressão é uma fonte de alguma ambiguidade (ver, FARRELL, 2016; SNOWDON, 2010; BYRNE, 2004). E, a ubiquidade do uso da expressão "algo como é" em caracterizações, declarações e formulação de questões sobre a consciência faz com que a ambiguidade da expressão "algo como é" seja um problema difícil de ser ignorado. Sem uma compreensão mínima do significado que essa expressão assume em cada caso, não será claro o significado que as caracterizações, declarações e questões sobre a consciência fenomênica expressam. Consequentemente, não será possível decidir, com segurança, se tais declarações e caracterizações são verdadeiras, já que não estaríamos em condições de identificar adequadamente o que faria delas verdadeiras ou falsas (ver, FARRELL, 2016, p.52).

Uma vez reconhecida a ambiguidade contida na concepção nageliana, cabe decidir se é desejável ou não preservar esta concepção. Uma razão que pode contar a favor da preservação

desta concepção é, mais uma vez, a sua ubiquidade nas discussões sobre a consciência. Se simplesmente optarmos por eliminar as referências à concepção nageliana como parte de um discurso sem sentido ou trivial (e.g, HACKER, 2002; SNOWDON, 2010), deveríamos igualmente considerar como sem sentido ou trivial a ampla literatura que adota essa expressão como uma maneira de se referir à consciência fenomênica? Deveríamos ignorar argumentos influentes como o quarto de Mary (JACKSON, 1982) e o argumento dos zumbis de Chalmers (1997), que são articulados em termos da concepção nageliana? Dada a repercussão desses argumentos em diversos campos da filosofia da mente, fazer pouco caso dessas discussões não parece uma opção viável. Se, ao contrário, adotarmos o posicionamento que se incumbe da tarefa de tornar mais claro o significado que a expressão "algo como é" assume em diferentes contextos de discussão sobre a consciência, podemos tentar resguardar as discussões articuladas em termos da concepção nageliana como peças informativas no debate sobre a consciência.

Mas, se uma teoria da consciência deseja manter o seu uso, é preciso que ela ofereça uma desambiguação do que significa para um estado mental ter "algo como é" no sentido que é relevante para a atribuição de consciência fenomênica. Uma parte central desta tarefa consiste em escolher qual dos possíveis conceitos ou expressões candidatos devemos favorecer na desambiguação da expressão original. Escolher qual expressão ou conceito entre um conjunto de candidatos deve ser adotado para a realização de uma desambiguação da concepção nageliana é uma tarefa que pode ser caracterizada como parte de um projeto em ética conceitual. Segundo Plunkett, a ética conceitual se caracteriza, grosso modo, pela preocupação com problemas normativos e avaliativos sobre conceitos, pensamentos, discurso e representações. Entre tais problemas, constam questões sobre quais conceitos deveríamos usar, quais defeitos conceitos podem exibir e o que deveríamos querer dizer através de nossas palavras (CAPPELEN; PLUNKET, 2020, p.4).

Vimos anteriormente, na seção 4.3 que há, no mínimo, duas intepretações alternativas sobre o que denota a expressão "algo como é" na concepção nageliana da consciência. E vimos também que abordagens de primeira-ordem e de segunda-ordem sustentam diferentes interpretações. Sendo assim, há bases para sustentarmos que, individualmente, cada abordagem está vinculada à tarefa de elucidar como deveríamos interpretar a concepção nageliana. Isto é, na medida em que as abordagens de primeira-ordem e segunda-ordem se engajam em um projeto no domínio da ética conceitual, temos uma boa motivação para interpretar o debate como estando envolvido em um problema de semântica normativa. Ou seja, com a questão sobre o que deveríamos querer dizer através do uso de um determinado termo ou qual deveria ser o significado que ele expressa.

3.5.2 Uma objeção a partir da ambiguidade da concepção nageliana da consciência

A ambiguidade da noção nageliana pode motivar também uma objeção à tese de que o debate satisfaz a condição da normatividade do significado. Esta objeção se vale da nossa comparação anterior entre o debate entre as abordagens de primeira-ordem e segunda-ordem da consciência e o caso de Oscar* e Callie*.

Notamos anteriormente, na seção 2, que não há um desacordo genuíno entre Oscar* e Callie* porque suas proposições tratam sobre diferentes objetos na realidade. Consequentemente, não há uma incompatibilidade entre o conteúdo proposicional sustentado por ambos. Isto é, a proposição que Oscar* sustenta a respeito da existência de um banco na frente da praça não implica em contradições ou outros tipos de incompatibilidade com a proposição que Callie* sustenta a respeito da não existência de um banco na frente da praça. Nesse sentido, algo similar poderia ser sustentado com relação ao debate a respeito da consciência fenomênica entre abordagens de primeira-ordem e segunda-ordem.

Poder-se-ia argumentar que não há incompatibilidade entre as (proposições sustentadas pelas) abordagens porque, semelhantemente ao caso de Oscar* e Callie*, as proposições de cada abordagem são sobre propriedades distintas. Portanto, quando abordagens de primeira-ordem negam que o princípio da transitividade – segundo o qual um estado perceptual ser *fenomênicamente* consciente consiste na consciência que um sujeito tem de si mesmo como estando nesse estado (de alguma maneira apropriada) – não é uma condição necessária para que um estado mental tenha consciência fenomênica, o que é negado não é o mesmo que as abordagens de segunda-ordem sustentam. Pois, abordagens de segunda-ordem sustentam que o princípio da transitividade é uma condição necessária para que um estado tenha subjetividade experiencial, e o que abordagens de primeira-ordem negam é que o princípio da transitividade seja uma condição necessária para que estados mentais tenham subjetividade mundana. Este argumento oferece um obstáculo para a tese de que o debate pode ser interpretado como sendo uma negociação metalinguística, porque nos leva a reconsiderar se o debate satisfaz a condição do desacordo[53].

