



**UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS**

DÉBORA DE OLIVEIRA SILVA

UMA ANÁLISE DA QUESTÃO DA CONTINUIDADE TEÓRICA NA CIÊNCIA

CAMPINAS

2024

DÉBORA DE OLIVEIRA SILVA

UMA ANÁLISE DA QUESTÃO DA CONTINUIDADE TEÓRICA NA CIÊNCIA

Dissertação apresentada ao Instituto de Filosofia e Ciências Humanas da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Filosofia.

Orientador: Prof. Dr. Silvio Seno Chibeni

ESTE TRABALHO CORRESPONDE À
VERSÃO FINAL DA DISSERTAÇÃO
DEFENDIDA PELA ALUNA DÉBORA DE
OLIVEIRA SILVA, E ORIENTADA PELO
PROF. DR. SILVIO SENO CHIBENI.

**CAMPINAS
2024**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Filosofia e Ciências Humanas
Cecília Maria Jorge Nicolau - CRB 8/3387

Si38a Silva, Débora de Oliveira, 1997-
Uma análise da questão da continuidade teórica na ciência / Débora de Oliveira Silva. – Campinas, SP : [s.n.], 2024.

Orientador: Silvio Seno Chibeni.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Filosofia e Ciências Humanas.

1. Filosofia e ciência. 2. Realismo. 3. Linguagem. I. Chibeni, Silvio Seno, 1958-. II. Universidade Estadual de Campinas. Instituto de Filosofia e Ciências Humanas. III. Título.

Informações Complementares

Título em outro idioma: An analysis of the issue of theoretical continuity in science

Palavras-chave em inglês:

Philosophy and science

Realism

Language

Área de concentração: Filosofia

Titulação: Mestra em Filosofia

Banca examinadora:

Silvio Seno Chibeni [Orientador]

Marco Antonio Caron Ruffino

Bruno Borge

Data de defesa: 28-02-2024

Programa de Pós-Graduação: Filosofia

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0003-4660-5796>

- Currículo Lattes do autor: <https://lattes.cnpq.br/1294009200467921>



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE FILOSOFIA E CIÊNCIAS HUMANAS

A Comissão Julgadora dos trabalhos de Defesa de Dissertação de Mestrado, composto pelos(as) Professores Doutores a seguir descritos, em sessão pública realizada em 28/02/2024, considerou a candidata Débora de Oliveira Silva aprovada.

Prof. Dr. Silvio Seno Chibeni

Prof. Dr. Marco Antonio Caron Ruffino

Prof. Dr. Bruno Borge

A Ata de Defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertações/Teses e na Coordenadoria do Programa de Pós-Graduação em Filosofia do Instituto de Filosofia e Ciências Humanas.

AGRADECIMENTOS

Aos meus familiares e amigos, pelo apoio constante na minha caminhada até aqui.

Ao colega Gabriel Sardi, cuja parceria foi de grande importância para a concepção de alguns aspectos do texto.

Ao meu orientador, prof. Silvio Seno Chibeni, pelo acolhimento, apoio e dedicação rigorosa que foram fundamentais para a elaboração do presente trabalho. Também não posso deixar de lembrar do prof. Luiz Helvécio, meu orientador durante a graduação, que foi fundamental para consolidar as bases dessa pesquisa. Ao prof. Marco Ruffino, pelas indicações bibliográficas valiosas para o texto final e correções sugeridas. Ao Pedro Merluzzi e Bruno Borge pelos comentários e sugestões úteis apresentadas no exame de qualificação e defesa da dissertação, respectivamente.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (Capes) – Código de Financiamento 001.

RESUMO

No debate entre realistas e antirrealistas científicos — i.e., na discussão sobre o estatuto epistêmico das teorias científicas que pressupõem a existência de entes e processos inobserváveis —, os realistas defendem que o sucesso dessas teorias nos compromete com sua verdade aproximada. No entanto, casos históricos que constituem episódios de substituição teórica mostram que, ao longo da história da ciência, diversas teorias bem-sucedidas em sua época foram posteriormente classificadas como falsas. Este trabalho se propõe a avaliar a possibilidade de manutenção da referência de termos que putativamente designam inobserváveis entre as teorias científicas sucessivas com recurso às teorias da referência, de modo a defender a continuidade teórica na ciência e salvaguardar a convicção realista de que o sucesso das teorias científicas é devido a sua verdade aproximada e acerto na referência ao nível inobservável. O Capítulo 1 apresenta, em termos gerais, o debate entre realistas e antirrealistas científicos. O Capítulo 2 introduz alguns argumentos do debate, considerações realistas sobre o problema da descontinuidade teórica na ciência e os casos históricos mais relevantes a serem examinados por nós. O Capítulo 3 analisa as teorias da referência e suas propostas de aplicação aos casos históricos. Tais teorias são avaliadas no Capítulo 4, frente ao seu desempenho em casos históricos escolhidos. O Capítulo 5 tece algumas considerações gerais sobre a análise empreendida nesta dissertação.

Palavras-chave: Inobserváveis; Realismo Científico; Argumento do Milagre; Metaindução Pessimista; Teorias da Referência; Continuidade Teórica.

ABSTRACT

In the debate between realists and antirealists in science — i.e., the discussion about the epistemic status of scientific theories that presuppose the existence of unobservable entities and process —, realists defend that the success of these theories involves their approximate truth. Although, historical cases that constitute episodes of theoretical substitution show that, throughout the history of science, several successful theories in their time were afterwards classified as false. This work aims to evaluate the possibility of reference maintenance of terms that putatively designate unobservable entities between successful scientific theories with the help of theories of reference, in order to defend the theoretical continuity in science and save the realist conviction that the success of scientific theories is due to their approximate truth and correct reference at the unobservable level. Chapter 1 shows, in general terms, the debate between realists and antirealists in science. Chapter 2 introduces some arguments of the debate, realist considerations about the problem of theoretical discontinuity in science and some relevant historical cases. Chapter 3 analyzes the theories of reference and their application to the historical cases. Such theories are evaluated in Chapter 4, based on their performance in chosen historical cases. Finally, Chapter 5 formulates some general considerations about the analysis made in this dissertation.

Keywords: Unobservables; Scientific Realism; No-Miracle Argument; Pessimistic induction; Theories of Reference; Theoretical Continuity.

SUMÁRIO

Capítulo 1: Introdução	7
1.1. Introdução ao debate realismo <i>versus</i> antirrealismo científico	7
1.2. Os inobserváveis e suas controvérsias	11
1.3. Inferência da melhor explicação	13
1.4. Realismo e antirrealismo científico: aprofundando a análise	16
Capítulo 2: Argumentos pró e contra o realismo científico	20
2.1. O argumento do milagre	20
2.2. A subdeterminação da teoria pela evidência empírica	24
2.3. A metaindução pessimista	26
2.4. O problema da descontinuidade teórica na ciência	30
Capítulo 3: O papel das teorias da referência na defesa da continuidade teórica	38
3.1. Teoria descritivista	39
3.1.1. Adaptação da teoria descritivista aos termos teóricos	42
3.2. Teoria causal	45
3.2.1. Adaptação da teoria causal aos termos teóricos	48
3.3. Teoria híbrida	52
3.3.1. Adaptação da teoria híbrida aos termos teóricos	53
3.3.1.1. Com recurso aos tipos naturais	55
3.3.1.2. Sem recurso aos tipos naturais	59
3.3.2. Crítica à teoria híbrida: projeção realista do presente sobre o passado	63
Capítulo 4: Teorias da referência perante casos históricos	69
4.1. O éter luminífero	69
4.2. O flogisto	72
4.3. Teoria descritivista	74
4.4. Teoria causal	76
4.5. Teoria híbrida	77
4.5.1. O éter luminífero e a projeção realista no passado	77
4.5.2. O flogisto e a negligência realista	82
5. Considerações finais	85
Bibliografia	91

Capítulo 1: Introdução

1.1. Introdução ao debate realismo *versus* antirrealismo científico

O debate entre realistas e antirrealistas científicos é uma discussão sobre o estatuto epistêmico das nossas melhores teorias científicas que pressupõem a existência de inobserváveis — aquelas entidades postuladas pelas teorias que não acessamos diretamente com os nossos sentidos, e cuja suposta interação conosco é mediada por instrumentos de observação ou inferências. Exemplos clássicos são elétrons, fótons, ondas eletromagnéticas, bactérias, vírus, etc. Devido a essa falta de acesso perceptual direto às entidades inobserváveis, os antirrealistas colocam sob suspeita o que as teorias expressam sobre elas, seja por não acreditarem ou suspenderem o juízo acerca da existência delas, inibindo qualquer crença ou comprometimento epistêmico com tais entidades.

Entidades tão pequenas, como as bactérias ou os vírus, só podem ser visualizadas com o auxílio de microscópios; nesse nível já não há a disponibilidade do acesso direto para garantir a correção daquilo a ser entregue pelo instrumento. Partículas subatômicas como elétrons, neutrinos e quarks não podem ser visualizados através de nenhum instrumento, tornando-nos então dependentes de inferências (que partem da ocorrência de certos fenômenos observáveis) para aferirmos algo sobre as entidades nesse nível. Seja como for, qualquer acesso indireto aos inobserváveis contém uma dependência teórica e instrumental que é incontornável. Refletindo sobre os exemplos previamente citados, um instrumento como o microscópio depende de teorias ópticas e/ou eletromagnéticas (a depender da sua complexidade), que embasam sua construção e utilização; já as inferências são ainda mais pautadas nas teorias para nos informar sobre essas entidades.

Dada a falta de recursos epistêmicos diretos para a justificação do conteúdo das teorias científicas sobre os inobserváveis, é em princípio possível que elas estejam equivocadas naquilo que expressam sobre esse âmbito do mundo. Tal problema suscita a discordância entre realistas e antirrealistas, fazendo-os assumir atitudes opostas no debate.

Quanto às teorias científicas que postulam inobserváveis, parece que a pressuposição da existência dessas entidades é crucial para que sejam bem-sucedidas. No presente contexto, o sucesso é entendido como sendo constituído principalmente pela adequação empírica e pelos poderes preditivo e unificador. Uma teoria é dita empiricamente adequada se expressa a verdade sobre os fenômenos observáveis (que podemos acessar diretamente com os nossos sentidos); é preditivamente poderosa se antecipa a ocorrência de

fenômenos novos, e é unificadora se unifica fenômenos aparentemente díspares. Neste ponto da análise não incluímos o poder explicativo como um fator constitutivo do sucesso teórico.¹ Adotamos tal postura porque I) não é consenso entre realistas e antirrealistas que o poder explicativo esteja incluído na definição de sucesso. II) A satisfação do requisito explicativo parece ser facilmente alcançada por qualquer concepção que seja proposta com um pouco de engenhosidade para explicar algum fato, pois ela só precisa incorporar os fatos e dar sentido a eles a partir das razões apresentadas; basta pensar nas explicações oferecidas por alguma mitologia para qualquer aspecto intrigante do mundo. É certo que admiramos a ciência devido ao seu potencial explicativo, mas a devida caracterização do sucesso da ciência deve incluir fatores que sejam realmente diferenciais para distingui-la de outras áreas investigativas. III) Mais adiante, veremos que também não há consenso sobre o que consiste uma boa explicação, de modo que falta clareza na própria caracterização do poder explicativo de uma teoria. IV) Há filósofos antirrealistas, como Bas van Fraassen (1980, cap. 5), que diminuem o papel das explicações na ciência, salientando que elas têm um caráter subjetivo, contextual, cultural, etc, devendo-se, portanto, evitar qualquer compromisso epistêmico com ela; na verdade, van Fraassen propõe uma teoria pragmática da explicação, compatível com o referencial antirrealista geral de sua filosofia da ciência.² No entanto, o poder explicativo será relevante em outros estágios distintos de análise deste trabalho, como nas considerações sobre melhor explicação ou em alguns argumentos do debate.

É inegável que a ciência é um empreendimento bem-sucedido, no sentido acima descrito; mas será que isso implica algum comprometimento com os inobserváveis supostos pelas teorias científicas? Os realistas alegam que sim. Já os antirrealistas discordam, pois a falta de acesso direto aos inobserváveis acaba, segundo eles, impedindo a obtenção de uma justificação adequada a favor da existência de tais entidades.

¹ Na presente ocasião não caracterizamos a explicação no sentido hempeliano (nomológico-dedutivo), no qual a explicação se reduz à previsão; nesse sentido, a explicação científica deve deduzir a ocorrência do(s) fenômeno(s) (*explanandum*) a partir de leis gerais e condições iniciais (*explanans*) (HEMPEL, 1965, p. 336). Para uma análise do que parece ter sido a formulação pioneira da noção nomológico-dedutiva da explicação científica, por parte de Berkeley, ver Chibeni (2008; 2010). Aqui entendemos a explicação segundo o senso comum (tal qual foi discutida no debate nas últimas décadas), em que explicar é poder compreender os fenômenos a partir das suas supostas causas. Essa foi também a concepção de explicação adotada, implícita ou explicitamente, por grande parte dos filósofos antigos, medievais e modernos, a começar com destaque para Aristóteles.

² Segundo ele, “[não] pode haver qualquer questão sobre o sucesso explicativo como algo que ofereça evidência a favor da verdade de uma teoria que vá além de qualquer evidência que tenhamos em favor do fato de que ela oferece uma descrição adequada dos fenômenos. Pois, em cada caso, um sucesso em explicar é um sucesso em descrever de maneira adequada e informativa. E, embora seja verdade que procuramos dar explicações, o valor dessa procura para a ciência é que a procura por explicação é *ipso facto* uma procura por teorias empiricamente adequadas e fortes.” (VAN FRAASSEN, 1980, pp. 156-157).

O debate sobre os inobserváveis tornou-se mais visível na filosofia da ciência na primeira metade do século XX, embora estivesse presente, sob outras denominações, em autores modernos.³ Esta discussão foi inicialmente fomentada pelo lado antirrealista da discussão, em especial, por filósofos adeptos do positivismo (ou empirismo) lógico, que, no início do século XX, buscavam repensar o empirismo face aos avanços das ciências físicas e formais, rejeitavam a metafísica com base em um critério empirista de significado e valorizavam o papel da lógica assim como a análise da linguagem para a eliminação de pseudoproblemas. O forte viés empirista desses filósofos, segundo o qual a fonte e a justificação do conhecimento remontam à experiência (ao que podemos captar com os nossos sentidos), os conduzia, ao menos na fase inicial do projeto positivista, a uma rejeição robusta de qualquer comprometimento acerca das entidades inobserváveis. No entanto, era preciso explicar como a menção aos inobserváveis poderia ser evitada ou ter sua importância diminuída. Os positivistas então propuseram uma espécie de “reducionismo”, segundo o qual, embora os enunciados teóricos sobre os inobserváveis sejam legítimos e dotados de valor de verdade, eles não devem ser interpretados literalmente, mas reduzidos a enunciados sobre observáveis, mediante certas regras de correspondência ou outros artifícios. Como o acesso direto a tais entidades observáveis nos permite atestar a correção daquilo que as teorias expressam sobre elas, os observáveis não suscitam as mesmas controvérsias dos inobserváveis; por essa razão, os positivistas reconheciam os observáveis como uma fonte genuína de referência (indireta) e significado dos enunciados das teorias científicas, mesmo aqueles que aparentemente se referem a inobserváveis. Rudolf Carnap (1928; 1936) foi um dos primeiros a tentar traduzir ou reduzir o significado do discurso “teórico” sobre os inobserváveis às entidades observáveis através de definições explícitas ou sentenças redutivas. Porém, como tal estratégia gerava uma multiplicidade de dificuldades e objeções, ela acabou não prosperando como se esperava. Na mesma linha, havia o *operacionalismo* (Bridgman, 1927), que visava reduzir a menção de termos referentes aos inobserváveis (i.e., os termos teóricos)⁴ a procedimentos experimentais de laboratório; contudo, por não assimilar

³ Como pode ser conferido em Chibeni (1990).

⁴ Por simplicidade, usaremos a expressão “termo teórico” ao longo do texto no lugar de “termos que putativamente referem a inobserváveis.” Logo adiante será mencionado que os próprios positivistas lógicos não conseguiram distinguir adequadamente os termos teóricos (sobre inobserváveis) dos termos observacionais (para observáveis). Em um sentido literal, todo termo da linguagem é teórico; na medida em que os termos resultam de definições, refinamentos, delimitações, etc., eles já são um constructo que excedem a experiência. E como muito bem coloca van Fraassen, “... qualquer entidade inobservável vai diferir daquelas que são observáveis na maneira pela qual ela sistematicamente carece de características observáveis. Enquanto não renunciarmos à negação, portanto, vamos ser capazes de enunciar no vocabulário observacional (como quer que ele seja concebido) que há entidades inobserváveis e, em alguma medida, como elas são. A teoria quântica, na versão de Copenhagen, implica que há coisas que às vezes possuem posição no espaço e às vezes, não. Acabo de enunciar essa

todos os casos de aplicação dos termos teóricos, o operacionalismo também não prosperou como se esperava inicialmente.

Outra proposta bem conhecida em filosofia da ciência — além do reducionismo dos positivistas — é o chamado *instrumentalismo* (Mach, 1893; Duhem, 1906), que considera que a menção teórica aos inobserváveis, na linguagem da ciência, é um mero instrumento para a predição e organização de sentenças sobre fenômenos observáveis.⁵ Diferente do que ocorre com o reducionismo, nesta proposta, os enunciados teóricos sobre os inobserváveis não são considerados legítimos, ou seja, não possuem valor de verdade; longe de descreverem aspectos inobserváveis do mundo, tais enunciados seriam apenas (em uma das versões) regras de inferência para a organização de fenômenos observáveis.

Uma abordagem bastante sofisticada do problema é a do teorema de Craig (Craig, 1956), que estabelece que se uma teoria cumpre certos requisitos é possível, mediante certas manipulações formais, expressá-la em um vocabulário observacional que só pressupõe a existência de entidades observáveis. Com isso tentou-se reformular as teorias científicas de modo a eliminar a menção aos inobserváveis. Entretanto, tal reformulação gerou formulações demasiadamente complexas e artificiais das teorias, que aparentemente não captavam uma parte considerável do poder inferencial da teoria original, uma vez que a aplicação do Teorema de Craig só permite a manutenção dos aspectos dedutivos de uma teoria. Como consequência dessa reformulação teórica, perdia-se o poder unificador das teorias e sua capacidade inferencial que nos permite ampliar o conhecimento sobre o mundo, aspectos resultantes do caráter indutivo das teorias científicas, incapazes de serem subsumidos em uma reformulação dedutiva das mesmas.

Com o insucesso dessas investidas iniciais dos antirrealistas — a saber, o reducionismo e o instrumentalismo —, eles acabaram por admitir que o discurso teórico sobre os inobserváveis tem um excesso de significado que não pode ser apropriadamente expresso em termos puramente observacionais sem afetar consideravelmente o fator de sucesso das teorias científicas. Outro aspecto que também minou as ambições iniciais dos antirrealistas foi a dificuldade em distinguir adequadamente o vocabulário teórico (sobre inobserváveis) do

consequência sem utilizar um único termo teórico. A teoria de Newton implica que há alguma coisa (isto é, o Espaço Absoluto) que nem possui uma posição, nem ocupa volume. Tais consequências não decorrem de forma alguma da imaginação sobre o que há no mundo observável, nem de como qualquer coisa observável é. A teoria reduzida T/E [i.e., teoremas da teoria expressos no subvocabulário observacional da teoria] não é uma descrição de uma parte do mundo descrito por T [teoria]; ao contrário, de uma forma claudicante e ineficiente, T/E é a descrição de tudo, feita por T (VAN FRAASSEN, 1980, pp. 54-55). Tendo isso em conta, a expressão “termo teórico” não será usada em um sentido literal, mas apenas para fins de simplicidade.

⁵ Em Popper (1963, cap. 3) e Nagel (1961, cap. 6) há uma fecunda apresentação e discussão do instrumentalismo.

observacional (para observáveis) a fim de expressar um em termos de outro. Em contrapartida, tal admissão não implicou um comprometimento antirrealista com os inobserváveis, mas inspirou outras posturas menos radicais. Carnap, por exemplo, tentou adotar uma estratégia neutra (que se distanciava tanto do instrumentalismo como do realismo); para tal, com o auxílio da sentença-Ramsey,⁶ ele defendeu que a estruturação do discurso teórico só implicava um comprometimento com a existência de entidades matemáticas (1963, p. 962).⁷ E mais recentemente, temos o empirismo construtivo proposto por Bas van Fraassen (1980), a variante atualmente predominante do antirrealismo, que abdica tanto do reducionismo como do instrumentalismo e sugere apenas uma postura agnóstica quanto aos inobserváveis, acoplada a uma interpretação literal da linguagem da ciência.

1.2. Os inobserváveis e suas controvérsias

O debate que mais especificamente nos interessa nesta dissertação inicia-se pela controvérsia sobre quais entidades de fato motivam a discussão do estatuto epistêmico das teorias científicas, já que a própria distinção entidades observáveis e inobserváveis, que fundamenta (no presente contexto) quais itens referidos pelas teorias são passíveis de questionamento, é por si só objeto de disputa. Comumente anuncia-se que tal distinção é uma incumbência que deveria preocupar os adeptos do antirrealismo, pois se eles fazem ressalvas quanto ao que as teorias dizem sobre o âmbito inobservável do mundo, é preciso que delimitem até onde se estende seu comprometimento epistêmico. Todavia, como os próprios realistas recorrem ao desempenho da teoria quanto aos (ditos) fenômenos observáveis para apoiarem suas convicções e defenderem os meios de acesso indireto às entidades inobserváveis (instrumentos ou inferências utilizados para tal), além de fazerem ressalvas sobre quais aspectos do mundo inobservável deveriam ser passíveis do nosso

⁶ A sentença-Ramsey foi proposta por Frank Ramsey em 1929 no artigo “Theories” como um recurso alternativo para a expressão de teorias. Uma teoria original (sem alteração) tem a estrutura lógica $TC(t, o)$ — com “t” representando os termos teóricos (relativos aos inobserváveis) e “o” os termos observacionais (sobre fenômenos observáveis). Para construir uma sentença-Ramsey de uma teoria precisamos eliminar seus termos teóricos (t), pois não temos certeza se eles realmente referem; após eliminá-los os substituiremos por variáveis existencialmente quantificadas, obtendo $TC^R \exists X TC(X, o)$. Essa é a forma ramseyficada da teoria original que expressa que “há um conjunto de entidades com tais e tais propriedades”. Esse conjunto de entidades com tais e tais propriedades satisfaz a sentença-Ramsey da teoria considerada, constituindo-se como o referente dela. Sem maiores restrições, a sentença-Ramsey de uma teoria pode ser satisfeita por qualquer tipo de entidade, seja ela concreta ou abstrata; por essa razão a sentença-Ramsey foi utilizada por Carnap para alegar que a reformulação do discurso teórico nos compromete apenas com a existência de entidades matemáticas.

⁷ Uma exposição aprofundada do início da discussão sobre os inobserváveis no século XX pode ser encontrada em Psillos (1999, cap. 1, 2 e 3) e Olegario da Silva (2020).

comprometimento — como é propagado por proponentes do realismo de entidades ou realismo estrutural (variantes do realismo que serão apresentadas no decorrer do texto) —, julgamos que os próprios realistas também deveriam se preocupar com tal distinção.

Como notaram alguns realistas (e.g. Maxwell, 1962), a presente distinção é relativa à espécie, pois há diversos animais que são capazes de perceber coisas que nos passam despercebidas — por exemplo, devido ao olfato avantajado dos cães eles têm mais facilidade em encontrar objetos escondidos, os morcegos escutam ultrassons e utilizam dessa vantagem para se guiarem no espaço, etc. Embora isso tenha constituído obstáculo para certas formas de antirrealismo científico, Bas van Fraassen manobrou para assimilar o ponto em seu empirismo construtivo, propondo que nessa discussão epistêmica devemos nos ater ao que a comunidade epistêmica humana pode ou não acessar diretamente a fim de delimitar até onde (supostamente) se estende o nosso conhecimento,

o que conta como um fenômeno observável é função do que é a comunidade epistêmica (que *observável* é *observável-para-nós*) ... mesmo que a observabilidade não tenha nada a ver com a existência (ela é, de fato, antropomórfica demais para isso), ela ainda teria muito a ver com a atitude epistêmica apropriada em relação à ciência” (VAN FRAASSEN, 1980, p. 19).

Tal distinção também não é de tipo, mas é de grau, por ser impossível traçar de modo categórico o início e o término do âmbito observável do mundo (MAXWELL, 1962, pp. 7-11). E comumente o filósofo e o cientista discordam sobre o que é diretamente observável ou não,

há um contínuo que começa com as observações sensoriais diretas e procede até ao enormemente complexo, [i.e.], métodos indiretos de observação. Obviamente, nenhuma linha precisa pode ser traçada nesse contínuo; é uma questão de grau. Um filósofo está seguro de que o som da voz da sua esposa, vindo do outro lado da sala, é um observável. Mas suponha que ele a escute pelo telefone. Sua voz é um observável ou não? Um físico certamente diria que quando ele olha para alguma coisa através de um microscópio ordinário [então] ele está observando isso diretamente. [Mas] esse também seria o caso quando ele usa um microscópio eletrônico? Será que ele observa o caminho de uma partícula quando ele vê o seu rastro em uma câmara de nuvens? Em geral, o físico fala sobre observáveis de uma maneira muito mais ampla do que o filósofo, mas, em ambos os casos, a linha que separa os observáveis dos inobserváveis é altamente arbitrária. (CARNAP, 1966, p. 226).

Embora falte uma distinção precisa e inteiramente objetiva entre observáveis e inobserváveis, a discussão envolvida no debate ainda é muito importante e pode ser adequadamente motivada. O debate por si só já é suscitado pela existência de certas entidades que não podem ser observadas de maneira nenhuma, isto é, que são inobserváveis em princípio. Como mencionado há pouco, quando observamos um rastro em uma câmara de nuvens não acessamos diretamente um elétron, a suposta partícula subatômica que produz

esse fenômeno observável. E para quem alegue que é possível detectar a presença de um elétron a partir de um rastro deixado na câmara de nuvens, temos de notar que detectar e observar são processos distintos. Ao observarmos uma entidade a acessamos diretamente, enquanto ao detectá-la estamos meramente inferindo a sua presença a partir de um efeito (supostamente) produzido por ela.⁸ Dado que a inferência não oferece a mesma segurança que um acesso direto à entidade (em condições adequadas), ela constitui um caso que envolve maiores possibilidades de equívocos. De todo modo, a mera existência de exemplares flagrantes de observáveis e inobserváveis — como o elétron — é suficiente para motivar a discussão sobre o estatuto epistêmico das teorias científicas. São essas entidades que em princípio não são passíveis de observação que fomentam integralmente a discussão, uma vez que elas não oferecem margem para a perspectiva de que a crença acerca delas possa ser posteriormente justificada através de algum acesso direto ao mundo.

A presente discussão incide majoritariamente sobre as entidades inobserváveis postuladas na física, química e biologia. Entretanto, uma vez que teorias de outras áreas científicas pressupõem aspectos do mundo natural trabalhados por essas ciências, tal discussão acaba igualmente gerando implicações para outras áreas da ciência.⁹

1.3. Inferência da melhor explicação

As entidades inobserváveis postuladas pela ciência têm um importante papel explicativo, tal como as ondas eletromagnéticas que explicam o funcionamento de vários aspectos do nosso cotidiano como a luz, o forno de micro-ondas, o *wi-fi*, etc., diversas outras entidades como essa oferecem uma compreensão ampliada de vários outros fenômenos observados por nós. Em um esboço simplificado, podemos dizer que na ciência a existência de um inobservável é suposta através de uma hipótese (ou teoria) que estabelece esse tipo de entidade como o fator responsável pela ocorrência de certos fenômenos observáveis; ela vem

⁸ Tanto que só observamos o (suposto) efeito produzido pela entidade, mas não a própria entidade, como explica muito bem van Fraassen através do seguinte exemplo “a linha cinza-prata resultante [do rastro de um elétron em uma câmara de nuvens] é similar (fisicamente, assim como em aparência) à trilha de vapor deixada no céu quando um jato passa. Suponhamos que eu aponte tal trilha e diga: “Olhe, lá está um jato!” Alguém poderia dizer: “Vejo a trilha de vapor, mas onde está o jato?” Então, eu responderia: “Olhe logo à frente da trilha, ... lá” Você o vê?” Ora, no caso da câmara de vapor, essa resposta não é possível. Assim, apesar de ser a partícula detectada por meio da câmara de vapor, e essa detecção estar baseada em observação, claramente, esse não é um caso de estar a partícula sendo observada” (VAN FRAASSEN, 1980, p. 17).

⁹ Essa discussão alastra-se para outras áreas como a geologia, paleontologia, arqueologia, história, etc.; nessas áreas, o acesso direto sobre seus respectivos objetos de estudo dificilmente está disponível. Dessa maneira, o debate apresenta-se como muito mais fecundo e interessante do que usualmente parece. Turner (2007) explora as consequências desses desdobramentos.

a ser escolhida (dentre outras) a partir de certas observações e suposições de fundo, sendo essa escolha guiada, sobretudo, pelo potencial explicativo da hipótese que supõe a existência dessa entidade.

Embora a explicação não esteja incluída na definição de sucesso teórico com a qual estamos trabalhando, inegavelmente ela é um fator que cumpre um papel relevante na prática científica, como na etapa de escolha entre hipóteses. Uma hipótese não científica também pode ser explicativa, contudo, quando temos duas hipóteses científicas em disputa, a capacidade explicativa de cada uma delas pode fazer diferença em tal processo. Sendo o sucesso teórico e a *inferência da melhor explicação* estágios distintos da nossa análise, não vemos maiores problemas em avaliar a explicação distintamente em cada uma delas.

O tipo de raciocínio que subjaz à escolha de uma hipótese tendo em vista o seu papel explicativo é a *inferência da melhor explicação* (IME). Charles S. Peirce (1934-1935) foi um dos primeiros a enfatizar a importância desse tipo de raciocínio, denominando-o inicialmente como *inferência abdutiva*; posteriormente na década de 1960, Gilbert Harman (1965; 1968) explicitou como essa inferência auxilia na comparação e escolha de hipóteses ou teorias, renomeando-a com a denominação atual.¹⁰ A IME é primordial para a escolha de uma hipótese explicativa, e é aplicada tanto na ciência como no nosso cotidiano — de modo simples, quando as nossas chaves somem misteriosamente, presumimos que esquecemos onde as colocamos e não que um gnomo as tenha escondido para nos injuriar, pois a primeira explicação é melhor do que a última face a tudo o que já observamos e supomos existir. Com o auxílio da IME, inferimos uma dada conclusão se ela é a melhor explicação para um certo conjunto de fatos. Porém, a verdade das premissas de uma IME não garante a verdade da sua conclusão, tornando-a falível (CHIBENI, 2006, p. 223). Se formalizarmos a estrutura da IME obtemos o seguinte: $H \rightarrow E$; E ; logo H .¹¹ Dedutivamente, essa forma de raciocinar é intitulada *falácia da afirmação da consequente*, que é inválida porque a obtenção da consequente (E = evidência) não nos permite inferir com certeza a antecedente (H = hipótese). A conclusão de uma IME amplia ou extrapola o que é informado nas premissas, impedindo a garantia de que a verdade das premissas se estenda à conclusão. A ampliação pode ser horizontal, na medida em que a conclusão menciona o mesmo tipo de entidade já citado nas premissas (e apenas aumenta a quantidade de entidades de um dado tipo), ou vertical, quando na conclusão é

¹⁰ Uma instrutiva fonte de consulta sobre essa inferência encontra-se em Lipton (2004); ver também Chibeni (2006). Alguns autores, como Campos (2011), têm proposto que a abdução de Peirce não deve ser identificada com a IME.

¹¹ Tal formalização remonta à Peirce (CHIBENI, 2006, p. 223n).

menção de um novo tipo de entidade não citada nas premissas, sendo esse o caso da postulação de novas entidades inobserváveis.

