



UNIVERSIDADE FEDERAL DO CEARÁ
INSTITUTO DE CULTURA E ARTE
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA

WISLEY SOUSA LIMA

O PROBLEMA DA REPRESENTAÇÃO MENTAL COMO COVARIAÇÃO

FORTALEZA
2023

WISLEY SOUSA LIMA

O PROBLEMA DA REPRESENTAÇÃO MENTAL COMO COVARIAÇÃO

Dissertação apresentada a Coordenação do Programa de Pós-Graduação em Filosofia do Instituto de Cultura e Arte da Universidade Federal do Ceará como requisito parcial para a obtenção do título de Mestre em Filosofia. Área de concentração: Filosofia.

Orientador: Prof. Dr. Cícero Antônio Cavalcante Barroso.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

L711p Lima, Wisley Sousa.
O Problema da Representação Mental Como Covariação / Wisley Sousa Lima. – 2023.
72 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Instituto de cultura e Arte,
Programa de Pós-Graduação em Filosofia, Fortaleza, 2023.
Orientação: Prof. Dr. Cícero Antônio Cavalcante Barroso.

1. Representação mental. 2. Covariação. 3. Representação simulada. 4. Semântica interpretativa. I. Título.

CDD 100

WISLEY SOUSA LIMA

O PROBLEMA DA REPRESENTAÇÃO MENTAL COMO COVARIAÇÃO

Dissertação apresentada a Coordenação do Programa de Pós-Graduação em Filosofia do Instituto de Cultura e Arte da Universidade Federal do Ceará como requisito parcial para a obtenção do título de Mestre em Filosofia. Área de concentração: Filosofia.

Aprovada em: 17/07/2023

BANCA EXAMINADORA

Prof. Dr. Cícero Antônio Cavalcante Barroso (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. José Gladstone de Almeida Júnior
Universidade Federal do Cariri (UFCA)

Prof. Dr. Ralph Leal Heck
Universidade Federal do Ceará (UFC)

AGRADECIMENTOS

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

Ao Prof. Dr. Cícero Antônio Cavalcante Barroso pela generosidade, paciência e excelente orientação.

Aos professores participantes da banca examinadora, Prof. Dr. José Gladstone de Almeida Júnior e Prof. Dr. Ralph Leal Heck.

À minha família.

A Fabiano Farias de Medeiros.

RESUMO

A partir da modernidade, as teorias covariacionistas da representação mental passaram a ser mais aceitas frente as teorias da similaridade, pois a covariação, ao afirmar que o conteúdo semântico das representações mentais é determinado pelos elos causais entre estas representações e o mundo, resolve alguns problemas típicos das teorias da similaridade, como o problema da abstração. Porém, as teorias covariacionistas apresentam sérias dificuldades para explicar como os erros de representação são possíveis. Além disso, os meios para lidar com esse problema são limitados, restringindo-se a estipulações sobre condições ideais de funcionamento do aparato cognitivo ou circunstâncias ideais de percepção. Robert Cummins argumenta que as teorias covariacionistas têm em comum a característica de serem descritíveis por uma metáfora computacional, mas, ao mesmo tempo, não são capazes de definir a representação de modo satisfatório dentro da Teoria Computacional da Mente. A solução proposta por Cummins é abandonar o covariacionismo em favor de uma semântica interpretativa que requer representações simuladas para fundamentar a representação. No entanto, a análise do debate entre Cummins e as teorias covariacionistas de John Locke e Jerry Fodor aponta para o fato de que as críticas de Cummins ao covariacionismo são por vezes inadequadas ou falhas e não encerram definitivamente o debate. Além disso, sua semântica interpretativa apresenta problemas para conectar a mecânica da representação simulada com a intencionalidade requerida pelas representações mentais. Esses fatores levam à conclusão de que tanto a crítica de Cummins não é pertinente quanto sua proposta de solução, embora interessante, não é capaz de concluir o debate sobre a representação mental de forma satisfatória.

Palavras-chave: representação mental; covariação; representação simulada; semântica interpretativa.