[53] Em apoio indireto a este argumento, poder-se-ia sustentar ainda que não só a subjetividade mundana é uma propriedade distinta da subjetividade experiencial, mas a primeira é uma parte constitutiva da segunda. Em razão disso, as abordagens cujo objeto de investigação é a subjetividade experiencial têm de ser, no mínimo, compatíveis com alguma versão de uma teoria a respeito da subjetividade mundana.

Mas, a objeção levantada acima ignora duas diferenças entre o caso de Oscar* e Callie* e o debate sobre a concepção nageliana da consciência fenomênica entre as abordagens de primeira-ordem e segunda-ordem: a convencionalidade lexical dos termos em cada caso e a consequência para o debate de se notar que o desacordo é terminológico.

Primeiro, enquanto o termo **"**banco" possui significados bem estabelecidos no léxico da língua portuguesa, os significados que a expressão "algo como é" na concepção nageliana da consciência fenomênica podem assumir não são tão bem estabelecidos. Uma possível explicação para este fato é a tese de que a expressão "algo como é", quando utilizada no contexto da concepção nageliana, não têm um significado que possa ser extraído da maneira que ordinariamente extraímos o significado de partes de uma sentença.

A hipótese padrão que assumimos ao encarar qualquer sentença é a de que o significado da sentença e de partes significativas dessa sentença podem ser extraídas a partir da consideração do significado das palavras que a compõem em conjunto com a sintaxe aparente da sentença. Nesse sentido, ao nos depararmos com a expressão "algo como é", deveríamos extrair o significado dessa expressão considerando a combinação do significado individual tal como definidos no léxico da língua portuguesa e a função sintática que os termos aparentam desempenhar na sentença.

Porém, é fácil encontrar exemplos que demonstram que a expressão "algo como é" e suas variações não podem ter o seu significado extraído da maneira ordinária, quando utilizadas no contexto da filosofia da mente contemporânea (FARRELL, 2016, p.52; SNOWDON, 2010, p. 17-21). Ou seja, a expressão não tem seu significado meramente determinado pela combinação dos significados dos termos individuais que a compõem e a função sintática que os termos desempenham na sentença.

Farrell (2016) demonstra este ponto de maneira sucinta ao discutir o uso da expressão "algo como é" por parte de Chalmers em seu argumento dos zumbis. A noção de um zumbi no sentido utilizado por Chalmers (1997) é a de um organismo fisicamente idêntico ao ser humano comum, mas que não tem experiências com qualidades subjetivas. Uma maneira através da qual Chalmers caracteriza sua noção de um zumbi é a seguinte:

> [...] ele estará acordado, é capaz de relatar o conteúdo dos seus estados internos, é capaz de focar sua atenção a vários lugares, e assim por diante. A única diferença é que nenhuma dessas funções estará acompanhada por experiências conscientes reais. Não haverá nenhuma sensação fenomênica. Não há nada como é ser um zumbi (CHALMERS, 1997, p. 95, tradução nossa).

Caso o significado da sentença "Não há nada como é ser um zumbi" fosse ordinário, ele seria o resultado da combinação literal dos significados dos termos individuais que compõem a

sentença e da sua ordenação. Portanto, a parte que diz “Não há nada como é ser um zumbi” significaria, aproximadamente, que não existem seres que se assemelham ao modo como são os zumbis (no sentido técnico especificado acima). Mas este não é o sentido que a porção relevante “como é” assume na discussão de Chalmers. A leitura segundo a qual o “como é” diz respeito às similaridades entre duas coisas é completamente equivocada no contexto da discussão de Chalmers sobre zumbis. Pois, o que faz dos zumbis um experimento mental filosoficamente relevante para a discussão da consciência fenomênica é o fato de que eles se parecem com seres humanos em todos os sentidos, exceto na sua vida mental. Portanto, seria equivocado assumir que a variante da expressão utilizada por Chalmers tem um significado extraído através do procedimento ordinário.

Se admitimos a tese de que expressão “algo como é” não possui uma interpretação que pode ser extraída da maneira ordinária, então temos de reconhecer que não é possível determinar o seu significado simplesmente recorrendo ao léxico convencional da língua portuguesa. Aparentemente, nós sustentamos, é preciso que a interpretação do significado dessa expressão esteja inserida no contexto de uma teoria da consciência, a qual estipula de que modo a expressão deveria ser interpretada. Nesse sentido, diferentemente do caso entre Oscar* e Callie*, pode haver uma disputa legítima entre os conteúdos literalmente compatíveis expressos pelas abordagens de primeira-ordem e segunda-ordem da consciência na medida em que eles se concentram numa disputa sobre qual teoria da consciência e, em conjunto com ela, qual interpretação da concepção nageliana da consciência que deveríamos adotar.

A segunda diferença entre o caso de Oscar* e Callie* e o debate entre abordagens de primeira-ordem e segunda-ordem que é ignorado pela objeção é o fato de que, uma vez que Oscar* e Callie* reconhecem que o uso do termo **“**banco” não está sendo utilizado com o mesmo significado, a aparência de que há uma disputa é derrotada. Porém, o mesmo não se aplica ao caso do debate sobre a concepção nageliana da consciência.

Oscar* e Callie discordam sobre se há um banco na frente da praça. Porém, esse desacordo é meramente aparente e se desfaz uma vez que ambos reconheçam que o uso do termo tem a intenção de denotar diferentes objetos quando utilizado nas sentenças de cada um deles. A razão pela qual o aparente desacordo se desfaz é o fato de que há uma convenção bem estabelecida sobre quando é legítimo afirmar que há um banco na frente da praça para ambos os sentidos que o termo pode assumir. Mas, no caso do debate sobre a concepção nageliana da consciência entre as abordagens de primeira-ordem e segunda-ordem, o debate não se resolve uma vez que seja apontado o modo como a concepção nageliana é canonicamente utilizada ou como essa noção foi originalmente articulada por Nagel (1974). A razão pela qual isso não

encerra o debate se deve ao fato de que, o que está implicitamente sendo debatido pelas abordagens não é apenas qual é o significado atual da concepção nageliana, mas sobre como deveríamos compreender essa noção e utilizá-la enquanto *explanandum* de uma teoria da consciência. Nesse sentido, ainda que houvesse uma convenção bem estabelecida sobre como essa noção fora compreendida no passado, ainda é possível dar continuidade ao debate se levantarmos a questão sobre como deveríamos interpretar a concepção nageliana ao articular uma teoria da consciência.