Uma dada conclusão é inferida em uma instância da IME em virtude de considerações explicativas; assim, se uma determinada hipótese é a melhor explicação para os fatos citados nas suas premissas, ela é escolhida, na conclusão. Uma boa explicação é constituída por virtudes teóricas como *consistência* e *coerência* (com teorias de fundo), *simplicidade*, *escopo* e *unidade*, etc. Uma explicação é *consistente e coerente* com outras teorias de fundo caso não as contradiga, é *simples* ao ter menos pressuposições, o *escopo* trata-se do conjunto de fenômenos subsumidos na sua explicação e a *unidade* consiste na unificação desses fenômenos a partir da hipótese explicativa apresentada.

O caráter falível da IME por si só já a torna alvo de suspeitas. No entanto, é essa forma de raciocínio que tipicamente embasa a escolha de hipóteses que introduzem a existência de inobserváveis. Devemos notar, é claro, que a ciência é um empreendimento também falível e passível de revisão de um modo geral, sendo esse o caráter que a distingue de outras práticas dogmáticas tornando-a uma fonte valiosa que nos auxilia na antecipação de fenômenos, soluções de problemas e produção de artefatos tecnológicos. Mas qual é o limite aceitável da falibilidade (sobretudo quando consideramos realistas e antirrealistas)? O requisito mínimo de uma boa explicação — e que sequer constitui as virtudes teóricas que auxiliam os cientistas na escolha entre hipóteses — é a sua adequação empírica, pois se uma explicação nem ao menos está preditivamente certa quanto àquilo que visa explicar, ela é liminarmente descartada; assim, as melhores explicações (escolhidas através da IME) devem, no mínimo, ser adequadas à experiência. No entanto, veremos que mais de uma hipótese explicativa pode adequar-se à experiência, suscitando a difícil questão de como se deve fazer a escolha entre hipóteses empiricamente equivalentes. Poder-se-ia objetar que as virtudes teóricas nos socorrem nesse caso de empate empírico, porém, não é consenso que tais virtudes nos apontem qual é a melhor hipótese a ser escolhida; ou, se apontarem, que peso epistêmico essas virtudes podem ter; ou seja, o quanto elas de fato justificam a escolha feita, como a que tem melhores credenciais para a verdade (e não apenas por razões pragmáticas). Na subseção 2.2, teremos a oportunidade de conferir essas duas críticas à IME. Adicionalmente, o questionamento do fator explicativo por parte do antirrealista atingirá a legitimidade da IME e uma de suas instâncias utilizadas por certos realistas para argumentarem a favor da sua posição, como em breve verificaremos na subseção 2.1.

1.4. Realismo e antirrealismo científico: aprofundando a análise

O realismo científico é uma tese sobre a extensão do nosso conhecimento. Segundo a tese realista, no campo da ciência, somos capazes de conhecer aspectos do mundo que ultrapassam a nossa capacidade de observação. Podemos compreender melhor o realismo científico através das suas três dimensões de comprometimento no debate, que evidenciam aquilo que um realista defende. Metafisicamente, as teorias científicas dizem respeito a uma realidade independente da mente humana. Semanticamente, os enunciados teóricos são interpretados literalmente, i.e., esses enunciados são dotados de valor de verdade e são compostos por termos que visam fazer referência a inobserváveis. E epistemicamente, as teorias nos fornecem (algum tipo de) conhecimento sobre o mundo, seja em seus aspectos observáveis ou inobserváveis. Qualquer forma de antirrealismo científico negará pelo menos uma dessas dimensões. Mas antes de caracterizarmos esse outro lado do debate, é importante distinguir alguns tipos de realismo *seletivo*, que assimilam ressalvas quanto ao comprometimento epistêmico com os inobserváveis, delimitando quais os componentes presentes nas teorias demandam uma postura realista. Podemos distinguir de maneira geral o *realismo explanacionista*, *realismo de entidades* e *realismo estrutural*. Segundo o explanacionismo (Kitcher, 1993; Psillos, 1999), devemos nos comprometer com aqueles pressupostos das teorias científicas que explicam o seu sucesso teórico — como inobserváveis, leis, etc., independentemente de eles serem específicos como entidades ou estruturas. O realismo de entidades (Hacking, 1983; Cartwright, 1983) sugere o comprometimento tão somente com a existência das *entidades* inobserváveis (e não com detalhes do que as teorias dizem sobre elas), na medida em que temos contato causal com elas, i.e., quando é possível manipulá-las e usá-las para intervir em outros fenômenos.¹² E o realismo estrutural (Worrall, 1989; Ladyman, 1998) adota um ceticismo quanto aos inobserváveis, reservando o comprometimento apenas com a *estrutura* do âmbito inobservável (representando nas relações descritas pelas nossas teorias científicas). O foco deste trabalho recairá sobre o realismo explanacionista, uma vez que tencionamos defender o conhecimento sobre os inobserváveis por pensar que essas entidades cumprem um papel importante na geração de sucesso na ciência.

¹² No entanto, alguns autores entendem que Nancy Cartwright seria instrumentalista, porque ela adota uma postura antirrealista quanto às leis científicas e a IME. Mas Cartwright acredita que, embora as leis sejam literalmente falsas, elas conseguem captar aspectos causais do mundo (que em sua visão são capacidades ou poderes) que embasam a crença na existência de entidades inobserváveis.

Em contrapartida, o antirrealismo científico defende que só podemos ter conhecimento do âmbito observável do mundo. Os antirrealistas duvidam da existência dos inobserváveis e negam que o nosso comprometimento epistêmico deva se estender a tais entidades. Uma postura antirrealista nesse debate pode ser motivada por: i) empirismo estrito, ii) economia ontológica, iii) *subdeterminação* da teoria pelos dados e iv) considerações históricas.

De acordo com o empirismo estrito (i), o conhecimento não só se origina e é ultimamente justificado pela experiência, mas também se limita a ela. Se essa experiência é concebida como experiência direta,¹³ em termos do que podemos acessar diretamente com os nossos cinco sentidos, então pode ser colocado o questionamento sobre os inobserváveis. Nesta concepção, só temos garantia epistêmica quanto ao que é expresso sobre as entidades observáveis, pois podemos atestar diretamente o que é dito acerca delas. Assim sendo, a inferência do observável ao inobservável é considerada precária e deve ser evitada. Nem mesmo os casos das chamadas “detecções” convencem um antirrealista, por não serem experiências diretas do mundo elas abrem margem para a possibilidade de equívoco nas inferências que as acompanham. Por exemplo, detectar a presença de um elétron através de um rastro deixado em uma câmara de nuvens não é o mesmo que observá-lo diretamente, por isso as dúvidas colocadas pelo antirrealista aplicam-se a meros casos de detecção. Aspirando uma economia ontológica (ii), o antirrealista propõe-se a prescindir do comprometimento com os inobserváveis e optar por uma caracterização alternativa adequada da ciência, em que seu conteúdo efetivo se limite ao nível empírico.

Por sua vez, a *subdeterminação* da teoria pelos dados (iii) é um problema que pode atingir qualquer hipótese (teoria) explicativa sobre a experiência. Nos casos de subdeterminação, a experiência não determina qual teoria deve ser escolhida, pois teríamos duas ou mais teorias diferentes implicando o mesmo conjunto de fenômenos, sendo empiricamente equivalentes. Nesse tipo de caso, se uma teoria que pressupõe inobserváveis equivale empiricamente a uma outra teoria que tem uma concepção totalmente diferente sobre tais entidades — seja caracterizando-as de outro modo ou sem supor sua existência —, a experiência, isolada de outros fatores, não nos ajuda a tomar uma decisão, situação que

¹³ Estamos tomando como experiência direta aquilo que no debate também é denominado como percepção direta, i.e., qualquer observação realizada com os nossos cinco sentidos sem o auxílio de instrumentos. No debate entre realistas e antirrealistas científicos são aquelas entidades que, em princípio, não podemos acessar somente com os nossos sentidos — i.e., os inobserváveis — que acabam suscitando toda a discussão, uma vez que a experiência direta é considerada pelo antirrealista como mais confiável do que a experiência indireta para adquirir conhecimento sobre o mundo; e, também por essa razão, a percepção de objetos observáveis não chega a ser problematizada.

sugere, novamente, evitar qualquer comprometimento com as entidades inobserváveis postuladas nessas teorias. De fato, a subdeterminação é um problema que afeta a confiabilidade da IME, ao minar a objetividade ou justificação da escolha teórica, enfraquecendo justamente a inferência que embasa a postulação das entidades inobserváveis. E, por fim, considerações históricas sobre episódios de substituição teórica (iv) também enfraquecem as convicções realistas, tornando-se uma razão adicional para posturas antirrealistas nessa discussão. Ao longo da história da ciência não faltam relatos de casos de teorias que foram exitosas na predição de fenômenos observáveis, embora posteriormente tenham sido abandonadas, comprometendo-se assim a crença quanto ao que as teorias científicas aparentemente estabelecem no âmbito inobservável do mundo. No decorrer do texto abordaremos novamente as motivações (iii) e (iv) — sendo esta última o foco deste trabalho —, nas respectivas subseções 2.2 e 2.3.

Ao apresentar as dimensões de comprometimento do realismo científico, indicamos que qualquer forma de antirrealismo negaria pelo menos uma delas. O reducionismo (envolvido na interpretação da ciência proposta pelos positivistas lógicos), assim como o instrumentalismo, negava tanto a dimensão semântica do debate quanto a epistêmica, na medida em que defendia uma interpretação não-literal da linguagem da ciência. No entanto, os reducionistas discordavam da dimensão semântica apenas por não interpretarem as sentenças teóricas putativamente sobre inobserváveis literalmente, mas ainda aceitavam que tais sentenças eram enunciados legítimos, dotados de valor de verdade. Já os instrumentalistas não interpretavam tais sentenças literalmente porque as consideravam meras elipses ou artifícios para a predição e organização de fenômenos observáveis, que funcionariam (numa versão mais conhecida) como regras de inferência em vez de proposições genuínas sobre o mundo; desse modo, na concepção instrumentalista, aquelas sentenças não seriam legítimas, sendo desprovidas de valor de verdade. Mesmo com suas devidas nuances, em ambas as posições, a negação da dimensão semântica por parte dos reducionistas e instrumentalistas evita o comprometimento epistêmico com as entidades inobserváveis. Como decorrência, a dimensão epistêmica é recusada, pois sequer há uma interpretação literal da linguagem que admita a existência de inobserváveis e a possibilidade do seu conhecimento.

No entanto, as reformulações reducionistas e instrumentalistas das teorias científicas não foram bem-sucedidas, se as considerarmos com a vantagem da retrospectiva. Isso acabou motivando e cedendo espaço para realistas que buscavam argumentar a favor dos inobserváveis por meio da explicação do sucesso da ciência, como Smart (1963) e Putnam (1975a). Dialeticamente, com a proposição do empirismo construtivo por Bas van Fraassen

em 1980, o lado antirrealista do debate fortaleceu-se novamente, na medida em que tinha agora um alvo mais específico, menos sobrecarregado de pressuposições alheias ao cerne da questão. Ciente das falhas prévias do antirrealismo (e sobretudo do instrumentalismo com a sua recusa em atribuir legitimidade ao discurso teórico sobre inobserváveis), van Fraassen concorda com o realista quanto à dimensão semântica e só rejeita a epistêmica. Segundo ele, devemos interpretar literalmente todo o discurso científico (quer seja sobre observáveis ou não), mas dada a falta de acesso direto aos inobserváveis não podemos garantir que as teorias nos fornecem conhecimento sobre essas entidades, abdicando então da crença sobre elas.

Após decidir que a linguagem da ciência deve ser compreendida literalmente, ainda podemos dizer que não há necessidade de acreditarmos que as boas teorias sejam verdadeiras, nem, *ipso facto*, que as entidades que elas postulam são reais. *A ciência visa nos dar teorias que são empiricamente adequadas; e a aceitação de uma teoria envolve apenas a crença de que ela é empiricamente adequada.* Esse é o enunciado da posição antirrealista que defendo; a denominarei *empirismo construtivo*. (VAN FRAASSEN, 1980, pp. 11-12).

Repare que van Fraassen indica que o objetivo da ciência não é desenvolver teorias verdadeiras *in toto*, mas sim, apenas, empiricamente adequadas, isto é, “... que diz a verdade sobre as coisas e eventos observáveis no mundo” (VAN FRAASSEN, 1980, p. 12). Como consequência, ele “... deseja meramente ser agnóstico sobre a existência dos aspectos inobserváveis do mundo descritos pela ciência” (VAN FRAASSEN, 1980, p. 72).

Vamos, no capítulo seguinte, examinar alguns dos argumentos mais relevantes no debate entre realistas e antirrealistas científicos.

Capítulo 2: Argumentos pró e contra o realismo científico

2.1. O argumento do milagre

O fato de as teorias científicas serem bem-sucedidas (i.e., empiricamente adequadas, dotadas de poder preditivo e unificador) é a principal razão alegada pelos realistas científicos a favor da existência de entidades inobserváveis. Segundo eles, caso tais entidades não existissem, teorias que as pressupõem não seriam tão bem-sucedidas como o são. Pelo sucesso das teorias que pressupõem a existência de inobserváveis estaríamos, em alguma medida, autorizados a inferir que essas entidades de fato existem e que as teorias que as postulam são aproximadamente verdadeiras.¹⁴ Tais convicções realistas são resumidas no célebre *argumento do milagre* (AM).

Um dos precursores do AM é o *argumento da coincidência cósmica* concebido por J. J. C. Smart (1963, p. 39). Smart sugere que se as nossas teorias científicas não fossem verdadeiras, seu sucesso seria inexplicável (ou, metaforicamente, devido à alguma “coincidência cósmica”). Seria demasiadamente improvável que tais teorias fossem falsas sendo tão bem-sucedidas, de modo que só uma grande coincidência de vastas proporções poderia explicar esse sucesso. Esse argumento foi desenvolvido por Smart para questionar a plausibilidade da visão instrumentalista das teorias, pois tal concepção tem dificuldade em explicar o sucesso teórico considerando a menção aos inobserváveis como meros instrumentos de predição de fenômenos observáveis.

Mas a versão do AM que veio a ser célebre na literatura do debate, tornando-se alvo de críticas, foi formulada por Hilary Putnam, na fase realista de seu pensamento (1975a),

o realismo é a única filosofia que não torna o sucesso da ciência um milagre... [Segundo tal concepção] os termos das teorias científicas maduras tipicamente referem... as teorias aceitas na ciência madura são aproximadamente verdadeiras (PUTNAM, 1975a, p. 73).

O AM é uma instância da IME, pois segundo ele: a melhor explicação para o sucesso das teorias científicas é que elas estão aproximadamente certas no que expressam

¹⁴ A noção de *verdade aproximada* permite que as teorias estejam equivocadas em alguma medida, no sentido de oferecerem descrições aproximadas do mundo. Mas mesmo assim essa noção carrega um comprometimento substancial com a ideia de que as teorias acertam consideravelmente no que expressam sobre os inobserváveis. Adicionalmente, essa noção implica que em casos de substituição teórica uma teoria posterior seria mais verdadeira do que uma anterior a ela (acerca do mesmo domínio de fenômenos). Em síntese, a noção de verdade aproximada não tem um caráter absoluto, uma vez que tolera a possibilidade de erro teórico e implica uma comparação entre teorias do mesmo domínio de fenômenos. Por carecer de uma definição precisa, tal noção é comumente criticada por antirrealistas — uma ótima fonte de informações e referências podem ser encontradas em Chakravartty (2017, seção 3.4).

sobre os inobserváveis. Assim, segundo Putnam ou concordamos com a explicação realista ou precisaremos admitir que tal sucesso é fruto de um “milagre” (também metaforicamente, claro). Com o auxílio desse argumento, o realismo visa a defender que as teorias científicas são aproximadamente verdadeiras e que os inobserváveis referidos por elas existem.¹⁵

Há filósofos, como Bas van Fraassen, que questionam essa explicação realista do sucesso científico. Segundo ele, o sucesso das teorias não deveria nos espantar e nem carece de uma explicação milagrosa, uma vez que só teorias bem-sucedidas sobrevivem na ciência, num processo de permanente filtragem e seleção:

... o sucesso das teorias científicas atuais não é nenhum milagre. Nem sequer é surpreendente para a mente científica (darwinista). Toda teoria científica nasce em uma vida de competição feroz, uma selva de dentes e garras ensanguentadas. Apenas as teorias bem-sucedidas sobrevivem — aquelas que *de fato* agarram as reais regularidades da natureza. (VAN FRAASSEN, 1980, p. 40).

O proponente do empirismo construtivo oferece, assim, o que parece ser uma explicação alternativa para o sucesso da ciência: as teorias usufruem de sucesso porque elas são desenvolvidas para serem assim e sobreviverem à disputa teórica. Tal explicação é denominada *seletiva* ou *darwinista*. A vantagem da explicação de van Fraassen é que ele pode incorporar tanto o sucesso das teorias que atualmente são consideradas verdadeiras como as que foram abandonadas e então classificadas como falsas. Entretanto, uma crítica remanescente à sua explicação é que ele não especifica o fator inerente a cada teoria que gera seu sucesso, enquanto o realista sempre remonta o sucesso de cada teoria à verdade aproximada dela, isto é, aos fatores que contribuem para a sua verdade aproximada.¹⁶

Pelo menos dois problemas iniciais são enfrentados pelo AM: a circularidade (FINE, 1984) e a falácia de taxa base (HOWSON, 2000, cap. 3; MAGNUS; CALLENDER, 2004; DE SOUZA, 2018). A circularidade afeta o AM porque tal argumento aparentemente pressupõe o mesmo tipo de raciocínio que já é questionado no nível da escolha de teorias que postulam entidades inobserváveis; a falácia de taxa base indica que se não soubermos qual é a chance de uma teoria qualquer ser, a priori, aproximadamente verdadeira, torna-se difícil sustentar a argumentação do AM de que o sucesso é um fator confiável para defender o acerto da teoria.

¹⁵ A Boyd (1984) também é atribuída a formulação do AM. Chibeni (1996; 2006) apresenta as versões mencionadas do AM, traça paralelos e distinções entre eles.

¹⁶ Uma exposição dessas críticas encontra-se em Lipton (2004, pp. 193-194), Chibeni (2006, seção 3.3) e Malavolta e Silva (2020, p. 26). Wray (2018, cap. 10) tenta responder a essa e outras críticas apresentadas à explicação darwinista de van Fraassen.

Dado que o AM é uma instância da IME, um dos primeiros problemas colocados a esse argumento é o questionamento do próprio tipo de raciocínio usado nele. Como indicado na apresentação da IME e da subdeterminação, o uso da IME para a escolha de teorias não está isento de críticas; desse modo, como o realista pode querer defender sua posição através do mesmo raciocínio que já é criticado no nível da escolha de teorias que postulam entidades inobserváveis? Parece que o realista sempre está assumindo (sem provas) a confiabilidade da IME seja para se comprometer com teorias que presumem a existência de inobserváveis quanto para defender que ele tem a melhor explicação/teoria filosófica para dar conta do aspecto bem-sucedido da ciência.

O cientista infere que sua melhor explicação para certos fenômenos naturais é provavelmente verdadeira, mas o filósofo anti-realista questiona essa inferência; o filósofo realista infere que o realismo científico é a melhor explicação para o sucesso da ciência, e o anti-realista o acusa de petição de princípio, por pressupor a forma de inferência em disputa. (CHIBENI, 2006, p. 230).

Além disso, explicações alternativas para o sucesso da ciência podem subdeterminar a escolha pela explicação realista: basta lembrar que embora a explicação proposta por van Fraassen não especifique o fator particular responsável pelo sucesso de cada teoria, ele consegue dar sentido ao sucesso de teorias que foram falseadas, algo que falta à explicação realista presente no AM. Fato é que a IME é falível; a questão é o quão falho esse raciocínio pode ser e ainda assim oferecer apoio para o realismo científico. Psillos (1999, pp. 79-87) questiona a objeção de circularidade, e Chibeni (2006, pp. 230-231) a afasta argumentando que as explicações científicas e filosóficas diferem, de modo que a explicação filosófica (nesse caso, o AM) não pressupõe a confiabilidade da IME no nível da aplicação científica.¹⁷

Por sua vez, a falácia de taxa base está relacionada com o uso de conceitos de probabilidade, que são relevantes para guiar o nosso julgamento sobre a chance de uma teoria de sucesso ser aproximadamente verdadeira ou falsa.

Howson (2000: 55-57, 2013) defende que a conclusão de AMH [do AM] apenas decorre de suas premissas ao assumirmos um valor não-negligenciável... para a probabilidade prévia de que *T* é aproximadamente verdadeira: AMH [o AM] comete, assim, a falácia de taxa-base. Em outras palavras, AMH [o AM] apenas consegue sustentar sua conclusão se já pressupõe uma probabilidade não muito baixa para a verdade aproximada da teoria *T*. (DE SOUZA, 2018, p. 51).

Consideremos a seguinte ilustração para esclarecer esse problema:

Há um teste para uma doença para o qual a taxa de falsos negativos (resultados negativos nos casos em que a doença está presente) é zero, e a taxa de falsos positivos (resultados positivos nos casos em que a doença está ausente) é um em dez (isto é, indivíduos sem doença são diagnosticados com positivo 10% das vezes). Se

¹⁷ Em Dicken (2016, pp. 71-74) há maiores informações e referências sobre a circularidade do AM.

alguém é diagnosticado com positivo, quais são as chances de ter a doença? Seria um erro concluir que, baseado na taxa de falsos positivos, a probabilidade seja de 90%, pois a real probabilidade depende de uma informação crucial adicional: a taxa de base da doença na população (a proporção das pessoas que a possuem). Quanto mais baixa a incidência da doença em geral, menor a probabilidade de um resultado positivo assinalar a presença da doença.

Por analogia, usar o sucesso de uma teoria científica como indicador de sua verdade aproximada (assumindo uma baixa taxa de falsos positivos — casos nos quais teorias afastadas da verdade são, todavia, bem-sucedidas) é possivelmente, da mesma forma, uma instância da falácia da taxa de base. O sucesso de uma teoria não sugere por si só que ela provavelmente seja aproximadamente verdadeira, e dado que não há meio independente de saber a taxa de base de teorias aproximadamente verdadeiras, as chances de ela ser aproximadamente verdadeira não podem ser avaliadas. (CHAKRAVARTTY, 2017, n.p).

Se não soubermos a taxa base de teorias arbitrárias serem aproximadamente verdadeiras (i.e., a chance de uma teoria qualquer (bem-sucedida ou não) ser aproximadamente verdadeira), não seremos capazes de determinar se o sucesso é um fator confiável para indicar o acerto da teoria a fim de construirmos alguma generalização acerca disso. De nada adianta considerarmos algumas teorias bem-sucedidas atualmente classificadas como aproximadamente verdadeiras se ignorarmos a distribuição geral de verdade aproximada entre as teorias científicas já propostas.¹⁸

Não nos estenderemos sobre esses dois problemas, por estarmos mais interessados em explorar neste trabalho as complicações *históricas* colocadas ao AM que motivam o recurso às teorias da referência. O tipo de realismo a ser discutido por nós é aquele que depende de alguma versão do AM, pois visamos defender o conhecimento sobre os inobserváveis recorrendo ao sucesso das teorias científicas. Em especial, nos interessa o realismo *explanacionista*, segundo o qual devemos nos comprometer epistemicamente com os pressupostos das teorias científicas que explicam seu sucesso teórico.¹⁹

¹⁸ De acordo com essa crítica, ao examinar o AM sob a perspectiva da probabilidade, o sucesso de uma teoria só pode ser considerado um indicador de sua verdade aproximada se assumirmos inicialmente uma probabilidade significativa, ou seja, não muito baixa, para a verdade aproximada da teoria. O problema surge pelo fato de o realista não oferecer uma justificativa adequada para essa probabilidade inicial de a teoria ser aproximadamente verdadeira, pressupondo então um valor que deveria ser devidamente justificado pela distribuição de verdade aproximada entre as teorias científicas. Desse modo, a constante relação entre sucesso e verdade aproximada das teorias é obtida de forma enviesada: por exemplo, o sucesso pode ser resultante de um acaso, mas se já assumirmos desde o início que é bem provável uma teoria científica ser aproximadamente verdadeira, adotaremos mais facilmente a verdade aproximada como explicação para o sucesso. Maiores informações e referências sobre como a falácia de taxa base ameaça o AM encontram-se em Dicken (2016, pp. 78-81; 120-124).

¹⁹ Entendemos por realismo explanacionista aquele tipo de posição que defende o AM ou que, assim como o AM, enfatiza a importância da explicação do sucesso da ciência e faz dessa explicação uma justificativa explícita a favor dos inobserváveis. Mesmo que fique apenas subtendido, os proponentes do realismo de entidades ou estrutural defendem que o sucesso da ciência é devido à existência real das entidades ou estruturas representadas pelas nossas teorias, mas eles tentam se afastar do enfoque explicativo do AM, priorizando aquilo que temos justificção para acreditar que existe — tanto que suas justificções a favor dos inobserváveis não consistem em explicações do sucesso da ciência, mas procedem através da manipulação instrumental ou equações das teorias, por exemplo.

Novamente, o recurso à explicação por parte do realista pode parecer inusitado no presente trabalho, dado que rejeitamos desde o início o poder explicativo como um fator constitutivo do sucesso da ciência. A começar pela consideração de uma melhor explicação (sendo essa caracterizada por sua *consistência e coerência, simplicidade, escopo e unidade*) como um fator relevante na etapa de escolha entre hipóteses científicas, isto não implica que devemos distinguir o sucesso da ciência pelo seu potencial explicativo, uma vez que a própria noção de boa explicação é questionada e outras hipóteses (que não científicas) também podem explicar fenômenos. Já o AM e o realismo explanacionista comumente pressupõem uma definição de sucesso científico que não contempla o fator explicativo, enquanto visam explicar este sucesso pressupondo a existência de entidades inobserváveis — seja este sucesso entendido de maneira geral (relativamente a todas as teorias científicas, como faz o AM) ou de maneira particular (quanto a uma teoria científica, como se propõe o realismo explanacionista). Em suma, a noção de explicação sempre acaba sendo evocada de alguma forma na caracterização do sucesso científico, na melhor explicação, no AM e realismo explanacionista; contudo, no caso do sucesso científico, que é um estágio distinto de análise no debate, extraímos uma conclusão diferente sobre a relevância da explicação.

2.2. A subdeterminação da teoria pela evidência empírica

A *subdeterminação* da teoria pelos dados da experiência (SUB) é um problema que pode ocorrer em circunstâncias de escolha teórica. Em tal cenário, parece não haver condições puramente empíricas que permitam escolher qual é a melhor teoria. Se duas teorias diferentes forem empiricamente equivalentes (implicarem os mesmos fenômenos), há casos em que elas podem ser inconsistentes entre si embora concordem sobre o mesmo conjunto de fenômenos. O problema é que apesar de oferecerem explicações diferentes para tais fenômenos, ambas teorias são igualmente confirmadas por eles, de modo que a própria experiência não nos permite saber qual delas é a mais acurada acerca do mundo.

A SUB é um corolário da “tese de Duhem-Quine”, que foi assim denominada em razão de ambos os filósofos, Pierre Duhem e Willard Quine, chamarem a atenção para o fato de que as teorias científicas não são testadas ou confirmadas isoladamente, dependendo para tal de assunções auxiliares, hipóteses de fundo, etc., de modo que qualquer teoria pode se tornar compatível com a experiência se for devidamente incorporada a esse pano de fundo

teórico.²⁰ Essa tese sugere que qualquer teoria pode se adequar à experiência com o auxílio de suposições adicionais apropriadamente concebidas. Assim, sempre é logicamente possível tornar uma teoria tão empiricamente adequada quanto qualquer outra já proposta.

Enquanto uma mera possibilidade lógica de empate entre teorias, a SUB não é considerada um grande problema. Mas caso seja possível que, na ciência real, uma teoria torne-se empiricamente equivalente à outra — dada a possibilidade de refinamento teórico —, a SUB vislumbra-se como um problema cada vez mais concreto.

A SUB é, portanto, uma ameaça potencial à IME. Esse problema evidencia que nem sempre podemos realizar inferências substanciais sobre o mundo com base na experiência apenas, pois ela não é um juiz absoluto que garanta objetivamente a escolha da melhor explicação em todos os casos. Suponha a existência de duas teorias divergentes quanto ao âmbito inobservável do mundo (seja por postularem diferentes entidades, caracterizá-los de outra maneira ou por uma delas nem sequer supor a existência de um inobservável) e que explicam os mesmos fenômenos, estando igualmente de acordo com eles. Nessas condições não há base para escolher qual teoria é a mais correta sobre o âmbito inobservável, já que a própria experiência não nos aponta qual delas é verdadeira *in toto* (e não apenas quanto ao que é observável). Não há como decidir empiricamente qual é a melhor explicação para os fenômenos em questão e, conseqüentemente, não estaríamos habilitados a inferir objetivamente qual teoria seria a mais correta sobre o reino do inobservável. Já frisamos que a SUB é um problema potencial inerente a qualquer tentativa de explicação da experiência; todavia, como a IME está sempre presente na escolha de teorias que pressupõem a existência de inobserváveis, as conseqüências da SUB são graves para um realista científico.

Talvez esse empate entre as teorias possa ser solucionado com o tempo, à medida que mais dados empíricos forem coletados. No entanto, a possibilidade lógica do surgimento de teorias que equivalem empiricamente a outras já concebidas é um entrave para que tal atitude seja tomada perante o problema da SUB. Uma outra saída ao problema é a busca de outros fatores, além do acordo com a experiência, que sejam capazes de desempatar efetivamente a disputa teórica. Pois bem, comumente os realistas alegam que de fato há outros fatores além da experiência envolvidos na escolha entre teorias. Desse modo, embora duas teorias sejam equivalentes quanto à experiência, não quer dizer que elas equivalem igualmente no nível do conjunto das evidências. As mesmas virtudes teóricas que auxiliam na

²⁰ Duhem (1906, cap. 6) sugere que quando um experimento falha ele não indica qual hipótese do corpo teórico deve ser abandonada. E Quine (1951; 1975) através do conceito de “holismo confirmacional” alerta que a experiência não confirma ou desconfirma crenças individuais, mas sim todo um conjunto de crenças.

escolha entre teorias explicativas — *consistência* e *coerência* (com teorias de fundo), *simplicidade*, *escopo* e *unidade*, etc — constituiriam virtudes epistêmicas que figurariam como evidências capazes de desempatar a escolha teórica. No entanto, tais virtudes podem não ser reconhecidas como epistêmicas pelo antirrealista, de modo a justificar que uma dada teoria nos forneceria conhecimento sobre o mundo e então desempatarasse plausivelmente a disputa em questão. Por exemplo, será que a simplicidade é um bom guia para a verdade? Assim, algumas virtudes teóricas podem ser meros aspectos que facilitam o trabalho científico, apresentando um mero caráter pragmático ao invés de epistêmico.²¹ Além do mais, não é óbvio qual é o peso de cada uma dessas virtudes teóricas e nem como elas se aplicam às teorias científicas, por exemplo, uma teoria pode ser mais simples tanto do ponto de vista ontológico como matemático; então como ficaria a escolha em caso de empate quanto à simplicidade teórica de dois tipos diferentes?

A SUB é uma possibilidade incômoda que ameaça qualquer inferência realista advinda da IME. Caso seja concretizada, o sucesso empírico de uma teoria não seria suficiente para ela ser escolhida como a melhor explicação em um dado caso *e então ser considerada como aproximadamente verdadeira*.²² Contudo, a SUB tem muito mais um caráter de possibilidade do que algo concreto.²³ Então mesmo que a SUB seja uma das motivações para o antirrealismo científico, optaremos por explorar as motivações históricas, que também descredita a inferência do aspecto bem-sucedido da teoria para sua verdade aproximada, por serem uma perspectiva mais concreta de desafio ao realismo científico. Nessa rota devemos lidar com outra poderosa objeção ao realismo científico, a chamada “metaindução pessimista”.