ABSTRACT

From the modern era, covariationist theories of mental representation have become more accepted in comparison to similarity theories because covariance, by asserting that the semantic content of mental representations is determined by the causal links between these representations and the world, solves some typical problems of similarity theories, such as the problem of abstraction. However, covariationist theories present serious difficulties in explaining how representation errors are possible. Moreover, the means to address this problem are limited, being restricted to stipulations about ideal conditions for cognitive functioning or ideal circumstances of perception. Robert Cummins argues that covariationist theories share the characteristic of being describable by a computational metaphor, yet they are unable to satisfactorily define representation within the Computational Theory of Mind. Cummins proposes abandoning covariationism in favor of an interpretive semantics that requires simulated representations to underpin representation. However, an analysis of the debate between Cummins and covariationist theories by John Locke and Jerry Fodor indicates that Cummins' criticisms of covariationism are sometimes inadequate or flawed and do not definitively settle the debate. Furthermore, his interpretive semantics poses problems in connecting the mechanics of simulated representation with the intentionality required by mental representations. These factors lead to the conclusion that both Cummins' critique is not pertinent and his proposed solution, although interesting, fails to conclusively address the debate on mental representation satisfactorily.

Keywords: mental representation; covariation; simulated representation; interpretational semantics.

SUMÁRIO

1	INTRODUÇÃO	7
1.1	A Teoria Computacional da Mente.....	10
1.2	A representação mental como covariação	11
1.3	As limitações do covariacionismo.....	12
1.4	Estrutura e método desta pesquisa	15
2	O COVARIACIONISMO DE LOCKE	16
2.1	O Ensaio Acerca do Entendimento Humano	16
2.2	A teoria das ideias	18
2.3	A teoria da linguagem.....	19
2.3.1	<i>Do uso dos termos gerais</i>	23
2.4	De como a ideia representa o objeto.....	24
2.5	Como Locke pode responder a Cummins?.....	29
3	O COVARIACIONISMO DE FODOR	32
3.1	A psicologia do senso comum	33
3.2	A Teoria Representacional da Mente	38
3.3	A Teoria Causal Bruta	39
3.4	Como Cummins critica a teoria de Fodor	42
3.5	Como Fodor poderia responder a Cummins	45
4	A SEMÂNTICA INTERPRETATIVA DE CUMMINS	50
4.1	Objecções à Semântica Interpretativa	54
4.1.1	<i>Naturalismo</i>	54
4.1.2	<i>O status causal do conteúdo de representações simuladas</i>	55
4.1.2.1	<i>Conteúdo e causalidade</i>	55
4.1.2.2	<i>Causalidade do conteúdo e explicação pelo conteúdo</i>	59
4.1.3	<i>Intencionalidade</i>	60
4.1.3.1	<i>Representações simuladas e intencionalidade</i>	60
4.1.3.2	<i>Se a Teoria Computacional da Mente pode ignorar a intencionalidade</i> ...	63
4.1.3.3	<i>Introduzindo a intencionalidade</i>	64
4.2	Avaliando a Semântica Interpretativa	66
5	CONCLUSÃO	70
	REFERÊNCIAS	71

1 INTRODUÇÃO

Para Robert Cummins (1989, p. 1-10), é necessário distinguir dois problemas a respeito das representações mentais. O *problema das representações* (no plural), é um problema teórico da ciência empírica e diz respeito ao modo pelo qual os processos e estados mentais atuam como representações. Já o *problema da representação* (no singular), é um problema paradigmático da filosofia que diz respeito à natureza da representação mesma e pergunta pelo que faz com que algo represente outra coisa.

Apenas quatro respostas já foram sugeridas para o problema das representações:

- a) a teoria aristotélica escolástica, que diz que a representação mental é o compartilhamento da forma substancial entre a mente e o objeto – ou seja, conhecemos uma substância material quando a forma (princípio de inteligibilidade) que informa uma porção de matéria (princípio de individuação) também informa nossa mente, de modo que ela se torna qualitativamente idêntica àquilo que os sentidos captaram diretamente do objeto, sendo uma representação completamente transparente;
- b) filósofos como Berkeley e Hume acreditavam que a representação mental é uma imagem do objeto – a mente representa os objetos através da recordação da imagem sensível que eles uma vez produziram, seja ela uma imagem pictórica, tátil, gustativa, auditiva, etc;
- c) filósofos como Hobbes sugeriram que a representação mental é um símbolo linguístico, uma espécie de descrição simbólica do objeto;
- d) a teoria mais aceita atualmente vem da neurociência e afirma que as representações mentais são estados neurofisiológicos, ou seja, estados que só são descritos e realizáveis em sistemas biológicos.