3.6 CONCLUSÃO E DESDOBRAMENTOS

Neste capítulo argumentamos que o debate entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência fenomênica pode ser interpretado como uma negociação metalinguística. Em suporte a essa hipótese, apresentamos diversas evidências de que o debate satisfaz as quatro condições (elencadas na seção 1) que, se satisfeitas, caracterizam um debate como interpretável como sendo uma negociação metalinguística. No curso desse argumento, pudemos observar alguns aspectos interessantes sobre o que está em jogo na teorização sobre a consciência fenomênica. No que concerne a argumentação sobre a condição do desacordo, vimos que é possível interpretarmos que as abordagens de primeira-ordem e segunda-ordem discordam não apenas sobre as condições factuais que devem ser realizadas para que um estado mental seja consciente, mas também sobre como representar o fenômeno que é alvo das explicações fornecidas por uma teoria da consciência. Em outras palavras, vimos que o desacordo no nível dos fatos entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência pode ser motivado por desacordo conceituais. Vimos na sequência, ao discutir como o debate satisfaz a condição da equivocação, que os conceitos utilizados por cada abordagem para representar o fenômeno alvo das teorias denota diferentes propriedades de estados mentais –a subjetividade mundana e a subjetividade experiencial – às quais possuem diferentes condições de instanciação.

Em virtude das teorias da consciência de primeira-ordem e de segunda-ordem adotarem diferentes propriedades como alvos de suas explicações, consideramos na seção 5 a possibilidade de que o debate entre as abordagens seja meramente verbal. Contudo, vimos que há boas razões para sustentar que o debate pode ser interpretado como dizendo respeito à semântica normativa, uma vez que a concepção nageliana da consciência, muito frequentemente adotada como *explanandum* de uma teoria sobre a consciência fenomênica, é ambígua. Portanto, ao articular teorias a respeito da consciência fenomênica, podemos

interpretar que defensores de abordagens de primeira-ordem e de segunda-ordem sobre a consciência estão engajados, implicitamente, na tarefa de resolver a ambiguidade presente nessa concepção. A tarefa de elucidar o que deveríamos compreender pelo conceito de consciência fenomênica, que figura como explanandum em teorias da consciência é imprescindível para o desenvolvimento da investigação teórica e, principalmente, empírica sobre esse fenômeno. Apenas na medida em que estamos de posse de um conceito que represente adequadamente o fenômeno sob investigação podemos formular as questões, argumentos e experimentos adequados para investigar teórica e empiricamente a consciência fenomênica.

Por fim, gostaríamos de considerar brevemente um desdobramento da hipótese de que o debate entre as abordagens de primeira-ordem e segunda-ordem sobre a consciência pode ser interpretado como uma negociação linguística. Se, como sugere a hipótese desse capítulo, as abordagens de primeira-ordem e segunda-ordem sobre a consciência adotam diferentes propriedades como seu *explanandum*, e se tais propriedades não são mutuamente exclusivas, então existe a possibilidade de sustentar que às abordagens de primeira-ordem e segunda-ordem da consciência podem desempenhar papéis complementares na explicação sobre a consciência fenomênica. Isto é, podemos defender uma posição "conciliatória" entre as duas abordagens. A ideia básica de uma proposta conciliatória entre a abordagem de segunda-ordem e primeira-ordem é, em nossa nomenclatura, que ambas as propriedades, a subjetividade mundana e a experiencial, desempenham um papel relevante na consciência fenomênica dos estados mentais.

Brown e Lau (2019, p. 185-191; ver também Kriegel, 2009) articulam um posicionamento conciliatório que se caracteriza por sustentar que tanto os aspectos dos estados mentais de primeira-ordem, como os aspectos do estado mental de segunda-ordem contribuem no modo como se parecem, para nós, as nossas experiências. Vejamos brevemente como esse posicionamento pode ser elucidado à luz de nossa discussão prévia sobre como as abordagens de primeira-ordem e de segunda-ordem sobre a consciência caracterizam seus respectivos *explananda*.

Sustentamos anteriormente que abordagens de primeira-ordem adotam a subjetividade mundana como principal *explananda* de uma teoria da consciência. Ou seja, o principal objetivo de uma teoria da consciência, segundo defensores de uma abordagem de primeira-ordem, é explicar a propriedade em virtude da qual estados mentais são capazes de representar qualitativamente o mundo e/ou o corpo do sujeito. Por sua vez, as abordagens de segunda-ordem adotam como *explananda* a subjetividade mundana, o que implica em sustentar que o principal objetivo de uma teoria da consciência é, para defensores dessa abordagem, explicar

como os sujeitos tem acesso subjetivo aos seus próprios estados mentais. E notamos anteriormente que a subjetividade experiencial é parcialmente constituída pela subjetividade mundana, pois ter um determinado conteúdo representacional disponível ou subjetivamente acessível é algo que exige, de antemão, a existência de um conteúdo representacional de primeira-ordem do qual se tem acesso. Nesse sentido, para que uma representação de segunda-ordem produza conteúdos como "experiência de Q" ou "estou tendo uma percepção vívida de Q", em que Q é uma propriedade qualitativa, a representação de segunda-ordem deve se *referir ao conteúdo particular* do estado de primeira-ordem relevante que tem o conteúdo Q (e.g., vermelho). Segundo um posicionamento conciliatório entre abordagens de primeira-ordem e segunda-ordem sobre a consciência, portanto, o modo como se parece, para nós, um estado mental fenomênicamente consciente é determinado pela combinação de aspectos da representação de segunda-ordem e do conteúdo do estado de primeira-ordem.