2.3. A metaindução pessimista

A história da ciência não consiste, se olhada a longo prazo, apenas em teorias bem-sucedidas e verdadeiras: muitas das que em uma época foram tidas como bem-sucedidas,

²¹ Essa é a posição explicitamente adotada por van Fraassen. Uma defesa famosa do peso epistêmico desses fatores é Churchland (1985, pp. 35-47). Uma análise mais recente pode ser encontrada em Wray (2018, cap. 8).

²² Malavolta e Silva (2020) indica que o problema da SUB pode motivar a atribuição de uma baixa probabilidade à chance de uma teoria bem-sucedida ser verdadeira. Como consequência, a SUB também motiva a postura antirrealista que diminui a importância de uma explicação para o sucesso da ciência ou constrói essa explicação em termos *darwinistas* e até de adequação empírica das teorias, como em van Fraassen (1980).

²³ Esse é um ponto polêmico. Há autores que defendem que há, sim, casos históricos concretos de equivalência empírica na ciência, em que o problema da SUB se patenteia. Para um exemplo, ver o tratamento extenso encontrado em Chibeni (1997) acerca da equivalência empírica entre a mecânica quântica não-relativista e a teoria de variáveis ocultas de Bohm (em uma de suas modulações teóricas).

e, portanto, verdadeiras, posteriormente passaram a ser classificadas como falsas. Esse tipo de consideração histórica está presente no argumento antirrealista intitulado *metaindução pessimista* (MIP), cuja mais poderosa e influente formulação é devida a Larry Laudan, que a apresentou em 1981 no artigo “A confutation of convergent realism” (LAUDAN, 1981).

A MIP é um argumento que consiste na análise histórica das teorias científicas que pressupõem a existência de entidades inobserváveis; no entanto, de maneira oposta ao AM, não leva a conclusões otimistas sobre a capacidade da ciência descobrir a verdade sobre o mundo inobservável. Muitas formulações da MIP possuem a estrutura de um argumento indutivo (POINCARÉ, 1905, p. 122; PUTNAM, 1978, pp. 22-25): se as teorias científicas do passado foram falseadas, então o mesmo infortúnio atingirá as nossas teorias atuais, não recomendando-se assim uma postura realista relativamente a elas. Já que não podemos inferir de modo garantido a falha teórica futura a partir da enumeração de falhas passadas, temos uma argumentação indutiva que não nos garante a verdade da sua conclusão. O caráter indutivo desses argumentos motiva críticas à conclusão pessimista presente neles.²⁴ Já o questionamento apresentado por Laudan é um *desafio cético*, representando um ataque à conexão, pressuposta pelo realista, entre sucesso, referência e verdade aproximada. Como vimos, essa conexão é crucial para a legitimidade do AM. O arrazoado de Laudan mostraria que uma análise histórica da ciência não parece apoiar tal conexão.²⁵

A partir de uma análise histórica da ciência, Laudan chama a atenção para o fato de que, contrariamente ao que o realista argumenta, o sucesso das teorias científicas comumente foi acompanhado por falhas. De fato, há vários casos históricos da ciência nos quais teorias bem-sucedidas que pressupunham a existência de inobserváveis foram posteriormente classificadas como falsas, de modo que não só as descrições dessas teorias sobre os inobserváveis foram consideradas pouco acuradas como os entes inobserváveis postulados por elas tiveram sua existência negada.

Laudan visa, portanto, abalar a conexão pressuposta pelo realista entre os fatores de *sucesso, referência e verdade aproximada das teorias*. Somente se esses aspectos

²⁴ Um exemplo de crítica é a falácia de taxa base (LEWIS, 2001), a qual também atinge o AM. De modo análogo ao problema colocado ao AM, no caso da MIP indutiva, se não soubermos a taxa base de teorias arbitrárias serem falsas, não seremos capazes de determinar o quanto o sucesso já esteve acompanhado da falsidade a fim de generalizar sobre isso. É preciso considerar a distribuição geral de sucesso e falsidade entre as teorias científicas já propostas. Maiores informações e referências sobre essa crítica podem ser encontradas em Dicken (2016, pp. 120-125). Como não nos ocupamos da MIP indutiva, não exploraremos essa crítica aqui.

²⁵ Wray (2018, cap. 5) apresenta competentemente algumas versões da MIP e Oliveira (2014) expõe maiores detalhes sobre os problemas inerentes às versões indutivas da MIP e inclusive defende que o questionamento antirrealista de Laudan é um desafio cético — dado que não é possível afastar a possibilidade de colapso teórico futuro (como o descarte de entidades pressupostas pelas teorias) não sabemos se as teorias atuais são verdadeiras e/ou se as entidades por elas supostas existem.

estiverem comumente associados às teorias é que o realista estará habilitado a inferir do fato de que uma teoria é bem-sucedida que ela também acerta na referência ao inobservável e é aproximadamente verdadeira. Tal versão da MIP tende a promover a incerteza sobre a conexão pressuposta pelo realista, deslegitimando sua explicação de que o sucesso das teorias científicas é devido a sua verdade aproximada.

Além da MIP, o *problema das alternativas não concebidas*, formulado por Kyle Stanford (2001; 2006), apresenta considerações históricas capazes de minar as convicções realistas expostas no AM.

Ao longo da história da investigação científica e em praticamente todos os campos científicos, ocupamos repetidamente uma posição epistêmica na qual só podíamos conceber apenas uma de algumas teorias que eram bem confirmadas pela evidência disponível, enquanto a investigação subsequente revelava rotineiramente (senão invariavelmente) outras alternativas radicalmente distintas tão bem confirmadas pela evidência previamente disponível como aquelas que estamos inclinados a aceitar pela força da evidência. (STANFORD, 2006, p. 19).

Esse outro problema histórico é, no fundo, uma combinação de SUB com MIP. Ao longo da história da ciência, teorias bem-sucedidas foram abandonadas em detrimento de outras consideradas mais acuradas (MIP), tais episódios de substituição teórica podem ser uma constante na ciência, a qual ocorre devido à disponibilidade posterior de melhores explicações para um domínio de fenômenos. Dado que nem sempre os cientistas formulam as melhores explicações em uma dada ocasião de escolha teórica, sempre há a possibilidade de formulação futura de uma teoria que seja uma melhor explicação que a anterior sobre um domínio do mundo (SUB), assim, as teorias já propostas podem ser comumente substituídas por outras melhores, de modo que o sucesso de uma teoria não bloquearia a sua falha.²⁶ O problema levantado por Stanford tem um caráter menos hipotético do que a SUB, porém, ainda não é tão concreto como a MIP porque alude a possibilidades futuras de substituições teóricas. Por ser um problema mais concreto, nosso foco neste trabalho será a MIP, de modo que não nos ocuparemos dessa outra consideração histórica não tão concreta assim.

No artigo de 1981, em que Laudan apresentou sua versão da MIP, ele formulou a seguinte lista de doze teorias que violariam as convicções realistas expostas no AM:

- as esferas cristalinas da astronomia antiga e medieval;
- a teoria dos humores da medicina;
- a teoria dos eflúvios da eletricidade estática;
- a geologia ‘catastrofista’, com seu comprometimento com o dilúvio universal (de Noé);
- a teoria do flogisto da química;
- a teoria do calórico;
- a teoria vibratória do calor;
- as teorias da força vital da fisiologia;

²⁶ Em Dicken (2016, pp. 109-115) há uma boa exposição desse problema e de respostas realistas a ele.

- o éter eletromagnético;
- o éter óptico;
- a teoria da inércia circular;
- teorias da geração espontânea. (LAUDAN, 1981, p. 33).

Laudan acrescenta que tal lista

[...] poderia ser estendida *ad nauseam*, envolvendo em cada caso uma teoria que foi bem-sucedida e confirmada, mas cujos termos centrais (tal como acreditamos agora) são não referentes [i.e., termos que dizem respeito aos inobserváveis]. Quem imagina que as teorias bem-sucedidas na história da ciência foram também genuinamente referentes em seus conceitos [termos] centrais, estudou apenas as versões mais liberais da história da ciência (i.e., aquelas que consideram apenas as teorias passadas que são referencialmente semelhantes às atuais). (LAUDAN, 1981, pp. 33-34).

Se algum desses casos de falha teórica for um contraexemplo legítimo à argumentação realista, ela só se sustentará se ignorarmos esse fato, ou se lhe dermos outra explicação.

Os casos históricos de teorias que falharam na referência ao inobservável são diretamente acionados para questionar a verdade aproximada das teorias, porque o realista parece pressupor a referência como um pré-requisito para a verdade. Tanto que esse aspecto é mencionado na formulação clássica do AM veiculada por Hilary Putnam e é devidamente reforçado por Laudan,

um realista jamais gostaria de dizer que uma teoria era aproximadamente verdadeira caso seus termos teóricos centrais falhassem na referência. Se não houver algo como os genes, então a teoria genética, não importa o quão bem confirmada ela seja, não seria aproximadamente verdadeira. Se não houvesse entidades similares aos átomos, nenhuma teoria atômica poderia ser aproximadamente verdadeira; se não houvesse partículas subatômicas, então nenhuma teoria quântica da química poderia ser aproximadamente verdadeira. Em suma, uma condição necessária — especialmente para um realista científico — para que uma teoria esteja próxima da verdade é que seus termos explanatórios centrais genuinamente refiram. (LAUDAN, 1981, p. 33).

E o problema histórico levantado pela MIP não depende de alguma versão específica de antirrealismo científico. Os casos históricos de falha teórica colocam em xeque a inferência realista do sucesso para a verdade, mas a apresentação desse desafio não depende de alguma tese específica de uma determinada posição antirrealista.

Aparentemente, portanto, a MIP abala seriamente o comprometimento epistêmico do realismo científico tal como exposto através do AM. O sucesso teórico já não seria mais uma boa razão para argumentar que as teorias nos fornecem conhecimento sobre a parte inobservável do mundo; como consequência, a verdade aproximada das teorias e o seu acerto na referência ao inobservável tornar-se-iam infundados.

O que analisaremos no prosseguimento deste trabalho é se, com o auxílio das teorias da referência, as teorias científicas podem, em que pese a evidência histórica manipulada pela metaindução pessimista, ser consideradas aproximadamente verdadeiras e referentes a fim de reabilitar o comprometimento epistêmico do realismo científico. Para tanto, averiguaremos se é possível defender a manutenção de referência entre as teorias. Há então neste trabalho um componente semântico. Mas esse componente não implica que estejamos nos ocupando da dimensão semântica de comprometimento do debate, pois ela diz respeito à interpretação literal do discurso teórico, que é algo diferente da questão da manutenção de referência entre as teorias, já que podemos interpretar literalmente um discurso e ainda assim admitir sua falha. Com isso queremos deixar claro que este trabalho se ocupa da dimensão epistêmica do debate, enquanto discorre sobre aspectos semânticos que não esgotam a dimensão semântica da discussão.

2.4. O problema da descontinuidade teórica na ciência

Os casos históricos representativos da MIP figuram em episódios de substituição teórica na ciência como aquelas teorias que foram abandonadas em favor de outras teorias. Além de serem contraexemplos às pressuposições realistas expostas no AM, tais casos históricos são uma evidência do alto grau de descontinuidade teórica na ciência. Como esses dois problemas estão estreitamente relacionados, caso o realista tencione defender suas convicções exibidas no AM ele precisará se ocupar da aparente descontinuidade teórica. Em uma análise inicial, a ocorrência de uma substituição teórica seria devida à disponibilidade de uma teoria mais acurada, abrangente, heurísticamente fértil, coerente, etc. sobre um certo domínio do mundo, que estimularia o abandono de uma teoria aceita, uma vez que há uma melhor alternativa disponível para nos informar sobre um certo conjunto de fenômenos. Quanto aos casos de teorias que pressupõem a existência de inobserváveis, é comum que a substituição teórica também inclua o abandono da referência a essas entidades, fazendo com que uma teoria posterior não compartilhe com a teoria anterior da mesma referência a um dado inobservável — seja por um mero refinamento da teoria anteriormente aceita ou pela sua total rejeição. Em decorrência dessa substituição, a teoria abandonada passa a ser classificada como falsa e a teoria posteriormente aceita é considerada verdadeira, ou pelo menos mais próxima de ser verdadeira.

Fato é que em episódios de substituição teórica ambas as teorias em questão são bem-sucedidas num determinado momento histórico — apesar da teoria posterior ser mais

bem-sucedida e assim ocupar o lugar da teoria anterior —, de modo que tanto a teoria aceita quanto a abandonada cumpririam o pré-requisito realista para classificar uma teoria como aproximadamente verdadeira e referente ao inobservável. Caso as substituições teóricas sejam constituídas por um alto grau de descontinuidade teórica (sobretudo no que concerne à referência aos inobserváveis) aquilo que é defendido pelo realismo científico através do AM não procede, já que esses episódios mostrariam que o fator de sucesso de uma teoria não nos permite inferir sua verdade aproximada e acerto na referência ao inobservável. Por essa razão, os realistas precisam lidar adequadamente com os casos históricos de substituição teórica na ciência a fim de salvaguardar o AM e continuar defendendo que as teorias são bem-sucedidas por serem aproximadamente verdadeiras e acertarem na referência às entidades inobserváveis.

A partir do diagnóstico da MIP, algumas estratégias foram adotadas pelos realistas científicos a fim de enfrentar adequadamente o problema da aparente descontinuidade exibida pelos casos históricos:

- a) alegar que a noção de sucesso teórico é mais refinada do que a considerada na formulação da MIP, de modo a diminuir a quantidade de casos históricos que descreditam o AM e enfraquecer o desafio cético lançado contra os realistas. Outros critérios mais exigentes para o sucesso — além dos comumente apresentados como adequação empírica, poder preditivo e unificador — constituiriam teorias genuinamente bem-sucedidas a serem avaliadas nessa discussão histórica (HARDIN E ROSENBERG, 1982; MUSGRAVE, 1988; WORRALL, 1989; PSILLOS, 1999; CHIBENI, 2006).
- b) a posição do realismo explanacionista de só sustentar o realismo quanto àqueles pressupostos das teorias científicas que cumpriram um papel importante na geração do seu sucesso e, portanto, o explicam. Somente quando a postulação de um dado inobservável foi crucial, indispensável, para o sucesso de uma teoria abandonada é que ela constitui um sério desafio histórico. Nesses casos, o realista tentará defender que os aspectos necessários para o sucesso das teorias abandonadas foram mantidos nas teorias atuais, de modo a argumentar que as pressuposições relevantes sobre o âmbito inobservável do mundo não falham, salvaguardando o acerto das teorias abandonadas e explicando efetivamente o aspecto bem-sucedido da ciência (KITCHER, 1993; PSILLOS, 1999);
- c) defender alguma forma de manutenção de referência a certos inobserváveis entre as teorias abandonadas e as atuais. Considerando a referência como um pré-requisito para a verdade aproximada, se uma teoria abandonada acerta na

referência ela pode ser classificada como aproximadamente verdadeira. As teorias atuais são mais bem apoiadas pela evidência do que as anteriores, sendo então consideradas verdadeiras; assim, qualquer teoria que pressupusesse a existência das mesmas entidades presentes nas teorias atuais também acertaria quanto à referência, satisfazendo o pré-requisito para ser classificada como aproximadamente verdadeira (HARDIN E ROSENBERG, 1982; KITCHER, 1993; PSILLOS, 1999);

- d) Dissociar as noções de “verdade aproximada” e “referência” para que a segunda não seja um pré-requisito da primeira. Ainda que uma teoria abandonada não compartilhe da mesma referência de alguma teoria atual, se algum outro aspecto da teoria foi mantido entre elas — como o significado dos enunciados das teorias presentes em suas descrições —, então é possível argumentar que a teoria abandonada acertou, em alguma medida, acerca do mundo, embora tal acerto não tenha sido em termos da referência ao inobservável (HARDIN E ROSENBERG, 1982; PAPINEAU, 2010).

A estratégia (a) foi parte de um movimento inicial dos realistas para deslegitimar os casos históricos inicialmente apresentados na formulação da MIP de Laudan, de modo que eles não constituíssem um sério desafio ao que é exposto no AM. Desde que foram propostos, os critérios que refinam a noção de sucesso teórico são usados para delimitar quais casos históricos são de fato uma ameaça ao realismo. A estratégia (b) é usualmente acompanhada pela (c) — como em Kitcher (1993) e Psillos (1999) —, pois a manutenção daquilo que é indispensável para o sucesso teórico remonta quase inevitavelmente à manutenção de referência entre as teorias científicas. Geralmente a adoção de (b) delimita ainda mais quais casos históricos constituem o tipo de descontinuidade teórica que descredita o AM — a saber, aqueles em que a pressuposição de um dado inobservável foi indispensável para o sucesso da teoria abandonada —, auxiliando a descartar os casos “inócuos” e a identificar quais deles são dignos de preocupação realista. Então (c) é acionada quando o diagnóstico daquilo que é indispensável para o sucesso teórico recai sobre a postulação de um inobservável abandonado, tornando necessária a defesa da manutenção de referência entre as teorias para resguardar que o fator responsável pelo sucesso de uma teoria abandonada ainda permanece em uma teoria atual de modo a classificá-la como aproximadamente verdadeira e certa quanto à referência, salvaguardando então a argumentação realista exposta no AM. A estratégia (d) — como em Hardin e Rosenberg (1982) e Papineau (2010) — é uma outra possível saída até para quem adota (b), mas usualmente ela não é tão explorada, porque dispensa uma característica cara ao

realismo: o acerto na referência a pelo menos algum inobservável. Sustentar a verdade aproximada das teorias em termos das descrições apresentadas acerca do mundo inobservável parece requerer a existência da entidade postulada por uma teoria, e esse é um aspecto substancial do qual o realista nem sempre está disposto a ceder. As estratégias (c) e (d) presumem algum tipo de continuidade entre as teorias para a defesa da verdade aproximada das teorias abandonadas, que depende de uma análise mais detalhada dos aspectos relacionados ao componente linguístico das teorias científicas — como a referência e o significado, que são trabalhados na área da filosofia da linguagem, especialmente com as teorias da referência.²⁷

O intento deste trabalho, a partir de agora, é avaliar a possibilidade de defesa da manutenção de referência entre as teorias — conforme a estratégia (c) —, a fim de sustentar que o sucesso das teorias científicas é, sim, devido à sua verdade aproximada e acerto na referência ao inobservável. Dado isso, não exploraremos a estratégia (d), uma vez que ela visa a defender a continuidade entre as teorias abdicando da manutenção da referência. Avaliaremos, a partir da estratégia (c), se as teorias abandonadas podem compartilhar da mesma referência que as teorias atuais presentes na ciência, para que de maneira oposta ao que aparentam ser, os casos históricos que seriamente desafiavam o realismo científico não constituiriam casos de descontinuidade teórica, mas sim de continuidade, em algum sentido a ser especificado. Assim, as substituições teóricas passariam a ser consideradas pelos realistas como um processo natural de aperfeiçoamento da ciência,²⁸ ao invés de uma ameaça às suas convicções quanto aos inobserváveis. Para os nossos intentos, examinaremos como as teorias da referência poderiam auxiliar-nos, uma vez que é no nível linguístico que poderemos analisar adequadamente a possibilidade de continuidade teórica na ciência. Eventualmente, a estratégia (b) também pode ser adotada no prosseguimento do trabalho para garantir a efetividade da estratégia (c), pois identificar o que é central na geração do sucesso das teorias

²⁷ Há outras posições mais robustas no debate como o realismo de entidades ou o realismo estrutural que também defendem a continuidade nos episódios de substituição teórica — seja relativamente à manutenção de referência das entidades ou estruturas —, mas que não se ocupam de maiores detalhes da filosofia da linguagem. De qualquer modo, atualmente as estratégias realistas demandam algum tipo de seletividade com relação às teorias. Esse tipo de atitude é bastante comum no debate contemporâneo, pois os realistas não ignoram os registros históricos de substituição teórica, voltando então a sua atenção para aqueles aspectos que parecem permanecer entre as teorias e, que, por isso, requereriam o nosso comprometimento epistêmico (ver Chakravartty, 2017, seção 2.3). Não abordaremos o realismo de entidades ou estrutural, pois nosso intuito é analisar algumas estratégias realistas que pressupõem as teorias da referência para lidar com o desafio cético apresentado pela MIP.

²⁸ Em Boyd (1984), por exemplo, as substituições teóricas são encaradas como processos de aperfeiçoamento da ciência, nos quais a própria metodologia científica seria ajustada pela realidade, melhorando cada vez mais.

ajuda a delimitar quais casos históricos exige o recurso às teorias da referência para a defesa da continuidade teórica.

Ainda neste capítulo, faremos uma breve exposição das estratégias (a) e (b); no quarto capítulo exploraremos os desdobramentos de (c) a partir da avaliação do desempenho das teorias da referência para a manutenção da referência entre teorias.

* * *

Inicialmente os realistas tentaram driblar a MIP apresentada por Laudan alegando que ele listou algumas teorias que não seriam dignas de considerações realistas, i.e., que não seriam genuinamente bem-sucedidas, afinal, os realistas defendem seu comprometimento *apenas* quanto às nossas *melhores* teorias científicas já concebidas. Nesse sentido, alguns realistas propuseram critérios adicionais para o refinamento da noção de sucesso teórico de modo a delimitar as melhores teorias científicas a serem consideradas nesta discussão, como anunciado na estratégia (a). Para uma teoria ser genuinamente bem-sucedida ela precisaria satisfazer os seguintes critérios: i) ser madura e não *ad hoc* (HARDIN E ROSENBERG, 1982, p. 609; WORRALL, 1989, pp. 113-114; PSILLOS, 1999, p. 103), além de ii) ter realizado previsões inéditas (MUSGRAVE, 1988; WORRALL, 1989, pp. 101-102, 113-116; PSILLOS, 1999, pp. 100-103). Uma teoria qualquer é madura se resistiu a diversos testes e persistiu por um tempo considerável, e não será *ad hoc* se não foi meramente construída para dar conta de algumas observações já conhecidas. A previsão inédita pode ser de dois tipos: *novidade temporal* (quando o fenômeno previsto pela teoria era desconhecido pelos cientistas até então) e *novidade de uso* (o fenômeno era conhecido na época em que foi previsto, mas não foi utilizado na elaboração da teoria; pela falta de expectativa que a teoria previsse tal fenômeno é surpreendente quando ela o faz). Um exemplo da novidade de uso foi a previsão realizada pela teoria da Relatividade Geral proposta por Einstein, segundo a qual a luz se curva em torno de objetos massivos.

Outra variante da estratégia (a) consiste em ressaltar que a lista de Laudan é heterogênea, sendo constituída por teorias com diferentes características e que foram propostas em contextos variados, de modo que conclusões distintas podem ser extraídas de cada caso histórico, incluindo-se o eventual descarte de vários deles (CHIBENI, 2006, p. 240).

Uma análise isenta parece recomendar a desqualificação de vários itens da lista como teorias claramente imaturas, ou mesmo meros fragmentos teóricos, que tiveram um sucesso bastante limitado, quer por seu escopo estreito, quer pela falta de acurácia das previsões, quer, finalmente, por seu conflito com evidência experimental e teórica disponível já em seu tempo. Se essa réplica puder sustentar-

se, teremos uma substancial redução da lista e uma suspeita adicional quanto à arriscada generalização que Laudan faz. (CHIBENI, 2006, p. 240).

Dessa forma, seria indispensável analisar as particularidades inerentes a cada caso a fim de reduzir a lista de Laudan aos casos históricos realmente pertinentes e extrair conclusões adequadas de cada um deles.

Porém, esta primeira estratégia não é capaz de exaurir as teorias dignas de consideração realista, uma vez que o próprio Laudan cita algumas teorias científicas que cumprem os requisitos adicionais típicos propostos pelos realistas, como as teorias do éter luminífero (na física) ou do flogisto (na química).²⁹ Quanto a esses casos históricos, podemos destacar as seguintes predições inéditas:

- a) Éter luminífero (ou óptico): a ocorrência de um ponto luminoso no centro de uma sombra, ao projetar-se um feixe de luz em um pequeno disco opaco (LYONS, 2002, p. 71);
- b) Flogisto: a existência de propriedades redutivas do hidrogênio, que é a capacidade do hidrogênio de doar elétrons em uma reação química, atuando então como agente redutor, transferindo seus elétrons para outro elemento químico (CARRIER, 1991, p. 29).

A suposição da existência do éter — entidade inobservável material que preencheria todo o espaço — remonta à Antiguidade e passou por diferentes reformulações ao longo do tempo. Aristóteles (384-322 a.C.), que rejeitava a noção de espaço vazio, pressupôs a existência do éter caracterizando-o como uma substância invisível e sutil presente nas partes do espaço denominadas “mundo supralunar” (onde estariam localizados os corpos celestes). No século XVII, René Descartes (1596-1650) pressupôs o éter para explicar a interação de corpos distantes entre si (negando assim a ação à distância), tal como Christiaan Huygens (1629-1695) para embasar a concepção ondulatória da luz. Em meados do século XIX, as pesquisas de Thomas Young (1773-1829) e Augustin Fresnel (1788-1827), levaram efetivamente à classificação da luz como um fenômeno de natureza ondulatória. Nessa teoria, o éter cumpria um papel explicativo importante relativamente à propagação da luz através do espaço aparentemente vazio, sendo então denominado éter luminífero. A concepção de que a

²⁹ A própria teoria do calórico (para o calor) também parece satisfazer esse critério, tal como ressaltado em Carrier, 1991; 1993. Ademais, em Lyons (2002) e Vickers (2013) temos a recuperação de outros casos históricos a partir do critério de predição inédita, os quais posteriormente requerem a análise por parte do realismo científico. Contudo, nessa breve exposição sobre o assunto seguiremos explorando o caso do éter e do flogisto, pois eles oferecem desdobramentos mais interessantes para a tentativa de defesa da manutenção de referência entre as teorias, que é o foco deste trabalho.

luz consistia em um fenômeno ondulatório parecia requerer aos físicos da época a existência de um meio que propiciasse a propagação da luz, uma vez que não se concebia que qualquer processo pudesse se disseminar no espaço vazio. Posteriormente, já no século XX, em virtude de diversos fatores experimentais e teóricos, a existência dessa entidade inobservável foi abandonada;³⁰ atualmente, a propagação da luz é explicada através dos campos eletromagnéticos (igualmente inobservável). Quanto ao flogisto, foi uma outra entidade inobservável postulada em teorias químicas anteriores à chamada “revolução química” do final do séc. XVIII (protagonizada por Antoine Lavoisier (1743-1794)) para explicar o fenômeno da combustão. Segundo a concepção teórica do flogisto, os corpos queimavam porque continham flogisto em sua composição, de modo que esse flogisto seria liberado na atmosfera no decorrer do processo de combustão. Essa concepção foi rejeitada por Lavoisier, que propôs que a combustão ocorre porque os corpos reagem com um outro tipo de entidade — igualmente inobservável, note-se — o oxigênio, presente na atmosfera. Com a ampla aceitação dessa nova teoria, a existência do flogisto foi abandonada.³¹

Outra via de resposta ao desafio cético da MIP consiste em propostas realistas explanacionistas (b) que visam sustentar o realismo somente quanto aqueles pressupostos das teorias científicas que foram de fato importantes para a geração do seu sucesso teórico (independentemente de eles serem específicos como entidades ou estruturas), como as de Kitcher (1993) e Psillos (1999).

Com a estratégia *divide et impera* (cujo significado é “dividir para conquistar”), Stathis Psillos argumenta que somente os itens essenciais para o sucesso das teorias científicas demandam o comprometimento realista e são eles que devem permanecer em episódios de substituição teórica.³² A ideia por trás dessa estratégia é que sempre devemos ser realistas, mas apenas com aqueles componentes da teoria que foram responsáveis pelo seu sucesso, separando assim o que é digno do nosso comprometimento ou não. No próprio AM, a argumentação realista a favor dos inobserváveis é motivada pelo fator de sucesso das teorias científicas, de modo que esse fator é o que também justifica o comprometimento com tal tipo

³⁰ Essa é a interpretação *standard* em textos de física, mas historiadores da ciência têm questionado sua fidelidade histórica. Maiores informações em Martins (2006), Whittaker (1910), Cantor (1981) e Navarro (2018).

³¹ Para mais detalhes acerca dessas teorias pode-se consultar Brito (2008); um estudo histórico aprofundado sobre o éter encontra-se em Psillos (1999, cap. 6), enquanto o caso do flogisto é exposto especialmente em Ladyman (2011).

³² A proposta de Phillip Kitcher é similar a de Psillos. Ele distingue entre os “postulados pressuposicionais” ou “peças inativas” das teorias dos seus “postulados funcionais”, sendo os últimos com os quais o realista deveria se comprometer por serem importantes para a derivação de predições inéditas (Kitcher, 1993, p.140-149). De modo auxiliar, Kitcher (1978; 1993) argumentou que os termos das teorias científicas têm *referentes potenciais*, de modo que o mesmo termo pode referir a entidades diferentes a depender do contexto de enunciação — críticas a essa concepção da referência dos termos científicos foram formuladas por Psillos (1997) e McLeish (2005).

de entidade. Quando a pressuposição de um inobservável não contribui efetivamente para o sucesso de uma teoria, não há razão para nos comprometermos com a sua existência, uma vez que ela não tem qualquer ligação com o fator que origina e justifica o comprometimento realista.

[...] é suficiente mostrar que as leis e os mecanismos teóricos que geraram o sucesso das teorias passadas foram retidas na nossa imagem científica atual. Chamarei a isto de movimento *divide et impera*. Ele é baseado na afirmação de que quando uma teoria é abandonada, seus constituintes teóricos, i.e., as leis e os mecanismos teóricos postulados por ela não precisam ser rejeitados *em bloco*. Alguns desses constituintes teóricos são inconsistentes com os que aceitamos agora e, então, devem ser rejeitados. Mas nem todos são assim. Alguns deles foram retidos como constituintes essenciais em teorias posteriores. O movimento *divide et impera* sugere que, se os constituintes teóricos responsáveis pelo sucesso empírico das teorias abandonadas forem aqueles retidos na nossa imagem científica atual, então uma versão substantiva do realismo científico ainda pode ser defendida. (PSILLOS, 1999, p. 103).

Pelas predições inéditas mencionadas, à primeira vista, parece que as teorias do éter luminífero e do flogisto dependiam em alguma medida da postulação desses inobserváveis para serem bem-sucedidas. Veremos mais adiante que, juntamente com o *divide et impera*, Psillos recorre ao auxílio de uma teoria da referência em especial para mostrar como os constituintes essenciais de uma teoria podem ser compartilhados com outras teorias ao longo do tempo, de modo a lidar com esses casos históricos remanescentes.

No entanto, Kyle Stanford (2003, pp. 563-572; 2006, pp. 155-159) é um filósofo antirrealista que questiona a nossa capacidade em distinguir adequadamente quais aspectos teóricos são centrais ou não em uma análise retrospectiva — algo que é crucial para estratégias realistas seletivas, como o *divide et impera* de Psillos. Se sua crítica proceder, não estaremos justificados em delimitar o que é central nas teorias abandonadas, já que não haveria garantia de que os teóricos do passado concordariam com a nossa ação; desse modo, poderíamos estar *projetando o presente sobre o passado*. Como esta crítica é similarmente apresentada à aplicação de uma teoria da referência aos casos históricos realizada por Psillos, exporemos a crítica de Stanford mais detalhadamente nas subseções 3.3.2 e 4.5.1.