A teoria aristotélica escolástica e as teorias do tipo da de Berkeley e Hume se fundamentam na similaridade, ou seja, no compartilhamento de propriedades entre a mente e o objeto. A desvantagem desse tipo de teoria é sua dificuldade em lidar com o problema da abstração, pois dificulta a explicação de

como uma mesma representação mental pode representar (isto é, ser semelhante a) diferentes objetos.

Vejamos um exemplo: se queremos pensar sobre os gatos, precisamos ter uma imagem, uma ideia que compartilha propriedades com os gatos. Mas como tal imagem pode representar ao mesmo tempo os gatos mais parecidos com ela e os mais exóticos? Gatos com os quais não estamos habituados a ver ou que talvez nunca tenhamos visto normalmente deveriam ser reconhecidos como gatos, mas se o que faz a representação é a semelhança, gatos muito diferentes não podem ser reconhecidos. Há ainda a questão de que ideias e gatos são coisas muito diferentes. Se acreditamos que o cérebro é um intermediário entre as coisas extramentais e a mente, que nossas ideias de algum modo são estados físicos no cérebro, não é possível (ou, pelo menos, não é nada simples) explicar como “a excitação do conjunto de neurônios x ” seja semelhante ou compartilhe propriedades com os gatos.

As teorias do tipo adotado por Hobbes tem como vantagem evitar a semelhança, ou seja, elas procuram se distanciar das condições e características individuais dos objetos, a fim de pensar uma representação mais abstrata, simplificada e generalizada. A vantagem de permitir a abstração é que, ao abstrair as características particulares de um objeto, é possível identificar os elementos essenciais e as relações fundamentais que estão em jogo sem cair nos mesmos problemas das teorias da similaridade.

Na medida em que se atêm aos aspectos abstratos dos estados mentais, essas teorias os analisam como representações simbólicas, ao mesmo tempo em que tratam os processos mentais como processos computacionais, ou seja, processos de manipulação simbólica. Dessa forma, a mente pode ser descrita como um sistema de representação que gera estruturas de dados por meio do encadeamento de entradas e saídas.¹ Essa característica computacional da mente é particularmente relevante para algumas teorias mais recentes sobre a mente que se valem da ciência da computação para descrever os processos e estados mentais. A computacionalidade permite formalizações sistemáticas e precisas que podem ser

1 Em computação, o termo “entrada” se refere à inserção de dados em algum software ou hardware. O termo “saída” se refere ao resultado da execução de algum programa a partir de uma entrada. O termo “estrutura de dados” denota o modo de armazenamento e organização de dados em um computador. Existem diversas estruturas de dados, como filas, listas, pilhas, árvores, etc.

descritas através de modelos matemáticos e simulações computacionais, possibilitando explorar as implicações e resultados dessas teorias de forma prática e eficiente.

A teoria dos estados neurofisiológicos tem as mesmas vantagens, com a diferença de que sua descrição da representação mental não se aplica a nenhuma arquitetura de computador, mas apenas a estruturas biológicas como o cérebro.

Quanto ao problema da representação, também só há quatro abordagens que se propõem a explicar a representacionalidade:

- a) similaridade – como já explicado acima, é o compartilhamento de propriedades entre a representação mental e o objeto representado;
- b) covariação – a representação se dá pelo efeito regular que um certo objeto causa na mente, não importando que efeito é este, mas apenas seu aparecimento contíguo ao objeto;
- c) papel adaptativo – é um caso de covariação mais brando em que a representação normalmente covaria com um objeto que é sua correspondência padrão, mas pode também covariar com outras coisas, caso isso cumpra uma função importante, por exemplo, as abelhas dançam para indicar às outras a direção onde encontraram alimento, mas, caso seja necessário, também podem executar a mesma dança para levar as companheiras a um local seguro;
- d) papel funcional – para o funcionalismo, os estados mentais não são constituídos por sua estrutura interna, mas pela função que eles desempenham no sistema cognitivo, relacionando estímulos sensoriais, outros estados mentais e comportamentos, assim, estruturas biológicas baseadas em carbono ou silício, ou mesmo máquinas, podem ter os mesmos estados mentais, contanto que esses estados cumpram a mesma função, causando comportamentos de dor, desejo, medo, etc (Cf. LEVIN, 2021).