Brown e Lau (2019, p. 187) propõe a seguinte maneira através da qual essa determinação conjunta de como se parecem os estados mentais fenomênicamente conscientes pode ser compreendida:

> A intensidade da experiência consciente é determinada pela representação de segunda-ordem (i. e., percepção vívida, em contraste com ter uma impressão fraca ou insegura), mas, em última análise, o conteúdo detalhado (e.g., de uma cor) é determinado pelo estado de primeira-ordem, em virtude de ser referido pelo estado de segunda-ordem.

Essa maneira de compreender a interação entre estados de primeira-ordem e segunda-ordem, em conjunto com a interpretação metalinguística do debate entre as teorias da consciência de primeira-ordem e segunda-ordem, permite preservar intuições de ambas as abordagens. Com relação às abordagens de segunda-ordem, esse posicionamento preserva a intuição de que os estados de segunda-ordem são relevantes para o modo como se parecem nossas experiências. Assim, por exemplo, o conteúdo representacional de um estado de primeira-ordem pode contribuir para o aspecto fenomenológico de uma experiência apenas na medida em que ele é representado por um estado de segunda-ordem para o efeito de que o sujeito está tendo o estado de primeira-ordem em questão. Além disso, o estado de segunda-ordem, na formulação de Brown e Lau (2019) dessa perspectiva, pode desempenhar um papel de modular essa experiência, por exemplo, em termos da vivacidade ou confiabilidade que é atribuída ao conteúdo do estado de primeira-ordem.

Com relação as abordagens de primeira-ordem, o posicionamento conciliatório esboçado aqui permite preservar a intuição de que as qualidades que são apresentadas em uma experiência são determinadas pelo conteúdo representacional do estado de primeira-ordem.

Nesse sentido, se o conteúdo representacional dos estados de primeira-ordem é, de alguma maneira, empobrecido ou ausente, haverá uma diferença na maneira como se parecem, para os sujeitos, as suas experiências, seja ela comportamental ou fenomenológica.

Um caso que ilustra as consequências da perspectiva conciliatória mencionadas acima é a percepção de cores na visão periférica. Nós comumente pensamos que enxergamos cores e outros detalhes dos objetos na periferia de nosso campo visual e parece haver uma experiência fenomenológica desses detalhes na periferia de nossa visão. Porém, há evidências de que nós não somos realmente capazes de discriminar diversos detalhes na periferia de nosso campo visual (ODEGAARD et al., 2018). Isto é, a fenomenologia e nossos juízos introspectivos sobre essas experiências da periferia de nosso campo visual são falsas.

A perspectiva conciliatória articulada acima pode ajudar a esboçar uma hipótese sobre porque temos esses juízos e falsa fenomenologia. A fenomenologia e os juízos introspectivos sobre essa fenomenologia podem ser o resultado de termos estados mentais de segunda-ordem que superestimam a qualidade ou confiabilidade dos estímulos representados pelo estado de primeira-ordem, resultando na impressão subjetiva de que estamos conscientes de qualidades particulares. Porém, na medida em que o conteúdo representacional do estado de primeira-ordem é, de fato, empobrecido ou ausente, não somos capazes de realizar identificações ou descriminações acuradas dos detalhes dos estímulos representados na periferia do nosso campo visual (e.g., maior taxa de repostas falso-positivas para a detecção de um estímulo na periferia, ver ODERGAARD et al., 2018).

A perspectiva conciliatória, porém, é vulnerável a diversas críticas (ver, BROWN; LAU, 2019, p.189-190). Defensores da abordagem de primeira-ordem podem criticar essa perspectiva conciliatória porque ela concede ênfase demasiada ao papel explanatório desempenhado pelos estados de segunda-ordem. Segundo essa crítica, na medida em que a perspectiva conciliatória assume que estados mentais de segunda-ordem são relevantes para a consciência fenomênica dos estados mentais, ela deve ser considerada como uma versão do representacionalismo de segunda-ordem e, portanto, vulnerável às mesmas objeções que as demais versões dessa abordagem (e.g., objeções a veracidade ou relevância do princípio da transitividade). Por outro lado, defensores de versões mais conservadoras do representacionalismo de segunda-ordem podem criticar a perspectiva conciliatória da seguinte maneira: se é a existência de um estado mental de segunda-ordem que faz referência ao conteúdo de um estado de primeira-ordem que explica porque porque um estado mental é fenomênicamente consciente, então não é claro qual papel o desempenhado pelo conteúdo do estado de primeira-ordem. Não poderia o estado de segunda-ordem, ele mesmo, representar esse conteúdo diretamente? E não poderia o estado de

segunda-ordem representar erroneamente um conteúdo de primeira-ordem e, ainda assim, dar origem a um estado mental consciente?

Apesar das críticas que podem ser levantadas contra a proposta de uma perspectiva que concilia as abordagens de primeira-ordem e segunda-ordem sobre a consciência, cremos que a perspectiva conciliatória esboçada aqui pode ser frutiferamente explorada na articulação de uma teoria sobre a consciência fenomênica que é simultaneamente mais robusta que ambas as abordagens de primeira-ordem e de segunda-ordem sobre a consciência, na medida em que herda diversas de suas virtudes teóricas e, potencialmente, é imune às objeções enfrentadas pelas abordagens originais. Por exemplo, a perspectiva conciliatória é imune à objeção enfrentada pela abordagem de segunda-ordem de que podem haver casos nos quais uma representação de segunda-ordem nos faz consciente de um estado mental de primeira-ordem inexiste. Pois, na perspectiva conciliatória a existência de um estado de primeira-ordem é uma condição constitutiva da representação de segunda-ordem em virtude da qual um sujeito é consciente de um dos seus estados mentais (ver, CARRUTHERS, 2005, p. 65-66). Pela mesma razão, a perspectiva conciliatória é também imune à crítica de que há uma duplicação de conteúdos representacionais por parte dos estados de segunda-ordem (BROWN; LAU, 2019, p.190). Por seu turno, a crítica enfrentada pela abordagem de primeira-ordem de que há evidências que sugerem que a atividade nas áreas corticais tipicamente vinculadas à processos cognitivos superiores (e a metacognição em particular) são relevantes para o modo com experienciamos nossos estados mentais não afeta a perspectiva conciliatória aqui esboçada, pois ela assume que tais áreas corticais são relevantes para o modo como experienciamos o conteúdo dos nossos estados mentais de primeira-ordem.