Capítulo 3: O papel das teorias da referência na defesa da continuidade teórica

Neste capítulo, vamos expor como o realista pode tentar acomodar ou mitigar a descontinuidade teórica na ciência utilizando recursos da filosofia da linguagem. Em especial, examinaremos a possibilidade de, mediante adequada interpretação, preservar a referência dos termos “teóricos”, ou seja, daqueles que putativamente denotam entes e processos inobserváveis, nos episódios de substituição teórica na ciência. A referência a um dado inobservável é um pressuposto teórico relevante capaz de garantir, ao menos parcialmente, uma continuidade substancial entre as teorias científicas sucessivas, na história da ciência.³³ Admite-se, naturalmente, que a teoria posterior ofereça descrições mais refinadas do referente compartilhado com uma teoria anterior, sendo então classificada como mais próxima da verdade do que aquela que a precedeu. Tal defesa visa atrelar a verdade aproximada de uma teoria ao acerto na referência a um dado inobservável, pois tolera a renúncia da acurácia descritiva (total ou parcialmente, a depender de cada vertente) acerca desse tipo de entidade. Caso possamos defender apropriadamente a manutenção de referência entre as teorias, haverá aí algum tipo de continuidade teórica que permitirá sustentar o núcleo do AM mesmo perante os casos históricos da MIP, reforçando a convicção realista de que o sucesso teórico é devido à existência dos inobserváveis referidos. Na filosofia da linguagem, o estudo das teorias da referência pode, talvez, oferecer recursos para a defesa da continuidade que nos interessa aqui, enquanto realistas científicos.

A *referência* (ou denotação) é uma relação entre expressões linguísticas (de certos tipos) e objetos (sejam eles concretos ou não)³⁴, na qual essas expressões são usadas para *fazer referência a*, ou *falar dos* objetos que, por sua vez, são os *referentes* dessas expressões ou *aquilo sobre os quais elas falam*. As teorias da referência tratam justamente dessa relação entre uma expressão aparentemente denotativa e o seu referente, especificando o que torna um objeto o referente de uma expressão denotativa ao delimitar as condições em que uma relação de referência é ou não é bem-sucedida. Outro aspecto estreitamente associado à referência é o *significado* — que é o conteúdo das expressões linguísticas ou aquilo que elas expressam —, pois o referente de uma expressão (quando o tem) também contribui para o seu significado. Embora sejam frequentemente apresentadas conjuntamente, as teorias da referência diferem

³³ Também existem propostas realistas que defendem a continuidade teórica prescindindo da manutenção de referência, como na estratégia (d). Porém, essa é uma estratégia menos debatida na literatura. Para uma análise dessa abordagem, veja-se Silva (2022).

³⁴ Objetos concretos podem ser entendidos em oposição aos objetos abstratos, pois esses não fazem parte do mundo material ou não têm localização espacial específica — como os números ou conceitos —, enquanto aqueles concretos são materiais ou localizáveis no espaço e no tempo.

das teorias do significado para as expressões denotativas: ao passo que as primeiras caracterizam a relação de denotação, as últimas se preocupam com a contribuição das expressões denotativas para o significado dos enunciados nos quais elas ocorrem. Enfatizo esse ponto porque, nesse primeiro momento, o nosso foco recairá no aspecto da referência; assim, nos restringiremos à exposição e discussão das teorias da referência, assinalando algo acerca do significado somente quando necessário.

Quanto às expressões linguísticas usadas para fins de denotação, dispomos dos *termos singulares* e os *termos gerais*, que são usados com o intuito de referir a um objeto ou a mais de um objeto: enquanto *nomes próprios* e *descrições definidas* são singulares, termos como *ser humano* ou *gato* são gerais. Dado o escopo deste trabalho, o nosso interesse recai naqueles termos que são usados com o intuito de referir aos objetos inobserváveis, nomeadamente, os termos teóricos. Eles podem ser tanto singulares como gerais: o termo *Big Bang* é singular por designar um único evento, já *elétron* é geral por aplicar-se a um conjunto de entidades.

O nosso objetivo é analisar o que as teorias da referência declaram sobre os termos teóricos (os termos que referem a entidades inobserváveis). No entanto, tais teorias não foram primariamente desenvolvidas para esses termos, sendo inicialmente elaboradas para o tratamento dos nomes próprios. Em virtude dessas circunstâncias, cada teoria da referência será inicialmente exposta (de forma resumida) de acordo com os seus propósitos originais e depois examinaremos sua possível adaptação aos termos teóricos. Respeitando uma ordem cronológica em que essas propostas surgiram, passaremos à exposição das teorias *descritivista*, *causal* e *híbrida* (*descritivo-causal*).

3.1. Teoria descritivista

A *teoria descritivista da referência* remonta aos trabalhos de Gottlob Frege (1892) e Bertrand Russell (1905; 1918). Segundo ela, um termo refere a um objeto se e somente se suas descrições são satisfeitas. As descrições atreladas a um termo delimitam as propriedades que um objeto deve possuir para ser o seu referente. Tais descrições identificadoras determinam a referência do termo apresentando-o de uma certa maneira, consistindo naquilo que denominamos *sentido* do termo. Esta é uma teoria da referência indireta, pois a ligação de um nome com o seu referente é intermediada por descrições (i.e., seu sentido).

De acordo com a concepção descritivista, o termo singular *Pirro de Élis* se refere ao objeto que satisfaz as descrições filósofo grego nascido em Élida e primeiro filósofo cético.

Igualmente, o termo geral *água* refere a tudo o que é um recurso natural abundante no planeta Terra, insípido (sem sabor) e tem a estrutura molecular H₂O. As descrições são condições *necessárias* (indispensáveis) e *suficientes* (bastam) para que um termo refira a um (ou mais) objeto(s); caso não exista qualquer objeto que corresponda às descrições atreladas ao termo, ele falhará na referência.

Tanto Frege (com a introdução da noção de sentido) como Russell (ao enfatizar o papel das descrições na referência) contribuíram, à sua maneira, para o que atualmente chamamos de teoria descritivista. Frege procurava explicar a diferença de significado entre enunciados de identidade com nomes correferentes. De acordo com a concepção dos nomes vigente à época, o significado dos nomes reduzia-se ao seu papel referencial. Esta é uma concepção dos nomes, introduzida por John Stuart Mill (1843). Porém, esta concepção não nos permite explicar por qual razão nomes que partilham da mesma referência acabam diferindo quanto ao significado; pior, a concepção milliana não explica o que liga um nome ao seu referente, que é justamente o que uma teoria da referência deveria fazer.

Os enunciados “Stephen King é Stephen King” e “Richard Bachman é Stephen King” dizem respeito à mesma pessoa, mas diferem quanto ao significado; o segundo enunciado é mais informativo que o primeiro: ele não expressa apenas algo trivial, pois também nos informa que Stephen King publicou livros sob o pseudônimo de *Richard Bachman*. Como isso pode acontecer? Frege explicou que essa diferença semântica consistia no fato de que um nome próprio tem associado a si não só um referente (quando o tem), mas também (ao menos) um sentido, que é um modo de apresentação do referente. Segundo Frege, o significado de um nome consiste não apenas do seu referente (quando o tem) como do seu sentido. No caso de identidade citado, os dois nomes têm sentidos diferentes: enquanto o nome *Stephen King* apresenta o referente em questão como o autor de *Carrie*, o outro nome desse referente — *Richard Bachman* — o apresenta como o autor de *Fúria*. Por vezes Frege expressava o sentido de um nome através de descrições definidas, como no caso do nome próprio *Aristóteles*, seu sentido poderia ser: o pupilo de Platão e professor de Alexandre O Grande ou o professor de Alexandre O Grande que nasceu em Estagira (FREGE, 1892, p. 210n). Embora esta tese tenha sido motivada por questões de significado, é a concepção de que o sentido determina a referência que nos interessa.

Já Russell estava interessado nas descrições definidas, que são expressões da forma “o/a tal e tal”. As descrições definidas podem ser analisadas como uma conjunção de condições. Como no caso de “o autor de *Aufbau* era judeu” podemos analisá-la em termos das seguintes condições: existiu pelo menos um autor de *Aufbau*, existiu no máximo um autor de

Aufbau e qualquer um que tenha sido o autor de *Aufbau* era judeu. A descrição definida transformou-se em enunciados de generalizações quantificadas, os quais serão conjuntamente verdadeiros se houver um único indivíduo que exemplifique as propriedades atribuídas a si nessas generalizações — nesse caso, elas são verdadeiras porque dizem respeito a Rudolf Carnap; como consequência, a descrição definida em questão é acerca dele também. A contribuição de Russell consistiu em indicar que as propriedades atribuídas a um objeto poderiam delimitar e intermediar a referência de uma expressão. Posteriormente ele estendeu as suas reflexões sobre as descrições definidas, estabelecendo que os nomes eram equivalentes a essas descrições; no entanto, como esta é uma tese sobre o significado dos nomes e ultrapassa questões sobre a referência, não trataremos dela aqui.

Tanto em Frege como em Russell temos uma concepção da referência indireta, que é mediada seja pelo modo de apresentação do objeto ou pela delimitação das suas propriedades. Em ambos os casos são as descrições que cumprem um papel relevante na determinação da referência. A tese de que as expressões têm sentido e referência não se estende apenas aos nomes, abrange todos os termos singulares e os gerais, assim como frases declarativas. A mediação das descrições atreladas a um termo (i.e., do seu sentido) na determinação do seu referente é o aspecto herdado de ambos os filósofos e foi mantido na formulação padrão da teoria descritivista da referência, assim como a concepção de que o significado de um termo é composto pelo seu sentido (descrições) e referente (quando o tem) — o mesmo não se sucedeu com a equivalência estabelecida por Russell entre nomes e descrições.

Embora a teoria descritivista assimile o fato de que as descrições cumprem um papel importante no aprendizado de um nome ou na caracterização do seu referente, ela padece de, ao menos, dois problemas: a *multiplicidade de descrições* e a *ambiguidade indesejada* (*unwanted ambiguity*) (DEVITT E STERELNY, 1999, pp. 48-49). O problema da multiplicidade de descrições é derivado do fato de que várias descrições podem ser associadas a um nome por diferentes pessoas. Por exemplo, o nome *Aristóteles* também costuma ser associado às descrições: o criador da lógica silogística, o autor da obra *Metafísica*, o criador da ética das virtudes, etc.; de modo que não é muito claro quais descrições são mais importantes do que outras para determinar oficialmente a referência do nome. Afinal, qual dessas descrições é mais relevante ou quantas delas precisam ser satisfeitas para que um nome refira de modo bem-sucedido? Já o problema da ambiguidade indesejada surge mesmo quando temos um princípio para associar certas descrições a um nome e assim determinar oficialmente sua referência, pois as pessoas ainda podem considerar que outras descrições,

distintas das previamente estabelecidas para um nome, são mais relevantes para delimitar seu referente.

Outra objeção à teoria descritivista é de que as descrições associadas a um nome não são necessárias e nem suficientes para a referência. Kripke (1980, pp. 83-7) enfatizou esse aspecto e ilustrou-o através dos exemplos de *Einstein* e de *Gödel*. Ainda há quem mencione Einstein como o criador da bomba atômica, apesar de isso ser um equívoco; tal descrição é equivocada acerca de Einstein, porém, quem se engana ao associar essa descrição ao nome *Einstein* ainda refere efetivamente a Einstein ao usar o nome que lhe foi atribuído. Nesse caso, ainda que a descrição não refira a Einstein, o nome *Einstein* refere, de modo que tal descrição não é necessária para que o nome *Einstein* refira de modo bem-sucedido. Quanto a Gödel, sabemos que ele provou o teorema da incompletude da aritmética. Agora suponha que a prova não tenha sido elaborada por ele, sendo roubada de um dos seus alunos chamado Schmidt. Desconhecendo esse fato, muitas pessoas associam ao nome *Gödel* a descrição: o homem que demonstrou o teorema da incompletude. Embora a descrição apresentada denote Schmidt, não parece razoável considerar que o nome *Gödel* refere a ele, de modo que essa descrição não é suficiente para determinar a referência desse nome. Esse exemplo sugere que a ligação entre um nome e o seu referente pode ser muito mais direta e explicada em termos de outros fatores. Isso motivou Kripke a defender uma outra teoria da referência: a teoria causal.³⁵ Em breve a analisaremos, mas antes vejamos o desempenho da teoria descritivista quando aplicada aos termos teóricos.

3.1.1 Adaptação da teoria descritivista aos termos teóricos

Pretendendo, inicialmente, analisar a referência dos nomes próprios, os proponentes da teoria descritivista não contemplaram diretamente os termos teóricos das teorias científicas, que tipicamente são nomes comuns. No entanto, a teoria descritivista da referência pode ser aplicada aos termos teóricos com algumas adaptações, sendo que a denotação de um inobservável será mediada pelas respectivas descrições fornecidas por alguma teoria.

Embora de maneira implícita, a teoria descritivista foi aplicada aos termos teóricos nos projetos (mal sucedidos) do positivismo lógico que buscavam estabelecer definições explícitas desses termos ou suas sentenças redutivas a fim de eliminar o discurso científico

³⁵ As famosas objeções modais à teoria descritivista formuladas por Kripke (1980) e o exemplo da Terra Gêmea de Putnam (1974) atingem somente a teoria descritivista do significado dos nomes (DEVITT e STERELNY, 1999, pp. 52-53); por essa razão não as abordaremos aqui — para conferir essas e outras objeções, veja Devitt e Sterelny (pp. 49-64; pp. 83-88) e Lycan (2008, pp. 39-43; pp. 58-60).

sobre os inobserváveis, como exposto no início do capítulo 1. Nesses projetos, eram utilizadas as descrições associadas aos termos teóricos na tentativa de expressá-los em um vocabulário puramente observacional; algo que claramente envolvia a concepção descritivista da referência. Contudo, a constatação ulterior, dos próprios positivistas, de que o significado dos termos teóricos não pode ser totalmente expresso por meio de relatos de observação, e a dificuldade em distinguir adequadamente o vocabulário teórico do observacional culminaram na falha desses projetos positivistas.

Como no caso dos inobserváveis as descrições são fornecidas pelas teorias científicas, pode haver uma menor chance de surgirem os problemas da multiplicidade de descrições e da ambiguidade indesejada, pois através da caracterização destas entidades os cientistas podem delimitar — mesmo que indiretamente — quais descrições são determinantes e relevantes para denotá-las.

Seguindo a teoria descritivista, pode haver continuidade de referência em episódios de substituição teórica se os termos teóricos das respectivas teorias pressupuserem i) as mesmas descrições para a referência ou ii) descrições diferentes que sejam consistentes entre si e captem o mesmo referente — pois descrições (ou sentidos) diferentes podem referir a um mesmo objeto. No primeiro caso, a manutenção da referência mediante a manutenção de descrições é ameaçada pelo *holismo semântico* — aspecto que será discutido em seguida. Já o segundo cenário parece difícil de ser realizado, uma vez que a mudança de descrições teria que ser suficientemente pequena para não implicar a caracterização de uma entidade diferente, levando então a uma indesejada descontinuidade na referência.³⁶

A conservação de descrições entre teorias pode, porém, ser um ideal inalcançável. A concepção descritivista de continuidade teórica é ameaçada por um aspecto que já foi ressaltado anteriormente na filosofia: o holismo semântico. Esse holismo é a tese segundo a qual o significado de uma expressão depende da totalidade (ou de uma parte considerável) da linguagem à qual ela pertence.³⁷ Embora o holismo seja um aspecto concernente ao

³⁶ Há casos, como o do átomo, que podem nos levar a reconsiderar esse ponto, pois consideramos que desde os gregos temos uma referência ao átomo apesar da pressuposição de descrições bem diferentes das atuais, como “entidade fundamental e indivisível”. Segundo nosso conhecimento atual, o átomo não é a entidade mais fundamental do universo, sobretudo porque é composta por outras partes, sendo então divisível. Se explorarmos a descrição antiga, não estaremos confortáveis em dizer que o átomo dos gregos é a mesma entidade que o átomo dos modernos, uma vez que há outras entidades mais fundamentais que o átomo (que fazem parte da sua composição) e ele não é indivisível. Em um sentido muito generalista concedemos que o átomo dos gregos é o mesmo átomo dos modernos, mas dada uma análise mais pormenorizada em termos de descrições, fica difícil sustentar essa concessão pela grande mudança descritiva presente em tal caso.

³⁷ O holismo semântico remonta aos trabalhos de Quine e Hempel. Segundo Quine, “é enganoso falar do conteúdo empírico de um enunciado individual” (QUINE, 1951, p. 43); “a unidade do significado empírico é toda a ciência” (QUINE, 1951, p. 42). Já Hempel diz que, “o significado cognitivo de um enunciado em uma

significado, ele acaba gerando implicações para a referência, uma vez que as descrições atreladas a um termo não só constituem seu significado como determinam a sua referência. No caso de um termo teórico, seu significado também depende de outros fatores teóricos que são usados na sua definição e de outros termos a ele relacionados — que, por sua vez, também dependem semanticamente de outros fatores —, fazendo com que seu significado seja derivado da teoria como um todo. Qualquer alteração teórica pode impactar os termos teóricos, seja diretamente ou por meio de outros fatores teóricos que afetam indiretamente seu significado. Em uma transição teórica sempre há algum tipo de alteração que faz uma nova teoria ser adotada, já que ela seria mais adequada, abrangeria mais fenômenos e seria mais precisa do que uma teoria anterior. Essa alteração já é suficiente para afetar o significado dos termos de uma teoria, de modo que nas substituições teóricas o significado dos termos em questão altera-se e, conseqüentemente, as descrições a ele atreladas. E basta que uma descrição de um termo teórico abandonado não seja mantida em uma teoria posterior para que ele falhe na referência; como consequência, não restaria espaço para a defesa da manutenção de referência a um dado inobservável entre duas teorias. É certo que as implicações indesejáveis do holismo semântico acabam inviabilizando a possibilidade de alterações ou refinamentos teóricos por parte dos cientistas, uma vez que as alterações descritivas são vistas como mudanças referenciais ao invés de serem encaradas como meros aperfeiçoamentos das características atribuídas às entidades inobserváveis. Como esses aspectos não são compatíveis com a imagem do aperfeiçoamento científico, também quanto a esse ponto a teoria descritivista da referência não parece ter um bom desempenho.

Outro problema que também surge é a *incomensurabilidade semântica*, tal como desenvolvida pelo filósofo Thomas Kuhn.³⁸ De modo geral, a incomensurabilidade consiste na falta de aspectos observacionais, metodológicos e semânticos que sejam comuns entre dois paradigmas, o que dificulta a comparação entre esses paradigmas. Nesta concepção, uma teoria faz parte de um paradigma, principalmente quando o paradigma é tomado como uma *matriz disciplinar*, pois ele será composto por elementos que guiam o trabalho de um

linguagem empírica é refletido na totalidade das suas relações lógicas com outros enunciados da linguagem” (HEMPEL, 1950, p. 59).

³⁸ Tanto Thomas Kuhn — no livro *A Estrutura das Revoluções Científicas* (1970) — como Paul Feyerabend — em “Explation, Reduction and Empiricism” (1962) — introduziram o conceito de incomensurabilidade; no entanto, trabalharemos com o conceito de incomensurabilidade tal como apresentado por Kuhn no livro *A Estrutura das Revoluções Científicas*. Também há outras dimensões da incomensurabilidade, como a *observacional* (mudança de dados empíricos) e a *metodológica* (mudança de método), mas só abordaremos a semântica porque ela tem implicações diretas para o aspecto de referência dos termos teóricos. Maiores detalhes sobre a incomensurabilidade podem ser conferidos em Oberheim (2018).

cientista, como teoria, hipóteses auxiliares, instrumentos, técnicas de experimentação, crenças, etc. Por outro lado, o paradigma também pode ser entendido como um *exemplar*, onde ele se configura como “soluções concretas de problemas que os estudantes encontram desde o começo da sua educação científica, seja em laboratórios, nos exames ou no final dos capítulos dos livros didáticos” (KUHN, 1970, p. 187). Mas em qualquer acepção que seja, segundo o que é estabelecido por Kuhn, a teoria sempre estará subsumida em um paradigma.

A incomensurabilidade semântica consiste na mudança de significado dos termos das linguagens das teorias entre a transição de paradigmas. Assim como o holismo, esse também é um aspecto relativo ao significado, porém acaba gerando implicações para a referência dado que o significado de um termo — composto pelo seu sentido (descrições) e referente (quando o tem) — também depende das descrições a ele atreladas que determinam a sua referência. Se de fato existir essa incomensurabilidade semântica em episódios de substituição teórica, tal como ocorre com o holismo, é bastante provável que alguma descrição atrelada a um termo teórico não seja mantida em uma teoria posterior, de modo que um termo abandonado falhará na referência, inviabilizando a manutenção de referência ao inobservável entre as teorias. Kuhn enfatizou esse aspecto ao discutir a possibilidade de derivação das leis de Newton a partir das leis de Einstein (o que permitiria defender alguma continuidade entre as teorias Newtoniana e Einsteiniana).

[...] os referentes físicos dos conceitos Einsteinianos não são de modo algum idênticos com os conceitos Newtonianos que têm o mesmo nome (a massa Newtoniana é conservada; a [massa] Einsteiniana é conversível em energia. Somente em velocidades relativamente baixas as duas podem ser medidas da mesma maneira, e mesmo nesse caso elas não devem ser concebidas como sendo a mesma). (KUHN, 1970, p. 102).

Em matéria descritiva, a dimensão semântica acaba tendo implicações desfavoráveis para a defesa da manutenção de referência entre as teorias. Como lição, tal defesa não pode ser totalmente vinculada ao conteúdo das teorias, pois suas respectivas linguagens serão alteradas com o refinamento teórico inerente aos episódios de substituição.

3.2. Teoria causal

A *teoria causal* (ou *histórico-causal*) da referência é devida a Kripke (1980), Putnam (1975b) e Donnellan (1970). Segundo esta teoria, um nome refere a um objeto em virtude de uma certa ligação causal entre eles. Esta ligação surge quando, em uma ocasião denominada *batismo*, o nome é atribuído ao seu objeto de referência por meio de um ato de

ostensão em que se aponta para tal referente. Após esse evento inicial que fixa a referência do nome (denominado fixação de referência do termo), sua referência é repassada ou transferida a outros membros da comunidade linguística através de uma cadeia histórica de comunicação (denominada *empréstimo de referência* do termo), em que aqueles que aprendem a utilizar o nome devem usá-lo com a intenção de referir ao mesmo objeto que aqueles que repassaram a referência desse nome. Tanto a introdução do nome (que é suscitada pelo contato causal com seu objeto de referência) como a cadeia histórica de comunicação (que contém elos de causação de repasse do nome) conferem a esta teoria o rótulo de *causal*. Esta é uma teoria da referência direta, pois para que um nome seja usado de maneira bem-sucedida na denotação de um objeto, não há nenhum aspecto mediando a relação de referência.

Alguém nasce, vamos dizer, um bebê; seus pais o chamam por um certo nome. Eles falam dele para seus amigos. Outras pessoas o conhecem. Através de várias conversas o nome é espalhado de elo em elo como se fosse uma corrente. Um falante que está no fim dessa corrente e ouviu falar de, digamos, Richard Feynman, no mercado ou em algum outro lugar, pode referir a Richard Feynman mesmo que ele não lembre de quem ouviu esse nome pela primeira vez ou de alguém que tenha usado esse nome. Ele sabe que Feynman é um físico famoso. Uma certa passagem de comunicação que surgiu a partir do próprio Feynman chega até esse falante. Ele está referindo a Feynman mesmo que não seja capaz de identificá-lo univocamente. [...] Uma cadeia de comunicação que remonta ao próprio Feynman foi estabelecida em virtude da sua participação em uma comunidade que passou o nome de elo em elo, ao invés de uma cerimônia na qual ele tenha feito em um estudo privado estabelecendo que: por “Feynman” eu quero dizer o homem que fez tais e tais coisas. (KRIPKE, 1980, p. 91).

A teoria causal parece adequar-se bastante bem às nossas práticas linguísticas de uso dos nomes próprios. Mesmo que as descrições cumpram um papel relevante no aprendizado e na caracterização dos referentes das expressões denotativas, ao usar um nome não parecemos ter essas descrições imediatamente em mente para julgar se ele refere ou não. A virtude desta teoria é mostrar que a relação de referência é muito mais informal e direta do que era estabelecido pela teoria descritivista, pois basta que um nome seja introduzido na linguagem para ser associado a um objeto. A teoria causal também foi estendida para termos gerais por Kripke (1980) e Putnam (1975b). No caso de um termo geral é impossível o contato com todos os seus objetos de referência, de modo que a introdução e a fixação da sua referência diferem um pouco. Supõe-se que um termo geral refere aos objetos que pertencem a uma mesma categoria ou tipo, isto é, aos objetos que compartilham da mesma natureza da amostra presente na ocasião de batismo — como exemplo, o termo geral *gato* referirá a todos os objetos que têm a mesma natureza de um gato utilizado para introduzir esse termo geral.

As objeções à teoria descritivista apontadas na subseção precedente não parecem aplicar-se à teoria causal, uma vez que ela dispensa o papel intermediário das descrições para

a referência. Quanto aos nomes próprios, via de regra, a teoria causal parece bastante adequada. Entretanto, existem certos casos em que a simplicidade dessa teoria pode levar a equívocos, justamente por desqualificar o papel auxiliar das descrições — e, como veremos na seção 3.3, tais problemas motivaram a elaboração de uma *teoria híbrida da referência*, que combina elementos causais e descritivos para tratar da referência.

Gareth Evans (1973) apresentou algumas objeções à teoria causal que acabaram por fomentar a consideração de que as descrições não podem ser totalmente descartadas na discussão da referência.³⁹ Exporei brevemente uma objeção que ressalta o papel ineliminável das descrições, que é o problema *qua* — assim denominado por Devitt e Sterelny (1999). O problema *qua* mostra que as pessoas podem ter crenças categoriais erradas sobre os referentes, de modo que quem introduz ou utiliza um nome não pode estar enganado e tem de ter a intenção de referir ao objeto que pertence à categoria (ou tipo) apropriada. Segundo Evans,

Nós aprendemos do exemplo de *Arthur of Britain* de E. K. Chambers que Arthur teve um filho chamado Anir “cuja lenda talvez o tenha confundido com seu local de sepultamento”. Se a noção Kripkeana de fixação da referência é tal que aqueles que dizem Anir foi o lugar de sepultamento de Arthur podem denotar uma pessoa, [...] certamente [Kripke] não está justificado pelas críticas feitas à Teoria Descritiva. (EVANS, 1973, p. 198).

Sem qualquer aporte descritivo, a teoria causal não parece ter recursos para delimitar apropriadamente que um objeto de referência pertence a uma certa categoria, deixando espaço para o surgimento de erros categoriais. Se ao menos alguma descrição categorial estivesse atrelada ao nome, esta confusão poderia ser evitada. Do mesmo modo, como apontam Devitt e Sterelny (1999, pp. 91-93), meramente estipular que, no caso de um termo geral, ele refere a objetos com a mesma natureza da amostra presente na ocasião do batismo também não nos permite especificar qual natureza é essa. As descrições poderiam nos ajudar nesse caso novamente. Em suma, em algum nível, as descrições podem cumprir um papel relevante para especificar apropriadamente a referência de um termo. Antes de discutir como isso poderia ser feito, vejamos a possibilidade de aplicação da teoria causal aos termos teóricos.

³⁹ Outras objeções podem ser consultadas em Devitt e Sterelny (1999, pp. 73-76; pp. 88-96) e Lycan (2008, pp. 55-58).

3.2.1. Adaptação da teoria causal aos termos teóricos

Inicialmente voltados para a análise dos nomes próprios, os proponentes da teoria causal não mostraram diretamente como aplicá-la aos termos teóricos. Apesar disso, Kripke (1980) já havia dado pistas de como isso poderia ser realizado ao admitir que alguma descrição poderia ser usada na fixação da referência de termos cujos referentes não pudessem ser apontados em um ato de ostensão. Considere o caso da descoberta do planeta Netuno:

Netuno foi suposto como o planeta que causou tais e tais discrepâncias nas órbitas de certos outros planetas. Se Leverrier realmente deu o nome ‘Netuno’ ao planeta antes de sequer vê-lo, então ele fixou a referência de ‘Netuno’ por meio da descrição mencionada. Naquela época ele não era nem capaz de ver o planeta através de um telescópio. (KRIPKE, 1980, p. 79n).

Kripke indicou que podemos recorrer ao (suposto) *poder causal* de uma entidade não observada (ou, no caso que mais nos interessa nesta dissertação, inobservável) para introduzirmos seu respectivo termo. A referência será bem-sucedida nesse tipo de caso mesmo quando a descrição apresentada para delimitar o poder causal de uma entidade não for satisfeita por ela: tal descrição causal cumpre apenas um papel auxiliar no ato de fixação da referência do termo (o batismo), sendo apenas um componente para vincular o termo à entidade. Na teoria causal, a descrição não intermedia a referência de maneira alguma, ela só é usada para justificar a introdução de um dado termo.

Segundo a teoria causal, um termo teórico refere a um objeto porque foi atribuído a ele em uma ocasião de batismo. No entanto, no caso dos inobserváveis, *não é possível realizar o ato de ostensão do batismo no qual se delimita o referente do seu termo, uma vez que não temos acesso perceptual a essas entidades*. Assim, a fixação da referência dos termos teóricos aparentemente traz complicações inexistentes para os termos observacionais. Para estes termos, o batismo é ostensivo; já para aqueles, a proposta seria apontar-se para um fenômeno observável (ou um conjunto deles) e pressupor que a entidade inobservável que causa (ou produz) tal(is) fenômeno(s) é o referente de um dado termo teórico. Pela teoria causal, o referente de um termo teórico seria o inobservável que causa um certo tipo de fenômeno; e do mesmo modo que ocorre com os termos observacionais, o seu referente seria repassado em uma comunidade linguística, fazendo com que qualquer um que utilize esse termo refira diretamente ao objeto (ou tipo de objeto) ao qual ele foi inicialmente atribuído e suscitou a sua introdução.

Hardin e Rosenberg (1982) se basearam na teoria causal para defender a manutenção de referência entre teorias. Segundo eles, “uma estratégia permissível para os realistas é

deixar a referência seguir o papel causal” (HARDIN; ROSENBERG, 1982, p. 613). Nesses moldes, duas teorias compartilharão da mesma referência a um inobservável se os seus respectivos termos teóricos denotarem a entidade que é responsável por um certo tipo de fenômeno. Assim, se ambas as teorias tratarem do mesmo domínio de fenômenos, a manutenção de referência entre elas estará automaticamente assegurada, uma vez que elas tratam do mesmo domínio observável que delimita o mesmo referente para ambos os termos.⁴⁰

À primeira vista, a teoria causal mostra-se bastante promissora para a defesa da manutenção de referência entre as teorias. Como as descrições não mediam a referência nesta concepção, o problema do holismo e da incomensurabilidade não atingem a aplicação da teoria causal aos termos teóricos. Não importa quanta alteração semântica tenha ocorrido em uma transição de teorias, pois o único fator a determinar a referência será a atribuição do termo a uma entidade que supostamente causa um certo tipo de fenômeno; como consequência, basta que as teorias tenham sido elaboradas para tratar do mesmo domínio de fenômenos para que elas pressuponham o mesmo agente causal como referente, mantendo assim a referência em episódios de substituição teórica. Por conseguinte, a comparação entre teorias (aspecto nefasto da incomensurabilidade) será regulada pela manutenção de referência dos termos, algo que parece ser mais facilmente alcançado com a teoria causal.