Tendo essas considerações em vista, em conjunto com as diversas discussões que foram promovidas ao longo deste capítulo sobre como compreender a relação entre teorias de primeira-ordem e de segunda-ordem sobre a consciência, concluímos que a interpretação metalinguística do debate entre abordagens de primeira-ordem e de segunda-ordem sobre a consciência é, no mínimo, uma chave de leitura frutífera para a investigação filosófica e empírica da consciência.

CONCLUSÃO

Nesta dissertação exploramos, sob diversas perspectivas, os conceitos de consciência e sua relevância na teorização a respeito desse fenômeno mental. No capítulo 1 pudemos observar que o termo "consciente" é polissêmico, podendo expressar diferentes conceitos que se distinguem segundo a aridade do predicado e o seu escopo de aplicação. A distinção e dos conceitos de consciência são uma ferramenta útil na teorização sobre a consciência, permitindo uma discussão mais sensível aos diferentes tipos de eventos mentais que se reúnem sobre o rótulo mais amplo de eventos mentais conscientes.

No capítulo 2 vimos como o representacionalismo de segunda-ordem se vale da variedade de conceitos de consciência para fornecer uma hipótese sobre o que faz com que estados mentais de primeira-ordem sejam *intransitivamente conscientes* a partir da ideia de que somos, via estados mentais metarepresentacionais, *transitivamente conscientes* desses estados. Ao longo de nossa apresentação do representacionalismo de segunda-ordem, vimos que o chamado *princípio da transitividade*, segundo o qual um estado mental é consciente apenas se um sujeito é, adequadamente, consciente desse estado, é uma intuição pré-teórica apoiada por diversas motivações, tais como: (a) ser uma concepção representativa de uma noção de senso-comum da consciência; (b) ser compatível com nossas evidências fenomenológicas sobre como é ter estados mentais conscientes; e (c) ser uma noção de consciência metodologicamente frutífera para a investigação teórica e empírica da consciência. Além disso, a estratégia explicativa derivada do princípio da transitividade, segundo a qual a consciência de estados mentais é realizada por processos metarepresentacionais exibe duas virtudes teóricas, na medida em que é capaz de gerar predições testáveis (virtude diacrônica da fecundidade), as quais são compatíveis com as evidências empíricas disponíveis (virtude da acurácia evidencial).

Por fim, no capítulo 3 vimos como pode ser frutífero interpretar o debate entre a abordagem de primeira-ordem e a abordagem de segunda-ordem como sendo uma negociação metalinguística sobre como devemos entender o conceito "consciência fenomênica", que serve de *explanandum* numa teoria da consciência. Em particular, nossa discussão sugeriu que estas duas abordagens identificam diferentes propriedades dos estados mentais como a propriedade que merece o rótulo de "consciência fenomênica" e que exige ser explicada por uma teoria da consciência fenomênica. Mas, vimos também, que na medida em que não há uma incompatibilidade, no nível semântico, entre as proposições sustentadas por cada abordagem, é possível articular uma perspectiva conciliatória entre elas. Segundo esta perspectiva conciliatória, as abordagens de primeira-ordem e de segunda-ordem sobre a consciência

desempenham papéis complementares na elucidação das condições que, conjuntamente, são suficientes para a realização de estados mentais fenomênicamente conscientes. E, uma vez assumida a perspectiva conciliatória, podemos preservar intuições e virtudes de ambas as abordagens, ao mesmo tempo em que evitamos alguns dos seus defeitos.

REFERÊNCIAS

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

BAARS, B. Global Workspace Theory of Consciousness: Toward a Cognitive Neuroscience of Human Experience. **Progress in Brain Research**, v. 150, n. 1, p.45-53, 2005. DOI: 10.1016/S0079-6123(05)50004-9.

BALCOMB, F; GERKEN, L. Does Implicit Metacognition Provide a Tool for Self-Guided Learning in Preschool Children?. **Proceedings of the Annual Meeting of the Cognitive Science Society**, v. 28, n. 28, p. 1003-1008, 2006.

BALCOMB, F; GERKEN, L. Three-year-old children can access their own memory to guide responses on a visual matching task. **Developmental Science**, v. 11, p. 750–760, 2008.

BAYNE, T. Conscious States and Conscious Creatures: Explanation in The Scientific Study of Consciousness. **Philosopical Perspectives**. V. 21, p. 1-22, 2007. DOI: 10.1111/j.1520-8583.2007.00118.x

BELLERI, D. Ontological Disputes and the Phenomenon of Metalinguistic Negotiation: Charting the Territory. **Philosophy Compass**, v. 15, n. 7, 2020. DOI: https://doi.org/10.1111/phc3.12684.

BERMÚDEZ, J. **Philosophy of Psychology**: A Contemporary Introduction. New York: Routledge, 2005.

BLOCK, N. Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience. **Behavioral and Brain Sciences**, N. 30, v.5, p. 481-548, 2007.

BLOCK, N. Mental paint. In: HAHN, M; RAMBERG, B. **Reflections and Replies: Essays on the Philosophy of Tyler Burge**. MIT Press. 2003, p. 165-200.