Todavia, a aparente capacidade da teoria causal em assegurar a referência dos termos teóricos é um fator que acaba, sob outro ponto de vista, prejudicando a sua legitimidade de aplicação nesses casos. Qualquer coisa que cause um certo tipo de fenômeno pode ser o referente de um dado termo teórico. O abandono total das descrições para a referência faz com que os inobserváveis sejam desconhecidos a tal ponto que a manutenção da referência será efetuada de modo demasiado fácil, tornando esta alternativa desinteressante para um realista que se preocupa com a apresentação de descrições cada vez mais refinadas acerca dessas entidades. Como Larry Laudan (1984) sinalizou, a teoria causal é demasiadamente tolerante quanto à referência, fazendo com que os termos teóricos sejam imunes à falha referencial, tornando a referência algo trivial. Além disso, sua aplicação aos episódios de

⁴⁰ O recurso ao poder causal das entidades inobserváveis para assegurar a referência também foi uma estratégia adotada pelos proponentes do realismo de entidades, como Hacking (1983) e Cartwright (1983). Segundo esta posição, a habilidade de manipulação causal dos inobserváveis permite o nosso comprometimento com a existência dessas entidades. Assim como ocorre com a teoria causal, o realismo de entidades não pressupõe o conteúdo das teorias científicas para assegurar a referência ao inobservável; só que apesar dessa compatibilidade, esta posição se apoia em considerações experimentais para tratar do papel causal das entidades, algo que não é contemplado pelas teorias da referência.

substituição teórica acaba confundindo agenda ontológica com agenda explanatória na tentativa de assegurar a manutenção da referência.

Duas coisas estão erradas com essa teoria funcional da referência. Em primeiro lugar, ela prova ser bastante *tolerante* para os propósitos realistas. [...] A concepção aristotélica de lugar natural cumpre muito do mesmo papel causal e explanatório que as forças gravitacionais newtonianas. Os vórtices cartesianos também cumprem uma função similar. Alguém está seriamente preparado para sustentar que Aristóteles e Descartes, cujas estruturas conceituais eram tão contrárias à ação à distância, estavam realmente ‘referindo’, [...], à atração gravitacional? O que Newton, Descartes e Aristóteles têm em comum é a convicção de que há algo que causa a queda dos corpos pesados. E em um sentido trivial, eles estavam, portanto, referindo à “causa da queda”. Mas isso não estabelece nenhuma semelhança interessante de referência no nível da estrutura explanatória. Em segundo lugar, e mais crucialmente, o relato de referência de Hardin e Rosenberg confunde uma agenda explanatória compartilhada (i.e., problemas em comum a serem resolvidos) com uma agenda ontológica compartilhada (i.e., as características das entidades explanatórias postuladas).

O fato de que (a) duas teorias abordam (muitos dos) mesmos problemas e (b) que essas teorias agrupam os problemas a serem resolvidos de maneira similar, certamente *não* é suficiente para estabelecer que (c) as ontologias explanatórias das duas teorias ‘referem’ às mesmas entidades. Especialmente para um realista, o qual considera que ‘aquilo que a teoria refere’ deve ter alguma conexão com as afirmações existenciais da teoria sobre a estrutura [...] [do mundo]. Tornar a referência parasitária [dependente] do que *é explicado* ao invés do que está *fazendo a explicação* implica que podemos estabelecer que uma teoria refere independentemente de qualquer análise detalhada do que essa teoria asseire. Se, para salvar sua teoria da referência, o realista científico é forçado a essa semântica voodoo, então o realismo dificilmente parece valer a pena. (LAUDAN, 1984, pp. 160-161).

Além dessas críticas, é preciso assinalar que a teoria causal pressupõe a existência metafísica da causalidade para explicar a referência dos termos teóricos. Esse não é um ponto pacífico no debate e precisa ser adequadamente defendido.⁴¹

A existência metafísica da causalidade é colocada em dúvida pela falta de experiências que possam confirmar a presença de poderes causais no mundo. Além disso, existem na história da filosofia várias críticas ao nexo causal do ponto de vista cognitivo. David Hume (1711-1776) já havia argumentado que apenas percebemos conjunções de fenômenos, mas não uma ligação necessária entre eles que nos permita dizer que um é a causa do outro — como quando meu computador está desligado, o pressionar do botão de ligar é sucedido pela ligação do computador, porém, não percebo qualquer fator adicional que conecte inevitavelmente esses dois eventos, de modo a dizer que o primeiro causa a ocorrência do segundo.

Mesmo que o realista científico venha a pressupor a existência de poderes causais ao longo do debate, o antirrealista (embora implicitamente) parece depender igualmente de tal tipo de existência ao se comprometer epistemicamente com as entidades *observáveis*. A

⁴¹ Para uma análise recente, ver Fabian (2022).

percepção de objetos observáveis é explicada pela suposição de que tal percepção é causada por contrapartidas no mundo externo aos indivíduos. Dessa maneira, a nossa interação com os observáveis seria mediada pelos dados dos sentidos, isto é, aquilo que nos é entregue pela percepção e que nos informariam adequadamente sobre tais objetos por serem produzidos a partir deles. Se o próprio antirrealista pressupõe em alguma medida a existência de poderes causais para a percepção e o conhecimento das entidades observáveis, o realista também estaria autorizado a fazer o mesmo em sua defesa dos inobserváveis.

Além disso, a própria problematização sobre a causalidade não se restringe à aplicação da teoria causal da referência, pois até o AM pressupõe de modo subjacente tal aspecto na explicação do sucesso da ciência. A existência dos inobserváveis seria a melhor explicação para o sucesso das teorias que os supõem porque essas entidades (supostamente) causam a ocorrência dos fenômenos observáveis, ajudando então na geração de previsões inéditas e unificações desses fenômenos. Desde o início, o realista que defende o AM também se compromete com a existência de poderes causais.

Aliás, com o auxílio do mesmo tipo de raciocínio utilizado no AM, i.e., a IME, podemos justificar a pressuposição da existência da causalidade para delimitar a referência das entidades inobserváveis. Afirmamos que há uma relação causal entre dois objetos porque comumente eles são contíguos e um deles sucede o outro com regularidade; desse modo, essa regularidade não tende a ser encarada por nós como um acaso ou algo gratuito, ao contrário, ela fomenta a ideia de que o primeiro objeto tem o poder de produzir a ocorrência do segundo. Podemos então dizer que a existência de poderes causais parece ser a melhor explicação para a conjunção regular dos objetos. Assim como pensamos que os inobserváveis podem ser defendidos pelo seu potencial explicativo do sucesso da ciência, a existência da causalidade também poderia ser justificada do mesmo modo, já que é uma boa explicação para a regularidade dos fenômenos — Galen Strawson (1989), por exemplo, é um dos expoentes desse tipo de justificativa. E mais: ele argumenta que o próprio Hume teria aceitado esse raciocínio que vai da regularidade ao poder causal real; isso contraria a interpretação usual, de Hume como alguém que teria reduzido a causação à mera regularidade — ver, a esse respeito, Chibeni (2012, Apêndice, pp. 246-249).

A suposição da existência metafísica da causalidade cumpre um papel explicativo importante quanto a: percepção de objetos observáveis, ligação entre entidades inobserváveis e observáveis e a regularidade dos fenômenos. Além disso, se a suposição desse aspecto ontológico está presente (mesmo que implicitamente) no comprometimento antirrealista com as entidades observáveis e é crucial para a explicação realista do sucesso da ciência através do

AM, temos não só razões para crer na existência da causalidade como também para justificar novamente a pressuposição desse aspecto ontológico, agora no que diz respeito à teoria causal.

Em um primeiro momento, a teoria causal da referência apresenta-se como bastante promissora por abdicar do conteúdo específico das teorias científicas na determinação da referência, evitando então os problemas do holismo e da incomensurabilidade semântica. No entanto, esta vantagem converte-se rapidamente em desvantagem por implicar a manutenção da referência a qualquer custo, deixando de lado a acurácia descritiva acerca dos inobserváveis e tornando as teorias imunes à falha referencial. Some-se a isso as dificuldades que acabamos de apontar, ligadas à própria noção de causalidade, e concluiremos que a teoria causal também não parece ser uma alternativa muito satisfatória para a defesa da manutenção de referência entre teorias.

3.3. Teoria Híbrida

Alguns casos que evidenciam a insuficiência da teoria causal para tratar da relação de referência motivaram a elaboração de uma *teoria híbrida (descritivo-causal) da referência*. Inicialmente introduzida pelos trabalhos de Lewis (1984) e Kroon (1987), tal teoria combina elementos descritivos e causais, seja para a delimitação da referência de um termo ou para sua difusão em uma comunidade linguística.

No caso dos nomes próprios, a teoria híbrida delimita que um nome próprio refere ao objeto ao qual foi inicialmente associado e, além disso, as descrições atreladas a este nome e que caracterizam tal objeto também devem ser satisfeitas para assegurar uma referência bem-sucedida. De acordo com essa proposta, o nome *Aristóteles* refere a um dado indivíduo não só porque foi atribuído a ele, mas também devido à satisfação das descrições: pupilo de Platão, professor de Alexandre o Grande, etc. Se um indivíduo não teve esse nome atribuído a si e não satisfaz tais descrições, ele não será o referente do termo, pois as duas condições essenciais para a referência não foram cumpridas. Esta proposta também é afetada pelo problema da multiplicidade de descrições (assim como a teoria descritivista, subseção 3.1), pois várias descrições podem ser relacionadas a um nome por diferentes pessoas, de modo que é difícil delimitar quais descrições são mais importantes na determinação da referência do nome. Até mesmo quando o aspecto descritivo é combinado com o aspecto causal ele ainda pode gerar certos problemas. De todo modo, essa combinação é eficiente em impedir que a mera satisfação de descrições resguarde a referência (como ocorria indesejavelmente no

exemplo de *Gödel*, uma das críticas à teoria descritivista, subseção 3.1) ou que erros categoriais sejam cometidos na tentativa de referir a um indivíduo (tal como indicado pelo problema *qua*, uma das críticas à teoria causal, subseção 3.2).

Esta alternativa pode ser interessante quando aplicada aos termos teóricos, pois visa igualmente suplantar as deficiências das teorias descritivista e causal nesse âmbito de aplicação. Passemos diretamente ao desempenho desta teoria quando aplicada aos termos teóricos.

3.3.1. Adaptação da teoria híbrida aos termos teóricos

A aplicação direta da teoria causal aos termos teóricos nos indica que o mero recurso ao papel causal de uma entidade inobservável não parece ser suficiente para fixar a referência desses termos. Nesse ponto, as descrições poderiam ajudar a qualificar adequadamente o papel causal de um dado inobservável, de modo a evitar a trivialização da referência do termo. Dada essa constatação, a teoria híbrida parece ser uma alternativa interessante para o tratamento dos termos teóricos, de modo a evitar os problemas das outras teorias da referência nesse âmbito e oferecer uma estratégia mais robusta para o realismo científico na defesa da continuidade entre teorias.

Cientes do papel das descrições para uma ligação mais adequada dos termos teóricos às entidades inobserváveis, filósofos como Enç (1976), Nola (1980), Sankey (1997), Bird (2000) e Psillos (1999; 2012) defenderam a adoção da teoria híbrida da referência. Segundo essa concepção, dados certos fenômenos que suscitam a introdução de um termo teórico, a fixação da sua referência será realizada mediante a apresentação de descrições que delimitem o poder causal da respectiva entidade inobservável, de modo que essas descrições causais possam intermediar a referência desse termo teórico. As descrições fixadoras da referência precisam ser satisfeitas para que a referência seja bem-sucedida. Não só a ligação causal com a entidade é indispensável para a referência, como também o são as descrições da suposta causa. De modo geral, segundo a teoria híbrida, a referência de um termo teórico seria assegurada por descrições mínimas que delimitariam o papel causal da entidade inobservável na produção dos fenômenos em consideração.

Há uma controvérsia sobre se as propriedades que contribuem para o poder causal das entidades inobserváveis deveriam ser fundamentas em propriedades de tipos naturais (categorias naturais) ou não, aspecto esse que divide os defensores da teoria híbrida nesse contexto de aplicação. Enç (1976) e Psillos (1999) defendem que sim; os que abdicam do

papel dessas propriedades categoriais são Nola (1980), Sankey (1997) e o mesmo Psillos, em artigo posterior (Psillos, 2012). No presente trabalho apresentaremos ambas as versões da teoria híbrida, seguindo de perto a formulação de Psillos, uma vez que ele se ocupou mais cuidadosamente dos pormenores da questão. Antes disso, porém, comentaremos brevemente algumas críticas que questionam a denominação “híbrida” desta teoria da referência.

Filósofos como Bird (2000) e Raatikainen (2007) alegaram que, no fundo, a teoria híbrida é uma teoria descritivista. Bird afirmou que a teoria híbrida é, na verdade,

uma teoria puramente descritivista, cuja característica é que algumas das descrições devem estabelecer que a entidade à qual se pretende referir cumpre um papel causal relativamente a alguns fenômenos específicos (BIRD, 2000, p. 135).

Contudo, esta acusação parece equivocada, pois segundo a teoria híbrida a fixação da referência de um termo teórico requer a presença do nexos causal, i.e., a causação de um fenômeno, que motiva a introdução da descrição fixadora do seu referente. Tal exigência não faz parte de uma teoria puramente descritivista, que abrange a possibilidade do aspecto causal constituir o conteúdo de uma descrição, mas não faz disso um pré-requisito para a delimitação da referência de um termo teórico tal como a teoria híbrida. Agora, se o poder causal que fundamenta a referência de um termo teórico existe ou não é uma dificuldade, claro; mas ela não faz, por si só, com que a formulação da teoria híbrida seja puramente descritivista.

Por outro lado, resta a acusação de que a teoria causal já previa o papel das descrições na fixação da referência dos termos teóricos. Porém, tal crítica parece ignorar que na teoria causal a referência é assegurada pela conexão causal do termo com a entidade, e que a descrição só serve para vincular mais particularmente o termo teórico ao seu respectivo inobservável. Nesta concepção, a descrição só auxilia na introdução do termo, mas não interfere posteriormente na referência a ele, pois ela pode até não ser satisfeita pelo referente pretendido e ainda assim a referência estaria assegurada pela conexão do termo com seu referente. Já na teoria híbrida a descrição detalha como o referente está causalmente conectado com os fenômenos produzidos por ele, sendo isso o que garante nosso (suposto) acesso cognitivo a ele. Nesse caso, as descrições apresentadas precisam ser satisfeitas para que a referência seja bem-sucedida. Psillos afastou esta acusação, respondendo à seguinte questão: se estamos errados sobre o referente, mas ainda em contato causal com ele, estaríamos mesmo assim nos referindo a ele? A resposta coerente é que não há referência nestas condições.

Deveria ser claro que o sucesso referencial não é meramente uma questão de contato causal, pois, como já enfatizado, requer-se que os termos teóricos rastreiem seus referentes, algo que é realizado pelas marcas identificadoras dos seus referentes, [...]

o contato causal ancora as descrições ao referente, as descrições habilitam o rastreamento do referente. (PSILLOS, 2012, p. 228).

Já que o mero contato causal não é suficiente para assegurar a referência de acordo com a teoria híbrida, ela não pode ser rotulada como puramente causal. Novamente, a referência só será bem-sucedida, nesta concepção, se as descrições apresentadas forem satisfeitas, aspecto esse que não é exigido pela teoria causal pura.

Também é preciso notar que a teoria híbrida pressupõe, em sua primeira versão, a existência de poderes causais e de tipos naturais, aspectos esses que precisariam ser adequadamente justificados para aqueles que, como os antirrealistas, não se comprometem com tais aspectos ontológicos. Quanto ao aspecto da causalidade, reitero aqui o que foi exposto no fim da subseção 3.2.1., de que esse aspecto ontológico pode ser pressuposto na medida em que é relevante para explicar a ocorrência de fenômenos, está presente de alguma forma no comprometimento antirrealista com as entidades observáveis (uma vez que explica a percepção dessas entidades) e também já é suposto na explicação realista do sucesso da ciência presente no AM. Qualquer que seja o caso, na sequência desta dissertação estaremos assumindo uma posição realista sobre a existência de poderes causais, para efeito de argumento. Sobre os tipos naturais, Psillos, em particular, não se dedicou a defender sua existência;⁴² porém, como ele apresentou uma versão da aplicação da teoria híbrida que não carrega o fardo dos tipos naturais, apresentaremos essa segunda versão posteriormente para avaliarmos se é possível evitar a pressuposição da existência de tipos naturais.

3.3.1.1. Com recurso aos tipos naturais

Em seu primeiro livro sobre o assunto, Psillos (1999) defendeu uma versão da teoria híbrida que pressupunha as *propriedades de tipo* para a delimitação da referência, isto é, propriedades que delimitam se um certo objeto pertence a um tipo ou categoria natural. Por exemplo: de maneira geral, os seres que têm as propriedades de serem *animais*, *mamíferos*, *felinos* e *domesticados* são *gatos*, ou seja, pertencem à categoria natural dos *gatos*. Pela teoria híbrida, a referência de um termo teórico seria assegurada por descrições mínimas que

⁴² Na ocasião, Psillos limitou-se a referenciar a defesa da existência de tipos naturais feita por outros filósofos na nota 4 do cap. 12 do seu livro de 1999: “não é minha intenção fazer uma diatribe sobre os tipos naturais. Remeto ao trabalho de Kornblith (1993, Capítulo 3) para a defesa requerida. Para algumas sugestões semelhantes, o leitor pode ver Boyd 1989; 1991; 1993. O que Boyd sublinhou corretamente é que a ideia de tipos naturais pode ser explicada pelo conceito de “conjuntos de propriedades homeostáticas”. O que torna um tipo natural é a presença de um mecanismo homeostático que provoca e sustenta a ocorrência simultânea de um certo número de propriedades” (PSILLOS, 1999, p. 304n).

delimitariam o poder causal de uma respectiva entidade inobservável. À época, Psillos denominou essas descrições como *descrições causais centrais*. Nesta primeira versão da teoria, Psillos ressaltou a necessidade de que as descrições causais centrais fossem baseadas nas propriedades de tipo da entidade. Essa exigência foi ressaltada por Berent Enç, ao dizer que,

o fardo da referência do termo será carregado pelas propriedades constitutivas de tipo atribuídas ao objeto e pelo mecanismo explanatório desenvolvido pela teoria, ao invés de expressões como ‘o que quer que seja responsável pelo fenômeno Φ ’. (ENÇ, 1976, p. 271).

Sob essa influência, Psillos propôs que,

quando uma entidade é postulada, há, normalmente, algum relato de certas propriedades fundamentais — constitutivas do tipo — que esta entidade deve possuir para exercer o seu papel causal... Assim, o termo que é empregado para denotar a entidade postulada é associado a uma *descrição causal central* das propriedades em virtude das quais a entidade exerce o seu papel causal em relação a um conjunto de fenômenos. Na medida que essas propriedades constitutivas do tipo são a origem causal da informação central associada ao termo, então pode-se dizer que o termo refere a esta entidade. (PSILLOS, 1999, p. 84).

Assim sendo, em episódios de substituição teórica, a manutenção da referência entre as teorias será assegurada se os seus termos teóricos partilharem das mesmas descrições causais centrais. Já que são essas descrições que delimitam a referência, quaisquer termos que as pressuponham captariam o mesmo referente. Formalmente, temos que,

1. Um termo t refere à entidade x se e somente se x satisfaz a descrição causal central associada a t .
2. Dois termos t'' e t denotam [referem] a mesma entidade se e somente se (a) seus referentes putativos cumprem o mesmo papel causal com relação a um conjunto de fenômenos; e (b) a descrição causal central de t'' diz respeito às mesmas propriedades constitutivas do tipo da descrição causal central associada a t . (PSILLOS, 1999, p. 285).

Como são somente algumas descrições que determinam a referência, mesmo em uma mudança teórica bastante significativa pode haver — em princípio, dependendo de cada caso histórico específico — a preservação das descrições centrais que asseguram a referência a um dado inobservável. Desse modo, as ameaças do holismo semântico e da incomensurabilidade podem ser afastadas, já que há somente uma exigência mínima de preservação de descrições para a manutenção de referência entre as teorias, fazendo com que a mudança de significado dos termos não implique diretamente a mudança de referência entre as teorias.

Para a correta identificação das descrições causais centrais associadas a um termo teórico em alguma teoria, Psillos estabeleceu que é preciso, 1) um exame cuidadoso das circunstâncias na qual uma entidade inobservável foi postulada e nomeada; 2) uma análise das descrições associadas a uma entidade em virtude do seu papel causal; e 3) um rastreamento

cuidadoso da história da entidade, de modo que mudanças possíveis na sua descrição causal [atribuída] sejam identificadas (PSILLOS, 1999, p. 287).

As críticas filosóficas que serão apresentadas a seguir foram formuladas por Bruno Borge em um artigo de 2017, intitulado “¿Soluciona el descriptivismo causal el problema de la referencia de los términos teóricos?”. Nele, Borge analisa as propostas apresentadas por Psillos (em ambas as versões) e tece críticas relevantes a ambas. Começaremos pelas críticas à versão da teoria híbrida que pressupõe a existência de tipos naturais e, posteriormente, na próxima subseção, analisaremos as críticas do autor à outra versão dessa teoria.

Dado que uma entidade inobservável pode cumprir papéis causais diferentes a depender do contexto no qual ela está inserida, a identificação das propriedades causais da entidade com as suas propriedades de tipo pode fazer com que a entidade instancie mais de um tipo natural (BORGE, 2017, p. 136). Aparentemente, esta crítica não se aplica se considerarmos que a teoria híbrida, tal como formulada nessa primeira versão, contempla esse aspecto sem maiores problemas. Na determinação da referência de um termo teóricojj é possível ir agregando descrições que abranjam os papéis causais descobertos da dita entidade, de modo a captar adequadamente seus aspectos causais que motivam o uso do seu respectivo termo teórico. Isso pode ser realizado tanto no batismo inicial do nome como em estágios posteriores de refinamento da sua referência. Consequentemente, essa entidade instanciará mais de um tipo natural. Só que assim como usualmente reconhecemos que os objetos instanciam mais de um tipo natural — como um gato que pertence às categorias de “animal”, “mamífero” e “felino” —, o mesmo também deve ocorrer com as entidades inobserváveis. Borge admite esta possibilidade, mas assinala que

a distinção entre propriedades causais e não causais (categóricas ou causalmente inertes?) está longe de poder ser estabelecida sem inconvenientes. De fato, há quem pense que o problema consista justamente na duvidosa possibilidade de estabelecer tal distinção. (BORGE, 2017, p. 136).

No entanto, a mera dificuldade em distinguir as propriedades que são ou não causalmente relevantes em um certo contexto não implica a impossibilidade de estabelecer esta distinção para tratar da referência.

Outro alerta feito por Borge é que o refinamento das propriedades causais de uma entidade inobservável — sobretudo se for uma tentativa de abranger todas as suas possíveis interações causais com outras entidades ou nossos instrumentos de detecção — impediria a continuidade de referência em uma teoria posterior (BORGE, 2017, p. 137). Se esse refinamento teórico fosse apenas uma extensão das propriedades causais já estabelecidas, não

haveria maiores problemas: caso as descrições causais centrais fossem mantidas, a referência ainda seria preservada e esse refinamento seria um indicativo de progresso científico. No entanto, se esse refinamento teórico alterasse substancialmente as propriedades causais atribuídas à entidade, não haveria maneira de defender a continuidade de referência, pois as entidades nem seriam classificadas como integrantes de uma mesma categoria. Fato é que tal crítica não pode ser afastada sem que tenhamos um critério explícito para mensurarmos o impacto de uma alteração teórica nas propriedades causais atribuídas à entidade; e isso não é trivial.

A crítica mais forte de Borge a essa versão da teoria híbrida é que as propriedades causais e constitutivas de tipo da entidade podem não coincidir (BORGE, 2017, p. 138). Esse problema decorre da possibilidade de que, “ao determinar o que uma determinada entidade *faz*, poderíamos dispensar as descrições que informam sobre o que essa entidade *é* essencialmente ou *a que classe ela pertence*” (BORGE, 2017, p. 137). Para reforçar sua crítica, o autor nos apresenta o seguinte exemplo fictício sobre *pregos e parafusos*:

Imagine um mundo no qual as cadeiras não são artefatos, mas entidades naturais; além disso, nesse universo específico, pregos e parafusos são objetos inobserváveis, ou seja, sua maneira de interagir com nosso aparato perceptual é tal que não podemos detectá-los a olho nu. Os cientistas desse estranho mundo inferiram com notável acuidade que as partes dessas cadeiras são unidas por pregos e os descreveram como entidades sólidas e alongadas, com uma ponta plana e outra afiada, etc. Eles estão orgulhosos de seu progresso, e com razão, mas estão errados, pois a natureza (ou algum demiurgo com habilidades de carpinteiro) fez as cadeiras com parafusos. Nossos fantásticos cientistas descreveram com precisão todas as propriedades em virtude das quais os pregos manteriam as partes da cadeira unidas, que são as mesmas propriedades em virtude das quais os parafusos fazem o mesmo. No entanto, pregos não são parafusos e, quando eles dizem “pregos”, não estão realmente se referindo a parafusos. (BORGE, 2017, pp. 137-138).

Este curioso exemplo mostra que se as propriedades que contribuem para o papel causal de uma entidade e suas propriedades de tipo não coincidirem, é possível usar um dado termo teórico para falar de duas entidades inobserváveis diferentes quando não é essa nossa intenção e nem deveria ser o caso. De fato, não há nenhum fator intrínseco às propriedades causais e propriedades de tipo de uma entidade que assegure a identificação entre elas. E como não há nenhuma argumentação que estabeleça tal identidade, é difícil sustentar que o componente causal da teoria híbrida seja o mesmo a assegurar que uma dada entidade pertence a um certo tipo, o qual nos permitiria rastrear a sua referência entre transições teóricas.

Vejamos agora se a outra versão da teoria híbrida, que não pressupõe a existência de tipos naturais no tratamento da referência, tem um melhor desempenho quanto a essa série de dificuldades.

3.3.1.2. Sem recurso aos tipos naturais

Psillos (2012) sustentou outra versão da teoria híbrida para tratar da referência dos termos teóricos, que, diferentemente da primeira, não requer que as propriedades causais estejam fundamentadas em propriedades de tipo. Segundo a nova concepção, qualquer propriedade identificadora estável contingente que contribua para o papel causal de uma entidade inobservável seria suficiente para fixar e rastrear a referência de um termo teórico. Psillos justifica a desvinculação das propriedades causais de uma entidade das suas propriedades de tipo, dizendo que,

Esse afrouxamento é importante porque quais propriedades são constitutivas de tipo pode ser um assunto resolvido somente em estágios bastante avançados de investigação, enquanto a referência bem-sucedida pode ser estabelecida em estágios mais iniciais. É esperado que as propriedades identificadoras façam parte das propriedades constitutivas de tipo das entidades postuladas, mas isso não é algo que tem de ser assumido *ab initio* ou exigido no fim. Na realidade, isso permite o refinamento da referência. (PSILLOS, 2012, pp. 227-228).

Assim, termos de teorias diferentes compartilharão da mesma referência ao pressuporem as mesmas descrições causais acerca de uma entidade inobservável, mesmo que as propriedades de tipo atribuídas a elas não sejam as mesmas.

Psillos considera que a sua nova versão da teoria híbrida tem as seguintes vantagens:

- 1) Não torna inevitável a continuidade referencial;
- 2) Explica de maneira não trivial o sucesso e a falha referencial;
- 3) Permite uma determinação mais detalhada do referente por meio da adição de descrições;
- 4) Deixa aberta a possibilidade de reidentificar a referência de um termo, quando ocorre por exemplo que $x = y$ mesmo que $D(x)$ [descrição causal de x] seja diferente de $D'(y)$. Isso será possível quando houver uma descrição unificadora que reúna tanto $D(x)$ quanto $D'(y)$;
- 5) Mostra como as descrições $D(x)$ associadas a um termo rastreiam a sua referência: isso ocorre quando as descrições são bem-sucedidas em identificar o agente causal de um dado conjunto de fenômenos (PSILLOS, 2012, p. 230).

Supondo-se que possam ser adequadamente defendidos, os pontos (1), (2) e (3) têm a vantagem de decorrerem automaticamente do componente descritivo da teoria híbrida. De fato, somente se as descrições causais dos termos teóricos forem preservadas durante os episódios de substituição teórica é que a continuidade referencial será assegurada. (1) é uma vantagem relativamente à teoria causal da referência, pois além da menção ao papel causal, na

teoria híbrida é preciso que o componente descritivo seja preservado para que possamos assinalar a manutenção de referência entre as teorias. Do mesmo modo, (2) segue do componente descritivo da teoria, o qual é capaz de explicar como um termo refere ou não refere de maneira bem-sucedida. Já (3) é uma vantagem em relação à teoria descritivista da referência, uma vez que na teoria híbrida a adição de novas descrições a um termo teórico não inibe o compartilhamento da referência com algum outro termo, já que o determinante para a referência é apenas um conjunto mínimo explícito de descrições relevantes; se o núcleo descritivo delimitado é preservado, ainda podemos referir à mesma entidade. As três primeiras vantagens são inerentes às duas versões da teoria híbrida, i.e., quer elas atrelem as propriedades causais às propriedades de tipo ou não.

Exporéi agora as críticas formuladas por Borge às supostas vantagens de (4) e (5), de modo a mostrar que mesmo a nova versão da teoria híbrida padece de sérias dificuldades.

Já que Psillos abdica da menção às propriedades de tipo, (4) indica que duas descrições causais diferentes — $D(x)$ e $D'(y)$ — podem captar a mesma entidade caso sejam subsumidas sob uma outra descrição unificadora. Esta descrição unificadora nos permitiria defender que duas entidades supostamente diferentes — dado que não causam os mesmos fenômenos — são, na realidade, a mesma entidade, uma vez que suas descrições passem a ser englobadas por uma única descrição. Um exemplo que pode ajudar a ilustrar esse ponto é o das ondas eletromagnéticas. Como veremos mais adiante, no capítulo sobre os casos históricos, a teoria eletromagnética unificou os fenômenos elétricos, magnéticos e ópticos. No entanto, inicialmente, o termo *onda eletromagnética* não englobava ondas de luz; posteriormente, quando se constatou que a velocidade das ondas eletromagnéticas era muito próxima da velocidade da luz (calculada a partir das equações de Maxwell), também a luz passou a ser entendida como uma onda eletromagnética. Borge critica (4), ressaltando que, “o conteúdo respectivo de $D(x)$ e $D'(y)$ [...] nunca é suficiente para estabelecer que $x = y$, a menos que se inclua descrições das propriedades constitutivas de classe correspondentes a x e y ” (BORGE, 2017, p. 141). Esta crítica enfatiza que sem a menção às propriedades de tipo não podemos estabelecer uma ligação entre papéis causais diferentes a fim de atribuí-las a uma mesma entidade.

No entanto, a impossibilidade de reidentificar a referência de um termo apenas pressupondo os papéis causais das entidades não descredita completamente a nova versão de aplicação da teoria híbrida, pois argumentaremos que é esperado que entidades diferentes não compartilhem dos mesmos papéis causais; consequentemente, essa distinção entre as entidades não deve significar uma tentativa de defender a identificação entre elas.

Do mesmo modo, quanto a (5), o rastreamento da referência pode ficar comprometido caso a descrição identificadora da entidade pressuponha apenas as propriedades causais prescindindo das propriedades de tipo. Borge alerta para a possibilidade de que,

Se as propriedades de identificação e as propriedades constitutivas de classe não coincidem, pode acontecer que duas classes compartilhem as propriedades de identificação e causem os mesmos fenômenos, mas ainda assim sejam distintas em virtude de alguma propriedade constitutiva de classe não detectada (ou indetectável) que não é responsável pelo papel causal das ditas classes. (BORGE, 2017, p. 142).

Esta crítica mostra que duas entidades podem ter as mesmas propriedades causais e ainda assim não serem idênticas, dado que elas podem pertencer a tipos diferentes. Esta consequência também indica que um termo teórico pode ter uma realização múltipla, pois ele pode referir a entidades que cumprem o mesmo papel causal mas que, no entanto, não pertencem às mesmas categorias existenciais.

Mas parece que as duas críticas precedentes se baseiam muito mais em uma mera possibilidade e não naquilo que realmente ocorre na ciência. Será de fato que duas entidades que pertencem a tipos/categorias diferentes podem compartilhar das mesmas propriedades causais, de modo a proporcionar a realização múltipla de um termo teórico? Pensamos que tal problema não se coloca. Aparentemente, os diversos contextos de produção de fenômenos observáveis atribuídos a uma dada entidade inobservável impedem que todas as suas propriedades causais (relevantes para a determinação da referência) possam ser compartilhadas por uma outra entidade que pertença a uma categoria existencial diferente.

Para reforçar essa consideração, resgatamos o exemplo dos *pregos e parafusos* apresentado no final da seção 3.3.1.1. Neste exemplo, Borge alertava para a possibilidade das propriedades causais de uma entidade não coincidirem com as suas propriedades de tipo. Assim, em um mundo fictício onde os pregos fossem objetos inobserváveis, poderíamos utilizar o termo teórico ‘prego’ em virtude de algumas propriedades atribuídas às entidades que auxiliam na junção das partes das cadeiras desse mundo, como *entidades sólidas e alongadas, com uma extremidade chata e a outra afiada*, etc. No entanto, os cientistas só teriam acertado na descrição das propriedades da entidade que a permite manter as partes das cadeiras juntas, pois nesse mundo o que mantém tal junção não são os pregos como se pensava, e sim os parafusos. Borge conclui que o termo teórico *prego* não refere a um parafuso, porque sua descrição contém apenas a menção a propriedades causais, mas não a propriedades de tipo que poderiam ajudar os cientistas desse mundo a distinguir essas duas entidades e captarem o equívoco de referência.

Porém, se explorarmos melhor o exemplo acima perceberemos que os parafusos não têm as mesmas propriedades causais atribuídas aos pregos, e isso já seria suficiente para distinguir as duas entidades. Além de serem *entidades sólidas e alongadas, com uma extremidade chata e a outra afiada*, os parafusos também têm *rosca*, característica essa que os torna mais eficazes na junção de partes de objetos mais frágeis, como madeiras que são menos sólidas, por exemplo. Ao considerarmos todos os contextos nos quais pregos e parafusos são capazes de unir partes de objetos, torna-se evidente que alguns contextos nos auxiliarão a distinguir tais entidades porque elas trazem à tona a diferença entre seus papéis causais.

Nossa crença na realidade de uma entidade inobservável se reforça quando ela é considerada ser a responsável pela manipulação causal não apenas de um fenômeno, mas de vários deles (preferencialmente de tipos diversos e em situações que não estejam envolvidas na postulação desse inobservável). Como argumentamos acerca do exemplo fictício apresentado por Borge, os pregos causam a junção de peças de madeira do tipo X, mas não do tipo Y, enquanto os parafusos exercem seu papel causal em ambos os casos. Um defensor da existência dos pregos não pode ficar satisfeito com o seu papel desempenhado apenas em X, que motivou a sua postulação, ele precisa ter seu desempenho testado em outros casos, senão sua postulação pode ser apenas uma estratégia *ad hoc*.

A possibilidade apresentada por Borge no seu exemplo é importante e deve ser avaliada, mas por tratar-se de um cenário muito específico ele parece ignorar certas complexidades do mundo real que são relevantes para a presente discussão: em seu exemplo, ele considerou apenas o papel causal das entidades na união das partes de cadeiras sem considerar qualquer outro fenômeno que fosse produzido por elas também. Os diversos contextos de produção de fenômenos indicam que as entidades inobserváveis possuem propriedades causais que as distinguem umas das outras, já que elas não compartilharão do mesmo conjunto de tais propriedades. Desse modo, pensamos que a delimitação e o rastreamento da referência de uma entidade inobservável podem ser pautados apenas em propriedades causais, sem a necessidade do auxílio das propriedades de tipo.

Além das críticas (4) e (5) que ressaltam a importância das propriedades de tipo, Borge questiona se as propriedades estáveis contingentes podem identificar uma entidade em contextos diferentes, nos quais ela não cumpre os mesmos papéis causais, e se essas propriedades estáveis identificadoras podem ser propriedades causais, ameaçando de uma vez por todas o empreendimento construído por Psillos (BORGE, 2017, p. 139). Discordamos do primeiro questionamento, pois se as descrições atribuídas a um termo teórico são aquelas que englobam o papel causal da entidade mesmo em contextos diferentes, é esperado que as

descrições relevantes para a identificação da entidade variem a depender do contexto. Isso não quer dizer que as propriedades contingentes não sejam estáveis, mas sim que elas se apresentam ou tornam-se evidentes à medida que o papel causal delas é requisitado.⁴³ Assim, a depender do contexto (isto é, do papel causal em questão), uma propriedade estável é mais relevante do que outra para a identificação da dita entidade.

Já o segundo questionamento é razoável, dado que assim como Psillos afirmava que as propriedades causais poderiam ser idênticas às propriedades de tipo, sem maiores defesas, nesta nova versão ele também não oferece nenhum argumento para sustentar que as propriedades estáveis de uma entidade possam ser causais e, assim, serem usadas para mediar a referência. Porém, responderemos a esta última crítica reafirmando que uma entidade cumpre diferentes papéis causais a depender do contexto, de modo que é esperado que pelo menos quanto às propriedades causais da entidade exista uma certa estabilidade ao longo do desenvolvimento da ciência, pois sempre que o seu papel causal for requerido sua propriedade causal estará presente, permitindo sua identificação tal como esperado.

Dada essa análise das críticas à versão da teoria híbrida sem tipos naturais, pensamos que ela mostra-se, no fim das contas, mais vantajosa do que a anterior. A abdicação da menção às propriedades de tipo torna-a ontologicamente econômica, evitando também a pressuposição de que as propriedades causais sejam idênticas às propriedades de tipo. A nova versão simplifica a identificação e o rastreamento da referência das entidades inobserváveis. Do ponto de vista teórico, a teoria híbrida da referência que pressupõe apenas as propriedades causais das entidades inobserváveis para a delimitação da referência parece ser, portanto, a mais adequada para a defesa da continuidade entre as teorias.

3.3.2. Crítica à teoria híbrida: projeção realista do presente sobre o passado

Outra vertente de objeção à aplicação da teoria híbrida aos casos históricos realizada por Psillos é a de que ela pode envolver uma projeção realista do presente sobre o passado. Esta é uma objeção não só filosófica, como contém um caráter histórico, o que ameaça a efetividade desta teoria da referência no tratamento dos casos históricos.

⁴³ Aliás, esse esclarecimento acerca das propriedades contingentes não as torna semelhantes às propriedades disposicionais? Do mesmo modo descrito, as propriedades disposicionais são características dos objetos que podem ser realizadas a depender da ocorrência de certas circunstâncias — por exemplo, a fragilidade de uma taça de vidro só é de fato atestada quando por algum descuido ela é estilhaçada. Tais propriedades foram famosamente discutidas por Carnap em “Testability and Meaning” (1936).

No que diz respeito à avaliação da manutenção de referência em episódios de substituição teórica, Psillos supôs que poderíamos nos basear no julgamento dos cientistas para delimitar quais descrições causais dos termos teóricos são centrais ou não. Afinal, quem estaria em melhor posição do que os próprios investigadores do mundo natural para identificar as descrições que intermediariam a referência aos inobserváveis?

A seleção das propriedades que integram o núcleo das descrições causais centrais é guiada por considerações sobre os requisitos das propriedades das entidades postuladas para o desempenho do seu papel causal pretendido com relação a um conjunto de fenômenos. *Isso é certamente limitado pela forma como os cientistas que postularam essa entidade a descrevem.* (PSILLOS, 1999, p. 287, grifo nosso).

Assim, a partir do que seja delimitado pelos cientistas, tornar-se-ia possível fazer uma reconstrução racional das teorias abandonadas a fim de defender uma correspondência das descrições causais com as fornecidas por teorias posteriores.

No entanto, veremos que essa confiança irrestrita na capacidade de os cientistas delimitarem aquilo que é central em uma teoria pode ser objeto de crítica, uma vez que a avaliação dos casos históricos parece justamente descreditar esta capacidade tão importante para tal estratégia realista. Mais especificamente, se esta capacidade de distinção (acerca do que é central) for falha, pode haver uma *projeção do presente sobre o passado* por parte do realista: uma vez que a avaliação das descrições causais centrais — indispensáveis para a defesa da continuidade de referência a um inobservável — é feita de maneira retroativa, o realista pode pressupor aquilo que é central nas teorias abandonadas para garantir a continuidade que ele tanto aspira. Essa não parece ser uma manobra teórica isenta de problemas.

É importante salientar que, para Psillos, a defesa da manutenção de referência entre teorias deve ser realizada após o diagnóstico de que a postulação de um dado inobservável foi um aspecto indispensável ou central para o sucesso de uma teoria — i.e., após a execução da estratégia *divide et impera*. Caso a entidade não tenha sido central, é desnecessário defender a manutenção de referência, já que o comprometimento realista deve se estender somente às entidades que cumprem algum papel no sucesso das teorias. No caso contrário torna-se, aí sim, necessário defender a manutenção de referência através da preservação de descrições causais centrais entre as teorias científicas.

Tanto a execução do *divide et impera* como a aplicação da teoria híbrida da referência envolvem algum tipo de identificação daquilo que é central em uma teoria: seja quanto à pressuposição de um inobservável para a derivação do sucesso de alguma teoria ou quais descrições causais determinam a referência a um dado inobservável. Assim, ambos os

recursos adotados por Psillos dependem de habilidade para a delimitação apropriada da centralidade desses aspectos; somente com a efetividade desta habilidade é que tais estratégias funcionarão. Porém, há dúvidas sobre a efetividade desta habilidade.

No artigo “Pyrrhic Victories for Scientific Realism” (2003) e no livro *Exceeding Our Grasp* (2006), Kyle Stanford critica algumas estratégias realistas seletivas,⁴⁴ como a que é adotada por Psillos. Estratégias de realismo seletivo são aquelas que restringem o comprometimento realista com as teorias científicas, de modo que somente algum tipo de componente teórico acerca dos inobserváveis demande tal comprometimento. Quando se trata da discussão histórica sobre as entidades inobserváveis, uma estratégia realista seletiva sustentará a continuidade somente quanto a uma parcela da teoria. Tanto o *divide et impera* como a teoria híbrida da referência fazem parte de uma estratégia unificada de Psillos, que visa defender o comprometimento apenas com as entidades inobserváveis efetivamente responsáveis pelo sucesso das suas respectivas teorias científicas e, consequentemente, com algumas descrições causais centrais que delimitam como esses inobserváveis produzem certos fenômenos. O *divide et impera* é seletivo na medida em que requer o comprometimento realista só quanto às entidades inobserváveis responsáveis pelo sucesso de alguma teoria. Do mesmo modo, a adoção da teoria híbrida é seletiva porque são somente as descrições causais centrais que exigem o comprometimento realista, uma vez que elas especificam como um inobservável produz certos fenômenos — explicitando o fator de sucesso teórico —, sendo essas descrições as que deverão ser mantidas em episódios de substituição teórica, de modo a garantir a manutenção de referência a um dado inobservável. A crítica de Stanford à estratégia de Psillos visa a descreditar os recursos utilizados por esse filósofo realista por meio do questionamento da habilidade dos cientistas em distinguirem, sem a vantagem da retrospectção, aquilo que é central em uma teoria. Como o teor da presente subseção é acerca da teoria híbrida, nela discutiremos as críticas de Stanford no que diz respeito à delimitação das descrições causais centrais (i.e., à aplicação da teoria híbrida nos casos históricos) — presente em Stanford (2003, pp. 555-563; 2006, pp. 147-155). Tais críticas, porém, são empregadas de modo similar ao *divide et impera*, já que em ambos os casos há a necessidade de delimitar aquilo que é central — em um caso são as descrições, no outro são as entidades inobserváveis.

⁴⁴ Além de Psillos, são criticadas as estratégias de Hardin e Rosenberg (1982), Philip Kitcher (1993), Jarret Leplin (1997) e John Worrall (1989). Uma análise mais recente das estratégias seletivas pode ser encontrada em Harker (2010).

A crítica de Stanford a esse novo tipo de realismo é que pela via da seletividade o realismo obtém apenas uma “vitória pírrica” — i.e., uma vitória conquistada a altos custos. Nesses moldes, um realista sustenta a verdade apenas parcial das teorias científicas perante o desafio dos casos históricos de abandono teórico ao prescindir consideravelmente do realismo científico e ceder bastante ao antirrealismo. Inicialmente, os realistas recorriam ao sucesso das teorias científicas para convencer-nos que elas eram bastante acuradas; agora, eles querem convencer-nos de que esse sucesso decorre apenas de uma parte da teoria, e de que somente esta seria acurada e demandaria o nosso comprometimento. Segundo Stanford,

[o realista] é forçado a conceder [...] ao seu oponente apenas os pontos substantivos que estão em disputa entre eles ou tudo que o antirrealista precisa para um caso histórico convincente contra o realismo. Embora as posições resultantes possam ser realistas apenas no nome, o ponto está longe de ser meramente terminológico: em cada um dos casos [...], aqueles que visam a usar o registro histórico para defender o realismo parecem estar inconscientes de que sacrificaram os princípios substanciais da posição realista no altar do seu nome. (STANFORD, 2003, p. 555).

Se, por um lado, o realista cede ao antirrealismo que o sucesso teórico não é suficiente para o comprometimento total com a teoria, por outro, ele ainda sustenta o comprometimento com alguma parte da teoria. Obviamente, o antirrealista pode criticar tal alteração, que diminui as implicações de uma posição realista. Por outro lado, porém, a alteração também pode ser encarada como um refinamento por parte do realista que tenha reconhecido as limitações da sua posição, dispondo-se a melhorá-la. A interpretação dessa mudança de atitude, a depender da perspectiva, pode ser encarada a favor ou contra o realismo.

Quanto à adoção da teoria híbrida por Psillos, a vitória pírrica envolveria, na verdade, uma manobra *ad hoc*: as descrições para a manutenção da referência poderiam ser escolhidas de modo artificial simplesmente para garantir a continuidade defendida pelo realista. Além disso, pode gerar um novo tipo de metaindução pessimista acerca da nossa habilidade em distinguir quais descrições causais são centrais.

A acusação de manobra *ad hoc* pode, talvez, ser afastada mediante um estudo cuidadoso dos casos históricos. Equívocos sempre são possíveis; mas se o próprio Psillos já havia alertado sobre a necessidade da realização de uma análise das descrições que seja historicamente fiel, ele não pode ser acusado de apresentar uma proposta sem as devidas ressalvas acerca da possibilidade *ad hoc*. O próprio Stanford concede que até podemos “deixar esta preocupação de lado e permitir que o realista titule [delimite] delicadamente as descrições causais centrais que ele associa aos termos cruciais das teorias passadas bem-sucedidas [...]” (STANFORD, 2006, p. 151). No entanto, Stanford alerta que, ainda assim a vitória de Psillos permanecerá sendo pírrica (STANFORD, 2006, p. 151).

A defesa da continuidade teórica com o auxílio da teoria híbrida, tal como defendida por Psillos, requer que os cientistas sejam capazes de identificar quais são as descrições causais centrais tanto das teorias passadas quanto das teorias atuais. A partir do julgamento desses especialistas é que avaliaríamos a continuidade nos casos históricos. Apesar de Psillos apresentar um critério para delimitar quais componentes teóricos requerem um comprometimento realista, esse critério, que consiste na delimitação das descrições causais centrais a partir de um estudo histórico cuidadoso, pressupõe uma habilidade humana que pode inexistir, ou ser falha. Os próprios casos históricos podem lançar dúvida sobre a habilidade dos cientistas em delimitar essas descrições tão importantes para o realismo científico.

Os pretendidos casos de manutenção de referência parecem, portanto, depender de uma leitura seletiva da história da ciência, que projeta no passado as descrições causais de interesse realista para a defesa da continuidade. Tal parece ser o caso do éter no século XIX, analisado pormenorizadamente no livro de Psillos. A projeção de descrições no passado por parte do realista, por si só, põe em questão a legitimidade da sua estratégia, pois é uma movimentação que visa defender a centralidade de certas descrições somente quando favorece o realismo científico.⁴⁵ A projeção realista no passado também acaba gerando um novo tipo de MIP. Se, contrariamente ao que foi argumentado inicialmente pelo realista, as descrições causais ditas centrais diferem daquelas que foram delimitadas pelos principais cientistas defensores de uma dada teoria, então os cientistas não parecem ser capazes de delimitar efetivamente tais descrições. Contudo, para defender a manutenção da referência pela correspondência das descrições causais de um termo abandonado com as de um outro termo adotado atualmente, o realista precisaria saber se estas descrições também são admitidas como centrais na atualidade. Para isso, bastaria que ele recorresse ao julgamento dos cientistas. Se, no entanto, ele divergiu dos cientistas quanto às descrições do passado, como pode justificar que os cientistas acertaram no presente? E caso o realista insista no seu próprio julgamento daquilo que é central ou não, como poderemos confiar na sua habilidade para distinguir isso, ao invés de suspeitar que ele está selecionando aquilo que melhor lhe convém? Ao que parece, os “realistas seletivos não estão em posição de indicar quais postulados de uma teoria atualmente aceita são os genuinamente funcionais a serem retidos no futuro e quais

⁴⁵ Harker (2010) apresentou uma proposta de realismo progressivo seletivo, que, segundo ele, teria a vantagem de evitar o problema da projeção realista no passado. Nesta proposta, o comprometimento realista se estende aos elementos responsáveis pelo progresso científico que foi promovido por uma teoria. Harker argumenta que como podemos identificar claramente quais elementos foram responsáveis pelo progresso em um caso histórico, então não haveria espaço para projeções realistas sobre o passado.

são meramente pressupostos suscetíveis de serem abandonados no futuro” (WRAY, 2017, p. 42). Desse modo, a projeção realista no passado gera um novo tipo de MIP nesse sentido: se anteriormente a MIP resultava no questionamento da conexão entre o sucesso e a verdade aproximada das teorias científicas, dado o relato dos casos históricos, agora a nova MIP resulta no questionamento da nossa habilidade (sejam os cientistas ou filósofos) em discriminar quais descrições são centrais ou não para a defesa da manutenção da referência entre teorias.

A razão é que esse caso do status referencial dos termos centrais nas teorias bem-sucedidas do passado simplesmente nos convida, a partir do registro histórico, a uma forma renovada de metaindução pessimista, a qual desta vez diz respeito à nossa habilidade de distinguir (no tempo em que uma teoria prevalece) quais das nossas crenças sobre uma entidade *atualmente* fazem parte das suas descrições causais centrais. (STANFORD, 2003, p. 559).

* * *

A teoria híbrida da referência é uma alternativa que, uma vez adequadamente desenvolvida e contornadas as críticas apontadas nas seções precedentes, pode vir a mostrar-se promissora, na medida em que contribua para equilibrar os elementos causais e descritivos para a determinação da referência. Ela poderia evitar os problemas do holismo e da incomensurabilidade semântica, além de não desmerecer a acurácia descritiva sobre os inobserváveis e nem trivializar o acerto referencial. Como vimos, a versão da teoria híbrida que pressupõe propriedades causais parece ser a mais eficaz para o rastreamento e preservação da referência entre teorias, candidatando-se então como a melhor alternativa para a defesa da manutenção de referência entre as teorias. Contudo, a acusação de projeção realista do presente sobre o passado leva à crítica de leitura seletiva da história da ciência e ao questionamento da nossa habilidade em determinar as descrições causais centrais que são indispensáveis para a defesa da manutenção da referência entre teorias.

4. Teorias da referência perante casos históricos

Vamos examinar qual seria o desempenho das teorias da referência quando aplicada aos casos históricos selecionados para estudo neste trabalho: o éter luminífero e o flogisto. Como são casos muito discutidos na literatura, aqui serão expostos de forma um tanto breve, para depois explorarmos algumas das consequências da aplicação das teorias da referência a tais casos, na esperança de que contribuam para a avaliação comparada do desempenho da aplicação das teorias da referência.

Aprofundaremos um pouco mais nos casos do éter luminífero e do flogisto, com isso, teremos a oportunidade de conferir efetivamente como as teorias da referência são aplicadas aos casos históricos, o que nos permitirá explorar algumas consequências prévias dessa aplicação já anunciadas na apresentação das teorias da referência (capítulo 3). Veremos como a teoria descritivista é afetada pelo holismo e incomensurabilidade semântica, a teoria causal pelas implicações indesejadas da trivialização da referência e, a teoria híbrida, não só tem o problema da projeção realista sobre o passado assim como sua proposta de equilíbrio entre as duas teorias da referência precedentes é colocada à prova.

4.1. O éter luminífero

A suposição da existência do éter — entidade inobservável material que preencheria todo o espaço — remonta à Antiguidade e passou por diferentes reformulações ao longo do tempo. Aristóteles (384-322 a.C.), que rejeitava a noção de espaço vazio, pressupôs a existência do éter caracterizando-o como uma substância invisível e sutil presente nas partes do espaço denominadas “mundo supralunar” (onde estariam localizados os corpos celestes). Já na modernidade, com o advento de teorias que conceberam a luz como um fenômeno de natureza ondulatória (Christiaan Huygens, no séc. XVII; Augustin Fresnel, no séc. XIX), parecia ser necessária a existência de um meio subjacente para a propagação das ondas luminíferas; presumiu-se então que esse meio seria o éter, que passou a ser denominado *éter luminífero*.

No século XIX o campo da óptica passava por uma revolução: a luz agora era vista como uma ocorrência de natureza ondulatória, em vez de corpuscular, como se acreditava no século precedente, pela ampla aceitação da teoria corpuscular newtoniana da luz. Dada essa mudança, a luz passou a ser caracterizada como uma perturbação (ou vibração) que se propagaria através de um meio, uma vez que à época as únicas ondas conhecidas eram

de natureza mecânica, que requeriam um meio de propagação (tal como as ondas sonoras, por exemplo). A classificação da luz como um fenômeno de natureza ondulatória foi um resultado das pesquisas de Thomas Young (1773-1829) e Augustin Fresnel (1788-1827), que ofereceram uma explicação mais robusta dos fenômenos de interferência, difração e polarização da luz ao tratá-la como uma onda. Essa explicação exitosa também dependia da classificação da luz como uma onda de tipo transversal, i.e., uma onda que oscila perpendicularmente à sua direção de propagação. E é nesse ponto que o éter luminífero se torna um protagonista: por ser um sólido elástico que preencheria todo o espaço, ele seria o meio que viabilizaria a propagação de ondas transversais, como a luz, segundo a teoria de Fresnel (mas não na sua predecessora, de Huygens). Por ser um sólido elástico, esse éter seria rígido o suficiente para resistir a qualquer coisa que pudesse distorcer a sua forma, sendo esse o fator que tornava transversais as ondas que se propagassem nele; e, por outro lado, deveria ser sutil o suficiente para permitir que os corpos celestes pudessem se movimentar livremente (WHITTAKER, 1910, p. 137).

Nesta época vigorava a concepção mecanicista do mundo físico, a qual rejeitava a noção de ação à distância.⁴⁶ Devido a isso, a explicação da propagação de fenômenos de natureza ondulatória reforçava a pressuposição de algum meio, tornando o éter luminífero mais indispensável ainda.

Posteriormente, ainda no século XIX, a luz foi classificada como uma onda eletromagnética por ter uma velocidade similar às ondas eletromagnéticas, satisfazendo as equações de Maxwell — formuladas por James Clerk Maxwell (1831-1879), o principal desenvolvedor da teoria eletromagnética. Com isso, tal como os fenômenos elétricos e magnéticos, os fenômenos ópticos (relativos à luz) passaram a ser entendidos como fenômenos eletromagnéticos, sendo então explicados pela teoria eletromagnética. O éter continuou a ser considerado um fator indispensável para a propagação da luz, pois dada a natureza transversal das ondas eletromagnéticas e a concepção mecanicista do mundo ainda vigente à época, o éter aparentava ser um meio crucial para a propagação dos fenômenos eletromagnéticos. A esta altura, quando Maxwell incorporou a óptica ao eletromagnetismo, o éter luminífero passou a ser denominado *éter eletromagnético*. Do mesmo modo que o

⁴⁶ A rejeição da noção de ação à distância foi compartilhada por muitos personagens ilustres da ciência, dentre eles, Isaac Newton, como exposto no seguinte trecho: “[...] é inconcebível que a matéria inanimada bruta possa, sem a mediação de alguma outra coisa que não seja material, operar sobre e afetar outra matéria sem contato mútuo[...]. [...] que um corpo possa agir a distância sobre outro através do vácuo, sem a mediação de alguma outra coisa, pela qual sua ação e força possam ser transportadas para o outro, é para mim um absurdo tão grande, que creio que nenhum homem que tenha em assuntos filosóficos uma faculdade competente de pensar possa jamais nele incorrer” (Newton *apud* Cajori, 1934, pp. 633-634).

anterior, esse éter foi caracterizado como um meio material e mecânico que viabilizaria a propagação dos efeitos eletromagnéticos, incluindo-se a luz.

No decorrer do século XIX foram executados diversos experimentos para detectar indiretamente a existência do éter, os quais tentavam verificar o movimento da Terra em relação ao éter. Todos, porém, apresentaram resultados negativos; o mais famoso deles foi o de Michelson-Morley.⁴⁷ E no início do século XX o advento da teoria da relatividade levou ao abandono da concepção do éter em associação ao eletromagnetismo.⁴⁸ Atualmente, as ondas eletromagnéticas são concebidas como o resultado das oscilações de campos eletromagnéticos, os quais não pressupõem mais o éter para a propagação dessas ondas do mesmo tipo que a luz. A questão que fica é se os campos eletromagnéticos não seriam, em certo sentido, os sucedâneos ontológicos do éter, já que têm com ele várias características comuns. Seja como for, o abandono do éter luminífero, tal como originalmente concebido, envolveu, pelo menos, a rejeição da hipótese de que ele seria um meio mecânico (no sentido da mecânica do século XIX). Mas a destituição do caráter mecânico do meio de propagação dos fenômenos eletromagnéticos, i.e., aquilo que é admitido atualmente, a saber, o campo eletromagnético, precisa ser mais bem esclarecida, uma vez que o campo eletromagnético é admitido como sendo material em alguma medida:

[...] a concepção contemporânea de campo eletromagnético atribui a ele um conteúdo de massa-energia bem definida (local) e, portanto, pode-se dizer que isto o qualifica como uma “substância material” em algum sentido atual desse termo [i.e., o termo campo eletromagnético]. Mas o sentido em questão está muito distante dos requisitos que Maxwell tinha em mente ao insistir que o éter deveria ser um “meio ou substância” ou ao negar que a energia poderia residir em qualquer lugar exceto na substância mais íntima das coisas materiais” (STANFORD, 2003, p. 560n).

Após o século XIX, houve uma gradual modificação do conceito de matéria, em que ela foi deixando de ser considerada exclusivamente como tendo um aspecto puramente mecânico; desse modo, diferentemente do éter luminífero, o campo eletromagnético é material (transporta energia, produz efeito sobre outros corpos, etc), mas não é classificado como mecânico, na acepção clássica desse termo, estabelecida pelos filósofos naturais do século XVII.

Assim, na análise de cada teoria da referência, a tentativa de defesa da manutenção de referência entre os termos *éter luminífero* e *campo eletromagnético* dependerá da consideração do aspecto mecânico atrelado ao primeiro termo.

⁴⁷ Maiores detalhes em Martins (2012).

⁴⁸ Veja Martins (2006).

4.2. O flogisto

O flogisto era uma entidade inobservável da química, proposta no século XVII por Georg Ernst Stahl (1659-1734) e defendida por Joseph Priestley (1733-1804) para explicar o fenômeno da combustão. Segundo a teoria do flogisto, os materiais inflamáveis seriam permeados por essa entidade inobservável que, ao ser liberada no ar, causaria sua combustão e os tornaria mais leves após o processo de queima. A teoria do flogisto foi amplamente aceita até o final do século XVIII quando ocorreu a chamada “revolução química” protagonizada por Antoine Lavoisier (1743-1794).⁴⁹

Segundo Joseph Priestley, quando o ar atmosférico estivesse demasiadamente carregado de flogisto, ele não seria adequado para a respiração e impediria que alguma combustão ocorresse naquele ambiente específico, de modo que o cessar da combustão de um corpo era derivado da liberação do seu componente flogístico no ar. Em contrapartida, o ar purificado de flogisto, e que previamente à combustão do corpo permitia a sua queima, foi denominado por Priestley como *ar desflogisticado*: um ar bom para respirar e bom para que ocorra combustão no ambiente.

Em clara oposição a essas hipóteses, Antoine Lavoisier publicou em 1789 a obra *Tratado Elementar de Química*, um novo compêndio e manual teórico propondo uma revolução, que ficou conhecida graças à alguns estudiosos como a *Nova Química*, que romperia drasticamente com o que os teóricos do flogisto acreditavam até então e também conseguia estabelecer relações com outro campo da química que era desconsiderado por esses teóricos, a saber: a química pneumática.

Quanto ao fenômeno da combustão, Lavoisier propôs que a sua causa residiria na reação de um material inflamável com um outro tipo de entidade, — igualmente inobservável, note-se — o oxigênio, presente na atmosfera. Essa nova explicação da combustão foi corroborada por experimentos cuidadosos realizados por Lavoisier e seus colaboradores, que permitiram que eles medissem com precisão a quantidade de oxigênio que seria consumida durante a combustão. Com a ampla aceitação dessa teoria, a existência do flogisto foi abandonada.

⁴⁹ A apresentação que se seguirá do caso histórico do flogisto, sobretudo nesta seção 4.2 e na subseção 4.5.2, é baseada na publicação do meu artigo escrito em parceria com Gabriel Sardi intitulado “Realismo, continuidade teórica e a revolução química” (Sardi; de Oliveira Silva, 2022).

A proposta de Lavoisier, além de ser mais fácil,⁵⁰ didática⁵¹ e capaz de explicar a totalidade de fenômenos que a teoria do flogisto explicava, mas sem a necessidade de recorrer ao uso dessa entidade inobservável, também resolvia uma série de problemas internos que a teoria do flogisto acumulou com o passar do tempo, tal como a questão do aumento de peso de alguns metais após a combustão (o que poderia ser considerado incoerente, visto que o metal perderia um componente – o flogisto – e deveria, conseqüentemente, perder peso).⁵² Esses foram alguns dos fatores que fizeram com que as teses científicas em defesa da existência do flogisto (e toda sua estrutura explicativa geral) fossem perdendo adeptos gradativamente e fazendo com que a proposta explicativa de Lavoisier ganhasse mais espaço nos círculos intelectuais e científicos da época.

Mas considerar que a teoria de Lavoisier simplesmente trocou o termo teórico *flogisto* pelo *oxigênio* não é um retrato correto desse episódio histórico. Para que o abandono do *flogisto* e o surgimento do *oxigênio* ocorresse, foi necessária toda uma reformulação da química vigente na época, apresentando uma nova estrutura conceitual e, conseqüentemente, reformulando o significado dos termos teóricos imersos na rede teórica, mesmo que alguns ainda possuíssem nomes que remetessem à antiga teoria do flogisto.

Fato é que, a pressuposição da existência do flogisto levou a algumas previsões inéditas, como a descoberta de novos tipos de ácido e que o aquecimento de alguns tipos de ácido [calx = óxido metálico] no ar inflamável gera metal puro. Outro ponto interessante é que no caso do flogisto parece haver uma menção indireta ao oxigênio (o real causador da combustão) a partir do uso de outro termo teórico desta teoria. Isso ocorreu quando Priestley — um dos proponentes da teoria do flogisto — usou o termo *ar desflogisticado* para se referir ao ar bom de se respirar e à substância obtida quando a substância emitida na combustão (flogisto) é removida do ar; ademais, ele também mencionou que as plantas desflogisticam o ar. Ora, o oxigênio é bom para a respiração e é emitido pelas plantas, então, ao enunciar as expressões citadas, Priestley parecia estar em contato com o oxigênio. Além dessa similaridade explicativa com o oxigênio, a suposição do ar desflogisticado também cumpriu um papel importante na teoria do flogisto no que diz respeito à unificação de fenômenos. Como James Ladyman nota,

⁵⁰ A tese do oxigênio era mais acessível na medida em que a obra de Lavoisier era também um manual que unificava o vocabulário, a metodologia e a cosmovisão dos químicos da época.