BLOCK, N. On a confusion about a function of consciousness. **Behavioral and Brain Sciences.** V. 18, n.2, p. 227-247. Jun, 1995. DOI: 10.1017/S0140525X00038188.

BLOCK, N. The Higher-order Approach to Consciousness is Defunct. **Analysis**, v. 71, n. 3, p. 419-431, 2011.

BRADDON-MITCHELL, D; JACKSON, F. Philosophy of Mind and Cognition. 2ed. Oxford: Blackwell, 2006.

BRENTANO, F. **Psychology from an Empirical Standpoint**. New York: Routledge, 2015.

BROWN, R; LAU, H; LeDOUX. Understanding the Higher-order Approach to Consciousness. **Trends in Cognitive Science**s**,** v. 23, n. 9, p.754-767, 2019.

BYRNE, A. Some Like it Hot: Consciousness and Higher-Order Thoughts. **Philosophical Studies**, v. 86, n. 2, p. 103-129, 1997.

CARRRUTHERS, P. Higher-Order Theories of Consciousness. In: ZALTA, E. (ed.) The Stanford Encyclopedia of Philosophy, 2016. Disponível em: <https://plato.stanford.edu/archives/fall2016/entries/consciousness-higher/ > Acesso em: 20, Maio, 2020.

CARRUTHERS, P. **Consciousness**: Essays from a Higher-Order Perspective. New York: Oxford University Press, 2005.

CARRUTHERS, P. **Human and Animal Minds:** The Consciousness Questions Laid to Rest. Oxford: Oxford University Press, 2019.

CARRUTHERS, P. **Phenomenal Consciousness**: A Naturalistic Theory. Cambridge: Cambridge University Press, 2000.

CARRY, M; REDER, L. Metacognition in Strategy Selection. In: CHAMBRES, P; IZAUTE, M; MARESCAUX, P. **Metacognition**. Boston: Springer, 2002.

CHALMERS, D. **The conscious mind**: In search of a fundamental theory. New York: Oxford University Press, 1997.

CHALMERS, D. Verbal Disputes. **Philosophical Review,** v. 120, n. 4, p. 515-566, 2011. DOI: 10.1215/00318108-1334478.

CHALMERS, D. What is a Neural Correlate of Consciousness? In METZINGER, T. (ed.). **The Neural Correlates of Consciousness**. Cambridge: MIT Press, 2000.

CLARK, D; CLARK, R. Neutral Point Testing of Color Vision in the Domestic Cat. **Exp Eye Res**. v. 153, n.1, p.23-26, 2016. DOI: 10.1016/j.exer.2016.10.002.

de SÁ PEREIRA, R. O Caráter Fenomenal como Modo de Processamento de Informação. In: SMITH, P; FILHO, W; SANTOS, P. **Crença, Verdade, Racionalidade: Ensaios de Filosofia Analítica.** Salvador: Edufba, 2014.

DEHAENE, S; CHANGEUX, J. Experimental and Theoretical Approaches to Conscious Processing. **Neuron**, v. 70, n. 2, p. 200-227, 2011.

DEHAENE, S; NACCACHE, L; COHEN, L; BIHAN, D; MANGIN, J; POLINE, J; RIVIÈRE, D. Cerebral mechanisms of word masking and unconscious repetition priming. **Nat. Neurosci**. v. 4, n. 1, p. 752–758, 2001.

DESMURGET, M. Movement Intention After Parietal Cortex Stimulation in Humans. **Science**, v. 324, n. 811, p.811-813, 2009. DOI: 10.1126/science.1169896.

DIENES, Z; PERNER, J. The Metacognitive Implications of the Implicit-Explicit Distinction. In: CHAMBRES, P; IZAUTE, M; MARESCAUX, P. **Metacognition**. Boston: Springer, 2002.

DIJKSTERHUIS, A**.** Think Different: The Merits of Unconscious Thought in Preference Development and Decision Making. **Journal of Personality and Social Psychology**, v. 87, n. 5, p. 586–598, 2004. doi:10.1037/0022-3514.87.5.586.

DRETSKE, F. **Naturalizing the Mind**. Cambridge: MIT Press, 1995.

DRETSKE, F. Perception without Awareness. In: GENDLER, T; HAWTHORNE, J. **Perceptual Experience**, Oxford: Oxford University Press, 2006. DOI: 10.1093/acprof:oso/9780199289769.003.0005.

FARELL, J. Higher-Order Theories of Consciousness and What-it-is-like-ness. **Philosophical Studies**, v. 175, n. 11, p. 2743-2761, 2017. DOI: 10.1007/s11098-017-0980-8.

FLEMING, S. **Know Thyself**: The Science of Self-awareness. New York: Basic Books, 2021.

FODOR, J. **The Modularity of Mind**: An Essay on Faculty Psychology. Cambridge: MIT Press, 1983.

GENNARO, R. Higher-order Theories of Consciousness: An Overview. In: Gennaro, R. **Higher-order Theories of Consciousness: An anthology.** Philadelphia: John Benjamin.

GENNARO, R. Representational Theories of Consciousness. In: Gennaro, R. **The Routledge Handbook of Consciousness**. New York: Routledge, 2018.

GOLDMAN, A. Consciousness, Folk Psychology and Cognitive Science. **Consciousness and Cognition**, v. 2, p. 364-382, 1993.

GOTTLIEB, J. On Ambitious Higher-Order Theories of Consciousness. **Philosophical Psychology**. v. 33, n. 3, p. 421-441, 2020. DOI: 10.1080/09515089.2020.1731445.

GOTTLIEB, J. Transitivity Principle and Transparency. **Analytic Philosophy**, v. 57, n. 4, p. 353- 379, 2016. DOI: 10.1111/phib.12091.

GOTTLIEB, J. Verbal Disputes in The Theory of Consciousness**. Ergo: An Open Access Journal in Philosophy**, V. 5, 2018. DOI: http://dx.doi.org/10.3998/ergo.12405314.0005.012.