⁵¹ Em o *Tratado Elementar de Química*, Lavoisier explicava o passo a passo de todos os seus experimentos, além de conter gravuras dos equipamentos e das etapas a fim de ilustrar como os resultados eram obtidos e permitir que outros químicos também os executassem. As gravuras foram obras da esposa de Lavoisier, a madame Marie-Anne Pierrette Paulze.

⁵² Algumas explicações, que beiravam o absurdo, foram propostas, como, por exemplo, o flogisto *de peso negativo* (LABINGER e WEININGER, 2005).

O aspecto mais importante da teoria do flogisto é o modo como ela unificou diferentes processos categorizando-os em termos dos recíprocos de flogisticação e desflogisticação. A combustão e a calcinação envolvem a desflogisticação do combustível ou metal e a flogisticação do ar, e a adição de um ácido ao metal envolvia a desflogisticação do metal para sair da base. Essa unificação é retida na química contemporânea pela dualidade entre reações de oxidação e redução. (LADYMAN, 2011, p. 89).

A teoria do flogisto possibilitou a realização de predições inéditas e unificações de fenômenos, portanto, ela foi bem-sucedida; sendo que a pressuposição da existência do flogisto e do ar desflogisticado contribuíram para a geração desse sucesso.

Porém, por mais que o termo *ar desflogisticado* tenha sido utilizado para descrever fenômenos que atualmente relacionamos com o oxigênio, a própria caracterização do seu referente pressupõe a existência do flogisto. Então considerando o aspecto ontológico da situação, se não for possível defender a manutenção da referência do termo *flogisto*, como dizer que o termo *ar desflogisticado* refere dada a sua pressuposição de uma entidade inexistente? Caso esse seja o caso, como apontado por Ladyman (2011, p. 95), se o termo original da teoria não refere (*flogisto*), o que dizer do seu termo derivado (*ar desflogisticado*)?

É preciso analisar, de acordo com cada teoria da referência, se há uma manutenção de referência do termo *flogisto* a fim de extrairmos conclusões sobre a referência do termo *ar desflogisticado*. Neste caso, o sucesso ou insucesso quanto ao primeiro termo afetará significativamente os resultados da análise do seu termo derivado.

4.3. Teoria descritivista

Segundo a teoria descritivista da referência, um termo teórico putativamente refere-se aos objetos que satisfaçam suas descrições. Para que exista uma manutenção de referência a um dado inobservável entre duas teorias, é preciso que seus termos teóricos i) compartilhem das mesmas descrições ou que ii) suas descrições sejam consistentes entre si e captem o mesmo referente. O cenário (i) é ameaçado pelo holismo semântico, pois como resultado das alterações ocorridas em episódios de substituição teórica haveria uma mudança de significado entre os termos teóricos das teorias, inviabilizando o compartilhamento das mesmas descrições. E parece ser inatingível o cenário (ii), porque a consistência entre descrições diferentes só será capaz de garantir a manutenção de referência caso a diferença entre elas não implique a caracterização de uma entidade muito diferente.

Além desses dois riscos iniciais à aplicação da teoria descritivista aos casos históricos típicos, temos ainda a incomensurabilidade semântica: a falta de aspectos semânticos em

comum nas linguagens das teorias entre transição de paradigmas pode dificultar a comparação entre elas.

Analisando o caso do éter luminífero, pelo menos uma descrição não é mantida entre os termos teóricos *éter luminífero* e *campo eletromagnético*: o suposto meio material mecânico que preenche todo o espaço. É justamente o abandono dessa descrição, que indicava a existência de um meio material que naquela época era classificado como um ente mecânico, que destituiu o meio de propagação dos fenômenos eletromagnéticos do seu caráter mecânico. Em um panorama descritivista, o termo *campo eletromagnético* não pode capturar a mesma entidade que o termo *éter luminífero* dada a diferença substancial entre as descrições que os caracterizam, inviabilizando assim uma manutenção de referência neste caso histórico. O aspecto mecânico era uma suposição importante da teoria do éter luminífero, e foi justamente o abandono dessa suposição que afetou sua manutenção de referência com a teoria posterior. Temos aqui uma consequência indesejada do holismo semântico, já que um determinado fator teórico acabou afetando o significado de um termo e, conseqüentemente, sua referência. E dada a falta de continuidade teórica neste caso, a comparação entre teorias é dificultada, fazendo surgir também a incomensurabilidade semântica.

No caso do flogisto, ambos os termos *flogisto* e *ar desflogisticado* tem uma discrepância descritiva muito grande em comparação com o termo *oxigênio*, de modo que a teoria descritivista falha novamente na manutenção da referência em um caso histórico. Não é mantido na concepção do oxigênio a descrição referente ao flogisto: uma substância emitida na atmosfera durante o processo de combustão. E essa descrição acaba afetando o termo *ar desflogisticado* apesar da sua semelhança explicativa com o oxigênio, pois uma das descrições que caracterizava seu referente era: substância obtida quando a substância emitida na combustão (flogisto) é removida do ar. Como tal descrição pressupõe a existência de uma entidade abandonada, ela não pode ser satisfeita, impedindo a correferência com o termo *oxigênio*. O holismo é muito bem representado neste caso: como o significado do termo *ar desflogisticado* era estritamente associado ao *flogisto*, bastou que este fosse considerado não referente para que a descrição daquele fosse incapaz de denotar algo. Este é um ótimo exemplar de como a dependência de significado entre os termos pode afetar sua referência. Novamente, a incomensurabilidade semântica ressurge pela falta de continuidade teórica neste caso.

De modo geral, a teoria descritivista falhou ao ser aplicada aos casos históricos analisados pela sua exigência excessiva de satisfação de descrições, razão que a faz ser enfraquecida pelos problemas do holismo e incomensurabilidade semântica. Certamente, as

descrições não podem ter tanto peso determinação da referência para viabilizar a continuidade entre teorias.

4.4. Teoria causal

A teoria causal busca evitar os problemas inerentes à pressuposição de descrições para a referência, pois é o (suposto) poder causal da entidade inobservável que determinaria a referência de um termo teórico. Basta que duas teorias tenham sido elaboradas para o tratamento do mesmo domínio de fenômenos para pressuporem o mesmo agente causal como referente, assegurando então a manutenção da referência em episódios de substituição teórica. O holismo semântico é afastado porque o significado de um termo não vai mediar sua referência e a incomensurabilidade semântica é solucionada pela manutenção de referência entre teorias que abrangem o mesmo domínio de fenômenos.

Segundo essa teoria, admitiremos que o termo *éter luminífero* capta o mesmo referente que o termo *campo eletromagnético* dado que ambos fazem menção à entidade que permite a propagação da luz, apesar de apenas o éter ser considerado um meio material mecânico que permearia todo o espaço. Do mesmo modo, os termos *flogisto*, *ar desflogisticado* e *oxigênio* serão denominados correferentes por nomearem a entidade que é responsável pela combustão e a respiração. Embora o flogisto e ar desflogisticado tenham sido caracterizados como entidades opostas e que sequer estavam presentes no mesmo ponto do espaço — como é exposto na descrição associada ao ar desflogisticado, segunda a qual, ele seria a substância obtida quando a substância emitida na combustão (flogisto) é removida do ar —, a divergência dos fenômenos que ambos eram supostos produzir acaba sendo ignorada pela identificação de seus papéis causais com os do oxigênio. Não importa quanta alteração teórica tenha ocorrido em qualquer desses casos históricos, a teoria causal nos permitiria ignorar as discrepâncias descritivas e assegurar facilmente a manutenção da referência. Mas tal desfecho não é satisfatório; assim como é injustificado admitir que as explicações da ocorrência dos raios oferecidas pela mitologia grega (que remete a Zeus) e a ciência (que evoca as descargas elétricas entre as nuvens e o solo) dizem respeito à mesma causa simplesmente por visarem explicar o mesmo tipo de fenômeno, também há algo de errado em admitir a manutenção da referência nos casos históricos do éter luminífero e flogisto dada a notável diferença explicativa existente entre os seus respectivos termos teóricos.

A renúncia descritiva da teoria causal não só ignora o refinamento na caracterização das entidades inobserváveis como trivializa a referência dos termos teóricos, tornando comum

e inevitável o acerto referencial. Não por acaso Laudan estendeu suas críticas a esse suposto acerto referencial indicando que a teoria causal também confunde as agendas ontológica e explanatória das teorias, pois a confluência teórica na explicação de um domínio de fenômenos leva à conclusão de que seus pressupostos existenciais são idênticos, apesar de todo o restante que as teorias declaram sobre eles. Embora prescindir das descrições evite os problemas do holismo e da incomensurabilidade semântica, não é somente pressupondo o aspecto causal que mitigaremos todos os problemas relativos à referência, senão corremos o risco de tornar a referência algo trivial.

4.5. Teoria híbrida

4.5.1. O Éter luminífero e a projeção realista no passado

Exporemos agora o diagnóstico de Psillos sobre o caso do éter do século XIX. No capítulo 12 (Reference of theoretical terms) de *How Science Tracks Truth* (1999), Psillos expôs como a teoria híbrida da referência poderia ser aplicada de modo bem sucedido a tal caso histórico. Cabe notar que, na ocasião, foi utilizada a versão inicial da teoria híbrida que pressupõe a existência de tipos naturais, como se vê nas próprias citações do texto. Porém, a versão posterior, que abdica do comprometimento com os tipos naturais e só pressupõe os poderes causais, é semelhantemente aplicada neste caso, pois era suposto na versão inicial que as propriedades causais e de tipo da entidade seriam idênticas, de modo que a conclusão sobre uma delas deve aplicar-se à outra. Ademais, as conclusões extraídas do estudo desse caso dizem respeito majoritariamente ao poder causal da entidade, reiterando a importância da versão posterior. Em suma, o diagnóstico realizado por Psillos em 1999 pode ser estendido à versão posterior da teoria híbrida que concluímos ser a mais adequada para a tentativa de defesa da continuidade teórica, de modo que tal diagnóstico também é legítimo para a avaliação da eficácia dessa nova versão.

Em especial, enfatizamos como a defesa da manutenção da referência nesse caso histórico parece envolver uma projeção realista no passado que acaba depondo contra a estratégia geral de Psillos. Em resumo, por meio de documentação histórica, Kyle Stanford argumentou que Psillos ignorou a pressuposição da natureza mecânica desse éter, excluindo-as das descrições causais centrais desse caso, a fim de defender uma manutenção de referência com o termo campo eletromagnético admitido atualmente. Esta é uma projeção realista no passado que contraria a sua confiança na habilidade dos cientistas em discriminarem as

descrições relevantes para a continuidade entre teorias, minando assim a aplicação adequada da teoria híbrida na discussão dos casos históricos. Tal discordância acerca das descrições, advinda da projeção realista, acaba reforçando a nossa falha em discriminar adequadamente as descrições causais centrais dos termos teóricos.

Aqui, a vitória do realista [...] assume um caráter pírrico, não porque nós podemos aceitá-la enquanto ainda negamos que podemos confiar totalmente no que as nossas teorias atuais dizem sobre a natureza, mas sim porque ela deixa-nos com aquilo que podemos chamar de uma *falha de discriminação*: ela nos permite confiar somente em *algumas* coisas que as teorias atuais nos contam sobre o mundo natural (e.g., as *reais* “descrições causais centrais” dos termos teóricos das nossas teorias bem-sucedidas, sejam quais forem essas) enquanto nos deixa completamente incapazes de confiarmos na nossa habilidade de discernir quais partes das nossas teorias *atualmente* constituem esta classe privilegiada de afirmações teóricas. (STANFORD, 2006, p. 153).

Aplicando a estratégia *divide et impera*, Psillos concluiu que o éter do século XIX — pressuposto pelas teorias ondulatórias da luz e do eletromagnetismo — cumpria um papel central na derivação do sucesso dessas teorias. A partir desta conclusão, com o auxílio da teoria híbrida da referência, ele argumentou que os termos *éter* da óptica do século XIX e o *campo eletromagnético* compartilhariam das mesmas descrições causais centrais, sendo correferenciais e então garantindo a continuidade de referência nesse caso histórico. Psillos (1999, pp. 285-287), defendeu que os termos *éter luminífero* e *campo eletromagnético* poderiam ser classificados como correferentes por fazerem menção àquilo que é responsável pelo fenômeno de propagação da luz. É certo que o campo eletromagnético não pode ser considerado um ente mecânico tal como o éter luminífero, pois envolve propriedades novas, aparentemente não redutíveis às propriedades mecânicas clássicas.⁵³ No entanto, segundo ele, tornar-se-ia possível defender que o termo *éter luminífero* não falhou na referência porque, assim como o termo *campo eletromagnético*, ele fazia referência àquilo que cumpriria papel causal importante na propagação da luz, de modo que a descrição causal central desses termos deve ser a mesma, assegurando a manutenção e referência entre esses dois termos,

[a] identidade de papel causal tem sido baseada na identidade de propriedades fundamentais — encapsuladas nas descrições causais centrais associadas aos termos relevantes. “Éter luminífero” e “campo eletromagnético” referem à mesma entidade precisamente porque seus referentes compartilham a mesma estrutura causal

⁵³ Além do que já foi dito no final da apresentação do caso histórico do éter luminífero na subseção 4.1 sobre o caráter mecânico de tal entidade, Kyle Stanford reforça que, “atualmente, as afirmações de Psillos parecem ser mais adequadas à concepção do campo eletromagnético do início do século XX ao invés das contemporâneas: é pelo menos controverso descrever o campo eletromagnético reconhecido pela eletrodinâmica quântica como aquele no qual as ondas de luz se *propagam* (cf. Psillos, p. 296). Mas eu não repousarei nesse ponto no que se segue: ao invés, mostrarei que a vitória realista de Psillos é vazia mesmo se considerarmos esta concepção do início do século XX do campo como a nossa própria” (STANFORD, 2003, p. 559n). Posteriormente, Stanford enfatiza que, “Psillos parece conceder muito [também] ao reconhecer que as crenças dos teóricos do século XIX sobre a natureza e a constituição do éter revelaram-se equivocadas” (STANFORD, 2003, p. 560n).

explanatória — tal como esta é especificada nas suas respectivas descrições centrais. (PSILLOS, 1999, p. 286).

Segundo a análise de Psillos, no presente caso, ambos os termos referem àquilo que satisfaz as descrições causais centrais das propriedades que delimitam o papel causal da entidade que propicia a propagação da luz. Como esse filósofo indicou, essas descrições dizem respeito a dois conjuntos de propriedades.

O primeiro conjunto de propriedades eram, em geral, cinemáticas. Dado que era experimentalmente conhecido que a luz se propaga com velocidade finita, suas leis de propagação deveriam ser baseadas em um meio ao invés de serem baseadas na ação à distância. O segundo conjunto de propriedades eram, em geral, dinâmicas: o éter luminífero era o repositório da energia potencial e cinética durante a propagação da luz. O termo ‘éter’ foi empregado — na verdade, foi emprestado — para denotar a entidade que, se existisse, deveria possuir as propriedades constitutivas de tipo acima expostos. O termo ‘éter’ foi associado com uma descrição causal central das propriedades em virtude das quais o éter era suposto cumprir o seu papel causal [...] [Já] a descrição causal central associada ao termo ‘campo eletromagnético’ assumiu a descrição causal central associada ao termo ‘éter’. A postulação de Maxwell do campo eletromagnético foi, em essência, associada aos mesmos conjuntos de propriedades que foram associados à postulação do éter, no entanto, sendo atrelada também a uma classe mais ampla de eletricidade e interações magnéticas [...]. (PSILLOS, 1999, p. 286).

Desse modo, a descrição partilhada entre esses termos delimitaria as propriedades causais da entidade inobservável que propicia a propagação da luz, o que corresponderia ambos os termos a uma mesma entidade e resguardaria a continuidade nesse episódio de substituição teórica. Em Psillos (1999), esse caso serve como um exemplo da efetividade de aplicação da teoria híbrida.

Na análise do caso histórico do éter no século XIX, Psillos (1999) apresentou um conjunto de descrições causais centrais (supostamente) unânimes sobre esse termo teórico, que excluía a menção a propriedades mecânicas. Ele procedeu dessa maneira, sobretudo, por reconhecer que naquela época havia uma indeterminação sobre o modelo a ser utilizado para caracterizar o éter.

O [...] estudo de caso — que discute as teorias do éter óptico dinâmico do século XIX — visa oferecer um serviço diferente ao realismo. Ele sugere que a teoria mais geral — em termos da dinâmica Lagrangiana e a satisfação do princípio de conservação de energia — que constituiu a espinha dorsal do programa de pesquisa em torno do comportamento dinâmico do portador das ondas de luz foi retido na estrutura subsequente do eletromagnetismo. Esta teoria geral foi empregada no estudo do éter luminífero, o qual foi tomado como a estrutura dinâmica que subjazia a propagação da luz, sustentando as ondas de luz e armazenando sua energia (*vis viva*), durante o tempo em que ela deixava a fonte e até um pouco antes de ela chegar ao receptor. Dado que o portador das ondas de luz foi uma estrutura dinâmica de constituição desconhecida, a aplicação da dinâmica Lagrangiana para o estudo do seu comportamento permitiu que a comunidade científica investigasse suas propriedades mais gerais (e.g. suas leis gerais de movimento) deixando de lado os detalhes da sua constituição. A investigação da possível constituição do portador das ondas de luz foi auxiliada pela construção de modelos (e.g. o modelo sólido elástico

do éter), onde esse modelo de construção foi baseado em analogias percebidas entre o portador das ondas de luz (e.g. sua habilidade para sustentar ondas transversais) e outros sistemas físicos (e.g. sólidos elásticos). Posteriormente, esses modelos foram abandonados. (PSILLOS, 1999, p. 108-109).

No entanto, parecia haver somente uma indeterminação sobre o modelo do *éter* no século XIX, mas não sobre seu caráter mecânico. Stanford discorda de Psillos citando as palavras do próprio James Clerk Maxwell, criador da teoria eletromagnética, como lembra o título dado até hoje às equações mais importantes da teoria (as quatro “equações de Maxwell”),

Se algo é transmitido de uma partícula para outra à distância, qual é a sua condição após ter deixado uma partícula e antes de ter alcançado outra? Se esse algo é uma energia potencial de duas partículas, como na teoria de Neumann, como conceberemos esta energia existindo em um ponto do espaço, sem coincidir com estas duas partículas? De fato, em qualquer momento que a energia é transmitida de um corpo para o outro no tempo, deve haver um meio ou substância no qual a energia existe depois de deixar um corpo e antes de alcançar outro, para a energia, como Torricelli observou, “é a quintessência de uma sutil natureza que não pode ser contida em qualquer receptáculo exceto a substância mais interna das coisas materiais. Por isso todas essas teorias levam à concepção de um meio no qual a propagação ocorre, e se nós admitimos esse meio como uma hipótese, eu penso que isto ocupará um lugar proeminente em nossas investigações, e nós deveríamos nos esforçar para construir uma representação mental desta ação, e isso tem sido um objetivo constante nesse tratado (MAXWELL, 1873, p. 493).

Como podemos depreender dessa passagem, contrariamente ao que Psillos argumenta, a natureza mecânica do éter do século XIX parecia ser um compromisso teórico robusto da época. Mas suponha que após essa constatação Psillos continuasse a insistir no caráter não mecânico desse éter; o que se seguiria desta atitude é que, indesejavelmente, ele iria contrariar a sua convicção inicial de que os cientistas seriam capazes de delimitar as descrições que importam para a defesa da manutenção da referência entre teorias.

Então, mesmo se Psillos puder argumentar convincentemente que as *atuais* “descrições causais centrais” do éter não incluem afirmações sobre seu caráter mecânico ou material, ele será forçado a conceder que os julgamentos cuidadosamente considerados dos principais defensores científicos da teoria acerca de quais descrições associadas aos seus termos centrais deveriam ser satisfeitos por uma entidade para que ela desempenhasse o papel causal associado ao termo (i.e., que figura nas descrições causais centrais) se provaram pouco confiáveis. Claramente, o que isto sugere, é que não podemos repousar em nossos *próprios* julgamentos sobre quais das descrições que associamos com nossos próprios termos fazem *genuinamente* parte das nossas descrições causais centrais. E caso não possamos manter qualquer confiança de que as descrições causais centrais que *nós* associamos com os termos centrais das nossas próprias teorias são de fato corretas, então os desafios colocados ao realismo [...] pela metaindução pessimista [...] reafirmam sua força total. (STANFORD, 2006, p. 153).

A discordância nesse caso histórico, entre o que Psillos alega ser central e o que os cientistas da época acreditavam, abala a convicção de que somos capazes de determinar quais

descrições causais são centrais. E esta falha de discriminação atinge não somente aquilo que é delimitado acerca dos termos abandonados, como também dos termos atuais: não poderíamos confiar também naquilo que os cientistas estabelecem sobre os termos teóricos atualmente, uma vez que eles se equivocaram no passado. Nesse cenário, não haveria nenhuma base segura de correspondência entre as descrições causais dos termos teóricos, inviabilizando qualquer tentativa de defesa da manutenção de referência entre as teorias. Esta projeção realista do presente sobre o passado indicaria a nossa falha em discriminar quais descrições causais são centrais, configurando um novo tipo de MIP que coloca em xeque a nossa habilidade de delimitar os itens cruciais para a defesa da continuidade entre as teorias, seja em casos passados, presentes ou futuros.

Como consequência, uma aplicação legítima da teoria híbrida da referência não deveria repousar em um critério de delimitação tão frágil. As descrições relevantes para a manutenção da referência seriam mais explícitas e o realista não correria o risco de fazer uma leitura seletiva da história da ciência que fosse arbitrária para defender a continuidade entre as teorias.

Todavia, há uma ressalva a ser feita no caso do éter luminífero. Considerando os registros históricos e a investigação realizada por Psillos (1999) sobre o caso do éter luminífero (ou óptico) do século XIX, Stanford tem razão ao acusá-lo de projetar descrições teóricas presentes sobre o passado. Porém, ressaltamos que, o sucesso da teoria eletromagnética do século XIX não dependeu da suposição da existência do éter; de modo que o éter referido por essa teoria, i.e., o éter eletromagnético, não se constitui como um caso histórico que ameaça o realismo científico segundo a estratégia *divide et impera*. As equações de Maxwell ofereceram um tratamento quantitativo dos fenômenos elétricos, magnéticos e ópticos, unificando-os como um mesmo conjunto de fenômenos e gerou previsões inéditas importantes, como a existência de ondas de rádio (LYONS, 2002, p. 71). Importante notar que essas equações não fazem nenhuma referência ao éter, quando aplicadas na predição de fenômenos, como aliás, fica claro do fato de que nas apresentações atuais das equações nos livros-textos de eletromagnetismo o éter nunca ser mencionado, a não ser, eventualmente, em notas de rodapé, como curiosidade histórica. Em sua análise sobre a manutenção de descrições causais, Psillos volta-se majoritariamente para o campo da óptica, enfatizando o papel do éter na explicação da luz (aí intitulado como éter luminífero); todavia, na teoria eletromagnética, esse éter foi intitulado eletromagnético após Maxwell incorporar a óptica ao eletromagnetismo. Como corroboração, Valter Bezerra afirma que, “a teoria de Maxwell [...] é também uma teoria do éter eletromagnético (que ele acabaria concluindo ser o mesmo que o

éter luminífero)” (BEZERRA, 2006, p. 191). Em termos explicativos, o éter eletromagnético era importante, mas em termos preditivos (realmente relevantes para a noção de sucesso teórico) ele era dispensável — as quatro equações de Maxwell podiam ser derivadas do modelo mecânico do éter, mas não dependiam crucialmente dele, tanto que Heinrich Hertz (1857-1894) conseguiu derivar as equações de Maxwell sem pressupor explicitamente a existência do éter eletromagnético (CORDERO, 2011, p. 1127). Ao apresentar a sua versão da MIP, Larry Laudan (1981, p. 33) listou o éter luminífero (em seu artigo, éter óptico) logo após o éter eletromagnético como um caso histórico que oferece problemas ao realismo. Como explanado, o éter eletromagnético não se constitui então um contraexemplo histórico ao que o realista defende no AM; contudo, conforme o próprio Psillos defendeu, o éter luminífero cumpria um papel central para a geração do sucesso da sua teoria, mas era inegavelmente mecânico, o que impede a manutenção de referência entre os termos *éter luminífero* e *campo eletromagnético* — justificando a crítica de Stanford à Psillos. A não ser que se admita que o *éter luminífero* é correferente com o *éter eletromagnético* — o que parece ser o caso, dada a preservação da caracterização mecânica do éter entre esses dois termos —, aí sim ele não oferecerá maiores problemas, pois o *éter eletromagnético* não precisa compartilhar da mesma referência que o *campo eletromagnético* para salvaguardar o realismo, como foi acima exposto. E assegurar a manutenção da referência a partir da análise mais pormenorizada desse caso afasta os problemas do holismo e da incomensurabilidade semântica, resguardando a continuidade teórica sem maiores obstáculos.

4.5.2. O flogisto e a negligência realista

Ao aplicarmos a teoria híbrida da referência ao caso histórico do flogisto constatamos que a descrição causal central que remete ao flogisto e delimitava o seu papel causal na produção dos fenômenos de combustão não é preservada na caracterização do oxigênio: uma substância emitida na atmosfera durante o processo de combustão. Por não preservar tal descrição, o termo *oxigênio* não refere à mesma entidade que o termo *flogisto*, de modo que eles diferem quanto à referência, inviabilizando assim a defesa da continuidade teórica nesse caso.

Psillos admite que a sua proposta falha na defesa da manutenção de referência entre os termos flogisto e oxigênio, mas o faz sob a égide de que a sua proposta explicaria adequadamente a falha de referência que já seria esperada neste caso por conta da divergência das descrições causais desses termos (PSILLOS, 1999, p. 288). Porém, a atitude adotada por

Psillos contradiz sua estratégia *divide et impera*, segunda a qual os itens que foram essenciais para o sucesso de alguma teoria científica deveriam permanecer em episódios de substituição teórica. Se a pressuposição da existência do flogisto foi essencial para a derivação de algumas previsões inéditas da sua teoria, então deveria haver uma manutenção de referência nesse caso. Quanto ao ar desflogisticado, a outra entidade inobservável pressuposta pela teoria do flogisto e que também contribuiu consideravelmente para seu aspecto bem sucedido através da unificação de fenômenos, Psillos sequer estende a aplicação da teoria híbrida ao seu termo teórico, contradizendo novamente a sua estratégia *divide et impera*.

Dada a negligência com o termo teórico *ar desflogisticado*, avaliaremos se é possível defender alguma continuidade teórica relativamente a esse termo a fim de avaliar o desempenho final da teoria híbrida no presente caso histórico. Com o abandono da existência do flogisto, como dar sentido à descrição que delimitava o papel causal do ar desflogisticado quanto aos fenômenos da facilitação da respiração e da combustão: substância obtida quando a substância emitida na combustão (flogisto) é removida do ar? Considerando o aspecto ontológico, percebemos que as descrições causais associadas a esse outro termo dependem do papel causal atribuído ao flogisto; como consequência, o abandono da sua existência afeta a satisfação do termo derivado *ar desflogisticado*, inviabilizando novamente a defesa da manutenção da referência com o termo *oxigênio* e reafirmando a falha de referência da teoria do flogisto.

A teoria híbrida da referência é duplamente falha ao ser aplicada ao caso do flogisto, e, sobretudo, quando pensamos na proposta de Psillos: ela falha quanto à suposição do flogisto e do ar desflogisticado — i.e., falhou na pressuposição de entidades que contribuíram para a realização de previsões inéditas e/ou a unificação de fenômenos, aspectos esses que constituem o sucesso de uma teoria e que, portanto, configuram essas entidades como essenciais para a geração do sucesso da teoria do flogisto, de modo que deveriam ter sido contempladas nessa proposta realista. Mesmo com a admissão de Psillos acerca da falha de referência do termo flogisto, a sua desconsideração em analisar seu termo derivado *ar desflogisticado* somada à ineficácia da sua proposta quanto a esse outro termo — sobre o qual inegavelmente deveria haver algum tipo de continuidade, pois além da centralidade teórica dessa entidade, ela tem uma grande semelhança explicativa com o termo vigente atualmente, o *oxigênio* — nos indica que Psillos foi negligente na avaliação desse caso histórico.

Como não há a manutenção de descrições causais centrais e, consequentemente, da referência, dos termos *flogisto* e *ar desflogisticado*, os problemas do holismo e da incomensurabilidade semântica não são afastados. A alteração de significado no caso do

flogisto foi tamanha que o núcleo central das descrições associadas a cada um dos termos foi incapaz de sustentar qualquer tipo de continuidade. Novamente, o holismo foi muito bem representado pelo pressuposto existencial do termo *ar desflogisticado*: seu significado estava tão estritamente associado ao do termo principal da teoria do flogisto que bastou tal termo perder sua referência para o mesmo infortúnio se suceder com o *ar desflogisticado*. Como resultado da falha de manutenção das descrições centrais para a referência, não há aspectos semânticos relevantes compartilhados entre as teorias do flogisto e a do oxigênio, o que dificulta a comparação entre tais teorias, como é ressaltado pelo problema da incomensurabilidade semântica.

Seja considerando ou não a proposta realista seletiva de Stathis Psillos — que depende tanto da aplicação da teoria híbrida da referência bem como da estratégia *divide et impera* —, concluímos que a teoria híbrida da referência não consegue assegurar a manutenção de referência no caso histórico do flogisto, seja com relação ao termo *flogisto* ou *ar desflogisticado*. Por não resguardar a continuidade teórica neste caso, depreendemos que a aplicação da teoria híbrida aos casos históricos não é infalível, embora somente uma análise histórica mais abrangente poderia embasar um julgamento apropriado dessa teoria da referência.

Capítulo 5: Considerações finais

No debate entre realistas e antirrealistas científicos sobre os entes inobserváveis postulados pela ciência, este trabalho se propôs a defender o lado realista da discussão, buscando reabilitar o AM diante dos problemas levantados pelos casos históricos de substituição teórica com o auxílio das teorias da referência.

As melhores teorias científicas que pressupõem a existência de inobserváveis são notadamente bem-sucedidas. É este o fato que comumente motiva uma postura realista no debate, que se compromete epistemicamente com tais entidades; no nosso caso, também não foi diferente: mesmo ciente de todas as controvérsias e questionamentos acerca dos inobserváveis, o sucesso das teorias científicas continuava a nos intrigar, fortalecendo a ideia de que, de algum modo, esse sucesso remontava à existência de tais entidades. Como vimos a partir do capítulo 2, no AM encontramos a síntese desse tipo de argumentação realista, que nos traz uma defesa da verdade aproximada das teorias científicas que pressupõem entidades inobserváveis em virtude do seu sucesso. O AM é uma instância da IME, no qual temos a defesa do comprometimento realista com os inobserváveis por ser essa a melhor explicação para o sucesso das teorias científicas.

O antirrealismo se opõe a esse tipo de argumentação realista, a começar pela adesão a uma forma estrita de empirismo, que limita o comprometimento epistêmico ao âmbito observável do mundo. Nessa linha, o sucesso das teorias não é encarado com surpresa e nem demandaria maiores explicações, pois seria o resultado natural de uma filtragem da ciência que seleciona apenas teorias empiricamente adequadas. Por sua vez, argumentos como a SUB e a MIP desafiam o realismo, mostrando que a própria etapa de escolha entre teorias coloca problemas para as noções de melhor explicação e verdade aproximada, além de apresentarem considerações históricas acerca de episódios de substituição teórica que desafiam a argumentação realista de que teorias bem-sucedidas acertam na referência ao inobservável.

Pela SUB, podemos ter duas teorias empiricamente equivalentes que sejam inconsistentes entre si no nível dos inobserváveis que postulam, de modo que apesar das suas explicações diferentes para o mesmo domínio de fenômenos, a experiência as confirma igualmente, impedindo-nos de determinar empiricamente qual delas melhor explica os fenômenos e é mais aproximadamente verdadeira. Os realistas alegam que certas virtudes teóricas (como consistência e coerência, simplicidade, escopo, unidade, etc) podem resolver esse problema, desempatando a disputa teórica. No entanto, antirrealistas não atribuem

estatuto epistêmico a tais virtudes, classificando-as como meros fatores pragmáticos que ajudam a facilitar o trabalho científico de escolha entre teorias, sem que indiquem efetivamente qual teoria está mais próxima da verdade. Em contrapartida ao AM, a SUB enfatiza que o sucesso de uma teoria nem sempre é um fator suficiente para determinar sua verdade aproximada. O fator explicativo que é tão importante para estrutura do AM também acaba sendo questionado, pois não há uma clareza sobre quais aspectos de uma explicação a tornam realmente boa para ser classificada como verdadeira. Consequentemente, a IME que auxilia na escolha entre hipóteses científicas — cumprindo assim um papel importante na suposição da existência de inobserváveis — também é seriamente criticada, já que o seu funcionamento depende essencialmente de considerações explicativas. Por apresentar mais um caráter de possibilidade do que algo concreto, não exploramos os problemas apresentados pela SUB, optando então por examinar as considerações históricas levantadas pela MIP. Todavia, reconhecemos que, sendo o fator explicativo relevante para o AM, é necessário que o realismo científico se preocupe em justificar como certos aspectos de uma explicação podem ser indicativos da sua verdade, sobretudo considerando críticas antirrealistas já formuladas.

Segundo a MIP, dada a ocorrência de casos históricos que compõem episódios de substituição teórica na ciência, nem sempre as teorias bem-sucedidas acertam naquilo que é pressuposto sobre os inobserváveis, pois apesar do seu sucesso elas podem vir a ser consideradas falsas. Em episódios de substituição teórica, uma teoria científica é adotada no lugar de outra por ser considerada uma melhor alternativa para nos informar sobre um conjunto de fenômenos, passando, assim, a ser considerada verdadeira ou mais próxima da verdade do que a anterior, que, por sua vez, é classificada como falsa. Os casos históricos da MIP são as teorias que foram abandonadas em episódios de substituição teórica, sendo, portanto, aquelas consideradas falsas e falhas na referência aos seus respectivos inobserváveis, apesar de terem usufruído de considerável sucesso em sua época. Dado que o AM supõe que uma teoria bem-sucedida é aproximadamente verdadeira e acerta na referência ao inobservável, a existência de casos históricos de falha teórica mostra que essa suposição pode ser bastante equivocada, de modo que seria injustificado aludir ao sucesso de uma teoria para defender o comprometimento com as entidades inobserváveis. Para continuar a sustentar o AM, que embasa a ideia de que o aspecto bem-sucedido das teorias científicas remonta à existência dos inobserváveis, é preciso tentar lidar adequadamente com os episódios de substituição teórica dos quais os casos históricos fazem parte; e foi o que tencionamos realizar com o auxílio das teorias da referência.

Os episódios de substituição teórica parecem envolver um alto grau de descontinuidade teórica, dada a grande divergência entre as teorias que os compõem. No que diz respeito aos casos históricos da MIP, as respectivas entidades inobserváveis pressupostas pelas teorias abandonadas foram consideradas inexistentes, de modo que seus respectivos termos teóricos passaram a ser não referentes. Diante desse problema, recorremos ao auxílio das teorias da referência para averiguar se seria possível defender que há continuidade entre as teorias científicas desses episódios de substituição teórica devido à manutenção de referência às entidades inobserváveis pressupostos pelos termos teóricos em cada caso. Assim os casos históricos não seriam constituídos por teorias bem-sucedidas que são falsas e falhas quanto à referência, mas sim de teorias que acertaram ao menos parcialmente na referência ao inobservável e são aproximadamente verdadeiras, sendo somente devido a avanços de uma teoria posterior mais aproximadamente verdadeira que elas vieram a ser substituídas.

A continuidade teórica almejada com o recurso às teorias da referência não distingue o caráter das entidades inobserváveis a serem mantidas, i.e., se são entidades ou estruturas, pois o que importa é a manutenção de referência a tais entidades. Devido a esse aspecto, não exploramos o realismo de entidades ou o estruturalismo neste trabalho; preferimos uma continuidade mais geral que uma teoria da referência pode oferecer e a explicação do sucesso da ciência que pode ser proporcionada por tal continuidade teórica. Consequentemente, o realismo explanacionista — que sustenta um realismo quanto aos pressupostos das teorias científicas que são essenciais para o seu sucesso e o explicam — foi o nosso tipo de postura realista escolhida.

No fim do segundo capítulo, vimos que os casos históricos que devem ser considerados genuinamente bem-sucedidos são aqueles constituídos por teorias maduras que não são *ad hoc* e que realizaram previsões inéditas e unificações de fenômenos. Ainda assim, restam casos que satisfazem essa exigência mais robusta de sucesso, como o éter luminífero e o flogisto que foram estudados neste trabalho. Adicionalmente, a proposta do realismo explanacionista, de só sustentar uma atitude realista quanto aos pressupostos das teorias científicas que cumpriram um papel importante na geração do seu sucesso, foi de grande auxílio para redirecionar nossa compreensão dos casos históricos.

No terceiro capítulo apresentamos com razoável detalhe as teorias da referência descritivista, causal e híbrida. Primeiro o fizemos de acordo com seus propósitos originais (que eram voltados para os nomes próprios) para depois mostrar como elas poderiam ser aplicadas aos termos teóricos (referentes aos inobserváveis) e indicar algumas consequências prévias dessa aplicação. Segundo a teoria descritivista, um termo teórico refere a um

inobservável através das descrições fornecidas pela sua teoria. A teoria causal estabelece que o suposto papel causal de uma entidade inobservável é o que define a referência de um termo teórico. Por sua vez, a teoria descritivista foi proposta como um meio termo entre as duas teorias da referência anteriores; combinando elementos descritivos e causais, instituiu que descrições causais centrais, que caracterizam o suposto papel causal de uma entidade inobservável, são os fatores que intermediam a referência de um termo teórico. O holismo e a incomensurabilidade semântica são dois problemas previamente anunciados que podem afetar o desempenho das teorias da referência. Sendo aspectos concernentes ao significado, eles podem acabar prejudicando as teorias da referência que pressupõem a intermediação das descrições para a referência. Por sua vez, a teoria causal pode garantir a manutenção de referência ao custo de trivializar tal conquista, pois dado o recurso ao papel causal das entidades o acerto referencial sempre estaria assegurado. Já a teoria híbrida também pode enfrentar o problema da projeção realista no passado, por envolver um estudo retroativo na tentativa de defesa da manutenção da referência.

No quarto capítulo apresentamos brevemente os casos históricos do éter luminífero e do flogisto, com o objetivo de analisar o desempenho das teorias da referência quando aplicada a tais casos. De modo geral, verificamos que, em maior ou menor grau, todas as teorias da referência apresentam problemas ao serem aplicadas aos casos históricos.

A teoria descritivista requer uma coleta mais extensiva de descrições para a referência, por essa razão ela acaba sendo enfraquecida pelo holismo e incomensurabilidade semântica, inviabilizando a manutenção da referência entre teorias. Não há continuidade no caso do éter luminífero pelo abandono da descrição que delimitava seu caráter mecânico, enquanto no caso do flogisto, o abandono do suposto referente do termo *flogisto* afetou a referência do *ar desflogisticado*. Em ambos os casos, o holismo se manifestou em decorrência da mudança teórica que gerou alterações descritivas, impedindo a correferência. Além disso, o holismo também contribuiu para o surgimento da incomensurabilidade semântica, uma vez que a descontinuidade teórica resultante dificulta a comparação entre teorias. Com a teoria descritivista dificilmente será possível defender a manutenção de referência entre teorias devido à sua exigência excessiva de descrições para o acerto referencial.

Já a teoria causal, ao prescindir da coleta de descrições evita as fraquezas da teoria descritivista, usufruindo de uma alta taxa de acerto referencial. Contudo, a teoria causal é significativamente enfraquecida pela trivialização da referência, uma vez que ela é imune a falhas mesmo quando a mudança teórica é muito discrepante para ser ignorada. Este resultado indesejado ocorre pela desconsideração das descrições na determinação da referência. Além

disso, tal proposta confunde as agendas ontológica e explanatória das teorias, já que a concordância sobre os fenômenos que as teorias explicam leva à conclusão de que suas pressuposições ontológicas são idênticas, apesar de qualquer diferença em como elas são descritas. Nessa linha, seria preciso ignorar a caracterização mecânica do éter luminífero ou a produção de fenômenos diferentes atribuídos ao flogisto e ar desflogisticado para admitir uma correferência em seus respectivos casos. Devido ao seu foco exclusivo no aspecto causal das entidades, a teoria causal da referência também não é uma escolha adequada para os nossos propósitos.

A seu turno, a teoria híbrida, por ser uma proposta de equilíbrio entre as duas propostas anteriores, parece ser mais promissora para a defesa da manutenção da referência entre as teorias. E nos casos históricos em que a teoria híbrida assegura a manutenção de referência, o holismo e a incomensurabilidade semântica não afetam seu desempenho, pois o significado dos termos não implica diretamente em uma mudança de referência entre as teorias. Concluimos que a versão mais recente de aplicação desta teoria da referência — a que pressupõe apenas aspectos causais para a referência — é a mais adequada a ser adotada, dado ter menos pressupostos ontológicos e não supor uma identificação entre as propriedades causais e propriedades de tipo de uma entidade. Contudo, a possibilidade de projeção realista no passado é um risco inerente à aplicação da teoria híbrida aos casos históricos, pois a delimitação de descrições causais centrais para a referência pode ser feita arbitrariamente sem considerar o que de fato era central no passado na determinação da referência de um termo teórico; o caso do éter luminífero tal como analisado por Stathis Psillos (um dos principais colaboradores para a aplicação da teoria híbrida neste contexto) é afetado por essa nefasta possibilidade, mas esse problema pode ser relevado se, como propusemos, admitirmos uma correferência com a teoria eletromagnética. Se no fim a teoria híbrida até consegue ter um bom desempenho no caso do éter, o mesmo desempenho já não se estende ao caso do flogisto. Tanto seus termos *flogisto* e *ar desflogisticado* (que foram essenciais para o aspecto bem-sucedido da teoria desse caso) tiveram suas respectivas descrições causais centrais abandonadas, falhando na referência ao inobservável. Como resultado, o holismo e a incomensurabilidade semântica são problemas que acabam remanesendo neste caso. Este caso histórico evidencia uma falha de aplicação da teoria híbrida da referência, prejudicando o seu desempenho na manutenção da referência entre as teorias.

De modo geral, a teoria híbrida requer mais atenção quanto à coleta de descrições centrais, e é enfraquecida pela i) possibilidade de envolver a projeção realista no passado, como ocorreu no caso do éter, ou por ii) pressuposições existenciais (ontológicas), que

ficaram evidentes na análise do flogisto. A teoria híbrida aparentemente falha na defesa da manutenção da referência no caso do éter luminífero, mas vimos que tal falha talvez possa ser contornada pela admissão de que há neste caso uma correferência com o termo *éter eletromagnético* e a aplicação da estratégia *divide et impera* desenvolvida por Stathis Psillos. Contudo, a mesma ressalva não pode ser feita quanto ao caso do flogisto, de modo que a ineficácia da teoria híbrida nesse caso parece mais séria, numa avaliação final dessa teoria da referência.

Dado o prognóstico final das teorias da referência, a teoria híbrida é a que tem o melhor desempenho, pois ela não é tão sensível à mudança descritiva entre teorias e nem resguarda a referência aos inobserváveis ignorando suas características mais relevantes. No entanto, dada a sua falha com a teoria do flogisto, fica o alerta de que ela pode não ser efetiva em todos os casos. E mesmo que adotemos um falibilismo quanto a este ponto a fim de relevar falhas como essa, ainda resta a incômoda consideração do quão falha uma estratégia pode ser para que ainda a consideremos eficiente dentre diversos casos possíveis de aplicação. Por essa razão, é imprescindível, evidentemente, o estudo de mais casos históricos a fim de avaliar o desempenho efetivo da teoria híbrida na defesa da manutenção de referência. Os casos históricos dessa dissertação representam uma fração mínima do que precisa ser analisado para embasar um julgamento apropriado sobre a efetividade dessa proposta. Nossa contribuição consistiu na delimitação da teoria da referência mais promissora para a defesa da manutenção de referência em episódios de substituição teórica escolhidos, na esperança que esse estudo possa ser expandido em trabalhos futuros. Analisamos os aspectos conceituais de cada teoria e algumas consequências prévias da aplicação de cada uma delas, além de realizar uma breve análise de dois casos históricos relevantes que nos permitiram confirmar as consequências previamente anunciadas e extrair nossas considerações finais. Concluimos que a teoria híbrida é a mais adequada para defender a manutenção da referência e garantir a continuidade teórica na ciência, sendo a melhor alternativa para reabilitar o AM perante a ameaça dos casos históricos da MIP. No entanto, somente uma análise mais extensa do desempenho da aplicação da teoria híbrida, tendo em vista diversos casos históricos, possibilitaria um julgamento abrangente sobre o real desempenho dessa teoria da referência. Por ser este um trabalho de cunho realista, gostaríamos de defender a continuidade teórica em todos os casos, mas isso não pode ser feito de modo geral, pois a aplicação da teoria híbrida não é infalível (como evidencia o caso do flogisto) e, sobretudo, devido à necessidade de uma análise histórica mais abrangente de diversos outros casos semelhantes.

Bibliografia

BEZERRA, Valter Alnis. “Maxwell, a teoria do campo e a desmecanização da física”. *Scientiae Studia*, v. 4, p. 177-220, 2006.

BIRD, Alexander. *Thomas Kuhn*. London: Routledge, 2000.

BORGE, Bruno. “¿Soluciona el descriptivismo causal el problema de la referencia de los términos teóricos?”. *ideas y valores*, v. 66, n. 163, p. 125-151, 2017.

BOYD, Richard N. “3. The Current Status of Scientific Realism”. In: LEPLIN, J. (ed.) *Scientific realism*. University of California Press, 1984. p. 41-82.

BOYD, Richard. “What realism implies and what it does not”. *dialectica*, v. 43, n. 1-2, p. 5-29, 1989.

BOYD, Richard. “Realism, anti-foundationalism and the enthusiasm for natural kinds”. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, v. 61, n. 1/2, p. 127-148, 1991.

BOYD, Richard N. *Metaphor and theory change*. 1993.

BRANQUINHO, João; MURCHO, Desidério; GOMES, Nelson Gonçalves. *Enciclopédia de termos lógico-filosóficos*. 2006.

BRIDGMAN, Percy Williams. *The logic of modern physics*. Macmillan, 1927.

BRITO, Armando A. “Flogisto, Calórico & Éter” . *Ciência & Tecnologia dos Materiais*, v. 20, n. 3-4, p. 51-63, 2008.

CAMPOS, Daniel G. "On the distinction between Peirce's abduction and Lipton's inference to the best explanation". *Synthese*, v. 180, p. 419-442, 2011.

CANTOR, Geoffrey N.; HODGE, Michael JS. *Conceptions of Ether: Studies in the History of Ether Theories 1740-1900*. Cambridge: Cambridge UP, 1981.

CAJORI, Florian. "An historical and explanatory apêndix". In: Newton, I. *Mathematical principles of natural philosophy*. Tradução A. Motte; revisão F. Cajori. Berkeley/Los Angeles: University of California Press, 1934. p. 627-80.

CARNAP, Rudolf (1928). *The Logical Structure of the World*. Berkeley: University of California Press, 1967.

CARNAP, Rudolf. "Testability and meaning". *Philosophy of science*, v. 3, n. 4, p. 419-471, 1936.

CARNAP, Rudolf. Replies and Systematic Expositions. In: SCHILPP, Paul Arthur. (ed.) *The Philosophy of Rudolf Carnap*, La Salle: Open Court, 1963.

CARNAP, Rudolf. *Philosophical foundations of physics*. (New York: Basic Books), 1966.

CARRIER, Martin. "What is wrong with the miracle argument?". *Studies in History and Philosophy of Science Part A*, v. 22, n. 1, p. 23-36, 1991.

CARRIER, Martin. "What is right with the miracle argument: Establishing a taxonomy of natural kinds". *Studies in History and Philosophy of Science Part A*, v. 24, n. 3, p. 391-409, 1993.

CARTWRIGHT, Nancy. *How the Laws of Physics Lie*. OUP Oxford, 1983.

CHAKRAVARTTY, Anjan. "Scientific Realism". *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2017/entries/scientific-realism/>>. Acesso em: 27 de mai. de 2023.

CHIBENI, Silvio Seno. "Descartes, Locke, Berkeley, Hume e o Realismo Científico". *Primeira Versão* v. 25, 1990.

CHIBENI, Silvio Seno. “A inferência abdutiva e o realismo científico”. *Cadernos de História e Filosofia da Ciência*, v. 6, n. 1, p. 45-73, 1996.

CHIBENI, Silvio Seno. *Aspectos da descrição física da realidade*. Campinas: UNICAMP, Centro de Lógica, Epistemologia e História da Ciência, 1997.

CHIBENI, Silvio Seno. “Afirmando o conseqüente: uma defesa do realismo científico (?)”. *Scientiae Studia*, v. 4, p. 221-249, 2006.

CHIBENI, Silvio Seno. “Berkeley e o realismo científico”. *XIII Encontro da Associação Nacional de Pós Graduação em Filosofia*, v. 10, 2008.

CHIBENI, Silvio Seno. “Berkeley and the role of hypothesis in natural philosophy”. *Scientiae Studia*, v. 8, p. 389-419, 2010.

CHIBENI, Silvio Seno. "Hume e as bases científicas da tese de que não há acaso no mundo." *Principia: an international journal of epistemology*, v. 16, n. 2, p. 229-254, 2012.

CHURCHLAND, P. M. “The ontological status of observables: In praise of superempirical virtues”. In: P. M. Churchland & C. A. Hooker (eds.) *Images of Science*. University of Chicago Press, Chicago, 1985. p. 35-47.

CORDERO, Alberto. “Scientific realism and the divide et impera strategy: The ether saga revisited”. *Philosophy of Science*, v. 78, n. 5, p. 1120-1130, 2011.

CRAIG, William. “Replacements of auxiliary assumptions”. *The Philosophical Review*, v. 65, p. 38-55, 1956.

DE SOUZA, Pedro Bravo. “O argumento do milagre comete a falácia da taxa-base? Apresentação, estado da arte e questões de formalização”. *Intuitio*, v. 11, n. 1, p. 46-64, 2018.

DEVITT, Michael; STERELNY, Kim. *Language and reality: An introduction to the philosophy of language*. Mit Press, 1999.

DICKEN, Paul. *A Critical Introduction to Scientific Realism*. Bloomsbury Publishing, 2016.

DONNELLAN, Keith S. "Proper names and identifying descriptions". *Synthese*, v. 21, p. 335-358, 1970.

DUHEM, Pierre Maurice Marie (1906). *The Aim and Structure of Physical Theory*. Princeton University Press, 1954.

ENĆ, Berent. "Reference of theoretical terms". *Nous*, v. 10, n. 3, p. 261-282, 1976.

EVANS, Gareth. "The Causal Theory of Names". *Aristotelian Society Supplementary*, v. 47, n. 1, p. 187–208, 1973.

FEYERABEND, Paul K. "Explanation, Reduction, and Empiricism". In: H. Feigl, & G. Maxwell (Eds.) *Scientific Explanation, Space, and Time, Minnesota Studies in the Philosophy of Science* (Vol. 3, p. 28-97). Minneapolis, MN: University of Minnesota Press, 1962.

FABIAN, Ulisses Cauê Bôa Ventura. *Hume e a existência de conexões necessárias na natureza*. 2022. 137 p. Dissertação (Mestrado) – Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 2022. Recurso Eletrônico. Disponível em: <<https://hdl.handle.net/20.500.12733/6080>>. Acesso em: 5 de jan. de 2024.

FINE, Arthur. "The Natural Ontological Attitude". In: LEPLIN, Jarrett. (ed.) *Scientific realism*. (Berkeley: University of California Press), 1984. p. 83-107.

FREGE, Gottlob. "Sense and reference" (1892). *The philosophical review*, v. 57, n. 3, p. 209-230, 1948.

HACKING, Ian. *Representing and intervening: Introductory topics in the philosophy of natural science*. Cambridge university press, 1983.

HARDIN, Clyde L.; ROSENBERG, Alexander. "In defense of convergent realism". *Philosophy of Science*, v. 49, n. 4, p. 604-615, 1982.

HARKER, David. "Two arguments for scientific realism unified." *Studies in History and Philosophy of Science Part A*, v. 41, n. 2, p. 192-202, 2010.

HARMAN, Gilbert H. "The inference to the best explanation". *The philosophical review*, v. 74, n. 1, p. 88-95, 1965.

HARMAN, Gilbert H. "Enumerative induction as inference to the best explanation". *The Journal of Philosophy*, v. 65, n. 18, p. 529-533, 1968.

HEMPEL, Carl G. "Problems and changes in the empiricist criterion of meaning". *Revue internationale de philosophie*, p. 41-63, 1950.

HEMPEL, Carl G. *Aspects of scientific explanation*. New York: Free Press, 1965.

HOWSON, Colin. *Hume's problem: Induction and the justification of belief*. Clarendon Press, 2000.

KITCHER, Philip. "Theories, theorists and theoretical change". *The Philosophical Review*, v. 87, n. 4, p. 519-547, 1978.

KITCHER, Philip. *The advancement of science: Science without legend, objectivity without illusions*. Oxford University Press on Demand, 1993.

KORNBLITH, H. *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. MIT Press Cambridge, 1993.

KRIPKE, Saul. *Naming and Necessity*. (Oxford. UK: Blackwell), 1980.

KROON, Frederick W. "Causal descriptivism". *Australasian journal of Philosophy*, v. 65, n. 1, p. 1-17, 1987.

KUHN, Thomas S. *The structure of scientific revolutions*. University of Chicago Press: Chicago, 1970.

LABINGER, J. A.; WEININGER, S. J. "Controversy in Chemistry: How Do You Prove a Negative? – The Cases of Phlogiston and Cold Fusion". *Angewandte Chemie International Edition*, v. 44, n. 13, p. 1916-1922, 2005.

LADYMAN, James. "What is structural realism?". *Studies in History and Philosophy of Science Part A*, v. 29, n. 3, p. 409-424, 1998.

LADYMAN, James. "Structural realism versus standard scientific realism: The case of phlogiston and dephlogisticated air". *Synthese*, v. 180, n. 2, p. 87-101, 2011.

LAUDAN, Larry. "A confutation of convergent realism". *Philosophy of science*, v. 48, n. 1, p. 19-49, 1981.

LAUDAN, Larry. "Realism without the real". *Philosophy of Science*, v. 51, n. 1, p. 156-162, 1984.

LAVOISIER, A. L. *Traité élémentaire de chimie*. 1789.

LEPLIN, Jarrett. *A novel defense of scientific realism*. Oxford University Press, 1997.

LEWIS, David. "Putnam's paradox". *Australasian journal of philosophy*, v. 62, n. 3, p. 221-236, 1984.

LEWIS, Peter J. "Why the pessimistic induction is a fallacy". *Synthese*, v. 129, p. 371-380, 2001.

LIPTON, Peter. *Inference to the best explanation*. Routledge, 2004.

LYCAN, William G. *Philosophy of language: A contemporary introduction*. Routledge, 2008.

LYONS, Timothy D. "Scientific realism and the pessimistic meta-modus tollens". In: *Recent themes in the philosophy of science*. Springer, Dordrecht, p. 63-90, 2002.

MACH, Ernst. *The Science of Mechanics: A Critical and Historical Account of Its Development*. Open Court, 1893.

MAGNUS, Paul D.; CALLENDER, Craig. "Realist ennui and the base rate fallacy". *Philosophy of Science*, v. 71, n. 3, p. 320-338, 2004.

MALAVOLTA E SILVA, Bruno. "The Anti-Realist Explanation for Science's Success: Semantics, Method and Attitude." *KRITERION–Journal of Philosophy*, v. 34, n. 4, p. 15-44, 2020.

MARTINICH, Aloysius P. *The Philosophy of Language*. New York: Oxford University Press, 1996.

MARTINS, Roberto de Andrade. "Do éter ao vácuo e de volta ao éter". *Scientific American, Os Grandes Erros da Ciência, Especial História*, v. 6, p. 92-98, 2006.

MARTINS, Roberto Andrade. "O éter e a óptica dos corpos em movimento: a teoria de Fresnel e as tentativas de detecção do movimento da Terra, antes dos experimentos de Michelson e Morley (1818-1880)". *Caderno Brasileiro de Ensino de Física*, v. 29, n. 1, p. 52-80, 2012.

MAXWELL, Grover; FEIGL, Herbert. *The ontological status of theoretical entities*. 1962.

MAXWELL, James Clerk. *A treatise on electricity and magnetism*. Oxford: Clarendon Press, 1873.

MCLEISH, Christina. "Scientific realism bit by bit: Part I. Kitcher on reference". *Studies in History and Philosophy of Science Part A*, v. 36, n. 4, p. 668-686, 2005.

MILL, John Stuart (1843). *A System of Logic, Ratiocinative and Inductive*. In: *Collected Works*, v. 7, 1973.

MUSGRAVE, Alan. "The ultimate argument for scientific realism". *Relativism and realism in science*, p. 229-252, 1988.

NAVARRO, Jaume (Ed.). *Ether and modernity: The recalcitrance of an epistemic object in the early twentieth century*. Oxford University Press, 2018.

NAGEL, E. *The Structure of Science*. Londres: Routledge & Kegan Paul, 1961.

NOLA, Robert. "Fixing the reference of theoretical terms". *Philosophy of Science*, v. 47, n. 4, p. 505-531, 1980.

OBERHEIM, Eric and Paul Hoyningen-Huene, "The Incommensurability of Scientific Theories", *The Stanford Encyclopedia of Philosophy* (Fall 2018 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/fall2018/entries/incommensurability/>>. Acesso em: 08 de set. de 2023.

OLEGARIO DA SILVA, Gilson. *Os empiristas vão à missa: compromissos ontológicos e frameworks linguísticos*. 2020. 239 p. Tese (Doutorado) - Instituto de Filosofia e Ciências Humanas, Universidade Estadual de Campinas, Campinas, 2020. Recurso Eletrônico. Disponível em: < <https://hdl.handle.net/20.500.12733/1638803>>. Acesso em: 5 de jan. de 2024.

OLIVEIRA, Tiago Luís Teixeira. "Algumas razões para levar a sério a metaindução pessimista". *Principia: an international journal of epistemology*, v. 18, n. 2, p. 269-290, 2014.

PAPINEAU, David. "Realism, Ramsey sentences and the pessimistic metainduction". *Studies In History and Philosophy of Science Part A*, v. 41, n. 4, p. 375-385, 2010.

PEIRCE, Charles Sanders. *Collected papers*. Cambridge: The Belnap Press of Harvard University Press, 1934-1935. v. 5-6.

POINCARÉ, Henry. (1905). "Science and Hypothesis". In: *The Value of Science: Essential Writings of Henri Poincaré* (2001). New York: Dover Publications, p. 1-178.

POPPER, Karl R. *Conjectures and refutations*. Londres: Routledge & Kegan Paul, 1963.

PUTNAM, Hilary. "Meaning and reference". *The journal of philosophy*, v. 70, n. 19, p. 699-711, 1974.

PUTNAM, Hilary. What is mathematical truth? In: _____ *Philosophical Papers*, v. 1, 1975a.

PUTNAM, Hilary. "The meaning of meaning". In: _____ *Philosophical papers*, v. 2, 1975b.

PUTNAM, Hilary. *Meaning and the Moral Sciences*. London: Routledge, 1978.

PSILLOS, Stathis. "Kitcher on reference". *International Studies in the Philosophy of Science*, v. 11, n. 3, p. 259-272, 1997.

PSILLOS, Stathis. *Scientific Realism: how science tracks truth*. New York: Routledge, 1999.

PSILLOS, Stathis. "Causal descriptivism and the reference of theoretical terms". In: RAFTOPOULOS, Athanassios; MACHAMER, Peter (Ed.) *Perception, realism, and the problem of reference*. Cambridge University Express, p. 212-238, 2012.

QUINE, Willard van Orman. "Two dogmas of empiricism" (1951). In: _____ *From a logical point of view*. Cambridge: Harvard University Press, 1953.

QUINE, Willard V. "On empirically equivalent systems of the world". *Erkenntnis*, p. 313-328, 1975.

RAATIKAINEN, Panu (2007). "Theories of reference and the philosophy of Science". Paper presented at EPSA07: 1st Conference of the European Philosophy of Science Association Madrid, 15–17 November, 2007.

RAMSEY, Frank P. "Theories" (1929). In: BRAITHWAITE, Richard B. (ed.) *The foundations of mathematics and other logical essays*, 1931.

RUSSELL, Bertrand. “On denoting”. *Mind*, v. 14, n. 56, p. 479-493, 1905.

RUSSELL, Bertrand. “The Philosophy of Logical Atomism” (1918). In: MARSH, R. (org.) *Logic and Knowledge*. Londres: Allen & Unwin, 1956.

SANKEY, Howard. “Incommensurability: The current state of play”. *Theoria. Revista de Teoría, Historia y Fundamentos de la Ciencia*, v. 12, n. 3, p. 425-445, 1997.

SARDI, G. C.; de Oliveira Silva, D. “Realismo, continuidade teórica e a revolução química”. *Sapere Aude*, v. 13, n.26, p. 575-590, 2022.

SEARLE, John R. “Proper names”. *Mind*, v. 67, n. 266, p. 166-173, 1958.

SILVA, Débora de Oliveira. *Realismo científico e descontinuidade teórica: uma análise da relação entre sucesso teórico, verdade aproximada e referência de termos teóricos*. 2022. 49 f. Monografia (Graduação) – Instituto de Filosofia, Arte e Cultura, Universidade Federal de Ouro Preto, Ouro Preto, 2022. Documento Eletrônico. Disponível em: <<http://www.monografias.ufop.br/handle/35400000/3942>>. Acesso em: 30 de ago. de 2023.

SMART, John Jamieson Carswell. *Philosophy and Scientific Realism*. London: Routledge & Kegan Paul, 1963.

STANFORD, P. Kyle. “Refusing the devil's bargain: What kind of underdetermination should we take seriously?”. *Philosophy of Science*, v. 68, n. S3, p. S1-S12, 2001.

STANFORD, P. Kyle. “Pyrrhic Victories for Scientific Realism”. *The Journal of Philosophy*, v. 100, n. 11, p. 553-572, 2003.

STANFORD, P. Kyle. *Exceeding our grasp: Science, history, and the problem of unconceived alternatives*. Oxford University Press, 2006.

STRAWSON, G. *The Secret Connexion*. Oxford: Clarendon Press, 1989.

TURNER, Derek. *Making prehistory: Historical science and the scientific realism debate*. Cambridge University Press, 2007.

VAN FRAASSEN, Bas C. *The scientific image*. Oxford University Press, 1980.

VICKERS, Peter. "A confrontation of convergent realism". *Philosophy of Science*, v. 80, n. 2, p. 189-211, 2013.

WHITTAKER, Edmund Taylor. *A History of the Theories of Aether and Electricity from the Age of Descartes to the Close of the Nineteenth Century*. Longmans, Green and Company, 1910.

WORRALL, John. "Structural realism: The best of both worlds?". *dialectica*, v. 43, n. 1-2, p. 99-124, 1989.

WRAY, K. Brad. "Success of science as a motivation for realism." In: STAATSI, Juha (Ed.) *The Routledge handbook of scientific realism*. Routledge, p. 37-47, 2017.

WRAY, K. Brad. *Resisting scientific realism*. Cambridge University Press, 2018.