GUZËLDERE, G. Is consciousness the perception of what passes in one's own mind? In: METZINGER, T. **Conscious Experience.** Padernborn: Ferdinand Schoningh, 1995.

HACKER, P. Is There Anything it is Like to be a Bat?. **Philosophy**, v. 77, n. 300, p. 157-174, 2002. DOI: 10.1017/s0031819102000220.

HAGGARD, P; EIMER, M. On the Relation Between Brain Potentials and the Awareness of Voluntary Movements. **Exp**. **Brain**. **Res**. v. 126, n. 1, p. 128–133, 1999. DOI: https://doi.org/10.1007/s002210050722

HASSON, U; SKIPPER, J; NUSBAUM, H; SMALL, S. Abstract coding of audiovisual speech: Beyond sensory representation. **Neuron**, v. 56, n. 1, p. 1116–1126, 2007.

HILL, C. **Consciência**. Trad. Alzira Allegro. São Paulo: Editora Unesp, 2011.

JACKSON, F. Epiphenomenal qualia. **Philosophical Quarterly**, v. 32, n. 127, 1982.

KEAS, M. Systematizing the Theoretical Virtues. **Synthese**, v. 190, n. 9, 2017.

KIM, J. (2010). Philosophy of Mind. 3ed. Boulder, Colorado: Westview.

KIRK, R. **Raw Feeling**: A Philosophical Account of the Essence of Consciousness. Oxford: Oxford University Press, 1994.

KRIEGEL, U. PANIC Theory and the Prospects for a Representational Theory of Phenomenal Consciousness. **Philosophical Psychology**, v. 15, n.1, p.55-64, 2002. DOI: 10.1080/09515080120109414.

KRIEGEL, U. **Subjective Consciousness**: A Self-Representational Theory. New York: Oxford University Press, 2009.

LAU, H. Consciousness, Metacognition & Perceptual Reality Monitoring. 2019. DOI: https://doi.org/10.31234/osf.io/ckbyf.

LAU, H; BROWN, R. The emperor's new phenomenology? The empirical case for conscious experience without first-order representations. In: PAUTZ, A; STOLJAR, D. **Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness**. New York: MIT Press, 2019.

LAU, H; PASSINGHAM, R. Relative blindsight in Normal Observers and the Neural Correlate of Visual Consciousness. **Proc. Natl. Acad. Sci. U. S. A.,** v. 103, n. 49, p. 18763–18768, 2006.

LAU, H; ROSENTHAL, D. Empirical support for higher-order Theories of conscious Awareness. **Trends in Cognitive Sciences, v.** 15, n. 8, p. 365–373, 2011.

LEVINE, J. **Purple Haze**: The Puzzle of Consciousness. New York: Oxford University Press, 2001.

LEWIS, D. Psychophysical and Theoretical Identifications. **Australasian Journal of Philosophy**. v. 50, n. 3, 1972. DOI: https://doi.org/10.1080/00048407212341301

LIBET, B. Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action. **Behavioral and Brain Sciences**, v.8, n. 4, p.529-566.
DOI: https://doi.org/10.1017/S0140525X00044903

LURZ, R. Reducing Consciousness by Making it Hot A Review of Peter Carruthers' Phenomenal Consciousness. **Psyche**, v. 8, n. 5, 2002.

LYCAN, W. A Simple Argument for a Higher-order Representation Theory of Consciousness. **Analysis**, v.61, n.1, 2001.

LYCAN, W. **Consciousness and Experience**. Cambridge, MA: MIT Press, 1997.

MANSON, N. What Does Language Tell Us About Consciousness? First-person Mental Discourse and Higher-order Thought Theories of Consciousness. **Philosophical Psychology**, v. 15, n.3, p. 221-238, 2002. DOI: 10.1080/0951508021000006076.

MARTIN, M. The Limits of Self-Awareness. **Philosophical Studies**, v. 120, p. 37-89, 2004. DOI:10.1023/b:phil.0000033751.66949.97.

MERKER, B. Consciousness without a cerebral cortex: a challenge for neuroscience and medicine. **Behav Brain Sci**, V. 30, n.1, p. 81-134, fev, 2007. DOI: 10.1017/S0140525X07000891.

MORALES, J; LAU, H; FLEMING, S. Domain-General and Domain-Specific Patterns of Activity Supporting Metacognition in Human Prefrontal Cortex. Journal of Neuroscience, V. 38, n. 14, p. 3534-3546, 2018. DOI: https://doi.org/10.1523/JNEUROSCI.2360-17.2018.

NAGEL, T. What is it like to be a bat? **The Philosophical Review**, V. 83, n. 4, p. 435-450, 1974. Disponível em: <http://www.jstor.org/stable/2183914>. Acesso em: 19, jun, 2019.

ODEGAARD, B. CHANG, M; HAKWAN, L; CHEUNG, S. Inflation Versus Filling-In: Why We Feel We See More Than We Actually Do In Peripheral Vision. **Phil. Trans. R. Soc.**, v. 373, n. 1755, 2018. DOI: https://doi.org/10.1098/rstb.2017.0345.

PICCININI, G; CRAVER, C. Integrating Psychology and Neuroscience: functional analyses as mechanism sketches. **Synthese**, v. 183, p. 283-311, 2011. DOI: 10.1007/s11229-011-9898-4.

PLUNKETT, D; SUNDELL, T. Disagreement and the Semantics of Normative and Evaluative Terms. **Philosopher's Imprint**, v. 13, n. 23, 2013.

PLUNKETT, D; SUNDELL, T. Metalinguistic negotiation and speaker error. **Inquiry**, v. 64, n.1, p. 1-26, 2019. DOI: 10.1080/0020174X.2019.1610055.

PLUNKETT. D. Which Concepts Should We Use?: Metalinguistic Negotiations and The Methodology of Philosophy. **Inquiry: An Interdisciplinary Journal of Philosophy**, v. 58, n. 7-8, p. 828-874, 2015. DOI: 10.1080/0020174X.2015.1080184.

PLUNKETT; D; CAPPELEN; A. Introduction: A Guided Tour of Conceptual Engineering and Conceptual Ethics. In: Burgess, A; CAPPELEN, H; PLUNKETT, D. **Conceptual Engineering and Conceptual Ethics**. Oxford: Oxford University Press, 2020.

PRINZ, J. **The Conscious Brain**. Oxford: Oxford University Press, 2012.

PROUST, J. Metacognition and Metarepresentation: Is a Self-directed Theory of Mind a Precondition for Metacognition? **Synthese**, v. 159, n. 2, p. 271–195, 2007.

PYLYSHYN, Z. Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception. **Behavioral and Brain Sciences**, V. 22, n. 3. p. 341–365, 1999. DOI:10.1017/S0140525X99002022.

RO, T; SINGHAL, N; BREITMEYER, B; GARCIA, J. Unconscious processing of color and form in metacontrast masking. **Atten. Percept. Psychol.** V. 71, p. 95–103, 2009. DOI: 10.3758/APP.71.1.95.

ROSENTHAL, D. Consciousness and its Function. **Neuropsychologia**, v. 46, n. 1, p. 829-840, 2008. DOI: 10.1016/j.neuropsychologia.2007.11.012

ROSENTHAL, D. **Consciousness and Mind**. [S.l.]: Oxford University Press, 2005.

ROSENTHAL, D. Higher-order awareness, misrepresentation and function. **Philosophical transactions of the Royal Society of London. Series B, Biological sciences**, v. 367, n. 1594 p. 1424-1438, 2012. DOI:10.1098/rstb.2011.0353.

ROSENTHAL, D. Higher-order Thoughts and the Appendage Theory of Consciousness. **Philosophical Psychology,** V. 6, n. 2, p. 155-166, 1993. DOI: 10.1080/09515089308573085.

ROSENTHAL, D. Rosenthal, David. Consciousness, Content, and Metacognitive Judgments. **Consciousness and cognition**, v. 9, p. 203-214, 2000. DOI: 10.1006/ccog.2000.0437.

ROSENTHAL, D. Thinking that One Thinks. In: DAVIES, M; HUMPHREYS, G. **Consciousness**. Oxford: Blackwell, 1993.

ROSENTHAL, D. Two Concepts of Consciousness. **Philosophical Studies**, V. 49, n. 3, p. 329-359, 1986.

ROSENTHAL, D; WEISBERG, J. Higher-Order Theories of Consciousness. **Scholarpedia**, v.3, n.5, 2008. Disponível em: < http://www.scholarpedia.org/article/Higher-order_theories_of_consciousness>. Acesso em: 27, Maio, 2022.

ROUNIS, E. et al. Theta-burst Transcranial Magnetic Stimulation to the Prefrontal Cortex Impairs Metacognitive Visual Awareness. **Cogn. Neurosci**. v, 1, p. 165-175, 2010.

SANTOS, C. O Debate Causalismo Versus Simulacionismo em Filosofia da Memória como Negociação Metalinguística. **Perspectiva Filosófica**, v. 42, n. 2, p. 143-188, 2019.

SCOTT, S. Metarepresentation in Philososphy and Psychology. **Proceedings of the Annual Meeting of the Cognitive Science Society**, V. 23, 2001. Disponível em: < https://escholarship.org/uc/item/1kr8079g >. Acesso em, 27, Maio, 2022.

SENNET, A. Ambiguity. In: ZALTA, E. **The Stanford Encyclopedia of Philosophy**. Disponível em: <https://plato.stanford.edu/archives/fall2021/entries/ambiguity/>. Acesso em: 02 maio 2022.

SNOWDON, P. On the What-it-is-like-Ness of Experience. **Southern Journal of Philosophy**, v. 48, n. 1, p.8-27, 2010.

STICH, S. **Collected Papers, Volume 1:** Mind and Language, 1972-2010. New York: Oxford University Press, 2011.

STRAWSON, P. **Análise e Metafísica: Uma Introdução a Filosofia**. Trad. Amando Moraes de Oliveira. São Paulo: Discurso Editorial, 2002.

SUN, R; FRANKLIN, S. Computational Models of Consciousness: A Taxonomy and Some Examples. In: ZELAZO, P; MOSCOVITCH, M; THOMPSON, E. **The Cambridge Handbook of Consciousness**. Cambridge: Cambridge University Press, 2008. DOI:10.1017/CBO9780511816789.

TYE, M. **Consciousness and Persons**. Cambridge, MA: MIT Press, 2005.

TYE, M. **Consciousness, Color and Content**. Cambridge, MA: MIT Press, 2000.

TYE, M. **Ten Problems of Consciousness:** A Representational Theory of Mind. Cambridge: MIT Press, 1995.

TYE, M. **Tense Bees and Shell-shocked Crabs:** Are Animal Conscious?. Oxford: Oxford University Press, 2017.

TYE, M. The Problem of Simple Minds: Is There Anything it is Like to be a Honey Bee?. **Philosophical Studies**, v. 88, n. 1, p. 289–317, 1997.

WEISBERG, J. Higher-Order Theories of Consciousness. In: KRIEGEL, U. **Oxford Handbook of Philosophy of Consciousness**. Oxford: Oxford University Press, 2020. DOI: 10.1093/oxfordhb/9780198749677.013.20.

WEISBERG, J. Misrepresenting Consciousness. **Philosophical Studies**, v. 154, n. 3, p. 409-433, 2011. DOI: 10.1007/s11098-010-9567-3.

ZAHAVI, D; KRIEGEL, U. For-me-ness: What it is and What it is Not. In: DAHLSTROM, D; ELPIDOROU, A; HOPP, W. **Philosophy of Mind and Phenomenology.** Abingdon: Routledge, 2015.