O tema deste trabalho se restringe ao problema da representação, mais especificamente ao problema da representação como covariação causal, visto que a covariação se mostra mais promissora frente às teorias da similaridade e é plenamente compatível com os papéis adaptativos e funcionais, estabelecendo uma

espécie de base conceitual comum.

Esta pesquisa se justifica pelo fato de que as teorias empíricas da cognição, como a atual neurociência, por exemplo, geralmente tomam a noção de representação como um primitivo explicativo. Porém, se a noção de representação não estiver devidamente explicada será impossível demonstrar sua compatibilidade com a teoria empírica que a pressupõe. Portanto, tal teoria deve ser considerada mal fundamentada e, deste modo, inadequada, pois não pode fornecer um entendimento completo sobre as representações mentais, contrariamente a sua motivação inicial (DENNETT, 1978, *passim*).

Assim, para resolver o problema das representações é preciso antes resolver o problema da representação. Nesse sentido, esta pesquisa deverá contribuir para o debate do problema da representação investigando a representacionalidade pela covariância, já que a teoria covariacionista pode lidar satisfatoriamente com o problema da abstração e pode servir de base conceitual comum também ao funcionalismo e à teoria de papéis adaptativos.

1.1 A Teoria Computacional da Mente

Cummins (1989, p. 12-13) explica que, para compreender a noção de representação mental sob um dado arcabouço teórico, é preciso entender o papel explicativo que tal arcabouço atribui a essa mesma noção de representação mental. Por isso, perguntar pela natureza das representações mentais é simplesmente fazer uma pergunta mal delimitada. A quem pergunta: qual a natureza da representação mental? É preciso pedir a especificação: em qual arcabouço/referencial teórico?

O referencial ou arcabouço teórico que delimita nossa indagação pela natureza da representação neste trabalho é a *Teoria Computacional da Mente* (doravante TCM). Essa teoria assume que sistemas cognitivos são sistemas formais de interpretação automática. Portanto, sistemas que tomam a cognição como manipulação disciplinada de símbolos.

Uma vez estabelecido esse referencial teórico, podemos então perguntar pelo papel explicativo que a noção de representação mental cumpre nesse referencial e, uma vez fundamentado esse papel explicativo, podemos dizer no que consiste a relação de representação.

As teorias computacionais, de um modo geral, respondem ao problema

das representações (no plural), ao assumir que as representações são estruturas de dados simbólicas tal como compreendidas pela ciência da computação. A ciência da computação demonstra como essas estruturas podem ser instanciadas e manipuladas, mas o modo como essas estruturas são fisicamente instanciadas não interessa a essas teorias. Não importa se a instanciação é numa arquitetura de Von Newman, numa Máquina de Turing, num cérebro ou num alienígena biologicamente baseado em sílica.

Nesse sentido, perguntar o que é, para uma representação mental, ter um conteúdo e o que determina qual conteúdo ela tem, no contexto da TCM, equivale a perguntar o que faz de uma estrutura de dados uma representação e o que determina aquilo que essa estrutura de dados representa. Respondendo a essas perguntas teremos respondido ao problema da representação (no singular).

Para tanto, iremos nos deter, como dito acima, na representação por covariação, mas abordaremos também a proposta de Cummins, que é uma concepção de representação como papel funcional, sempre verificando se cada uma dessas concepções satisfaz às exigências da TCM e explica de modo satisfatório a representação.

1.2 A representação mental como covariação

As teorias da representação mental que se fundam na covariação afirmam, de um modo geral, que o conteúdo semântico das representações mentais é determinado pelos elos causais entre estas representações e o mundo, isto é, que nossas ideias são adequadas porque são produzidas de modo regular e natural por causas externas (CUMMINS, 1989, p. 35).

Por não apelar à semelhança para fundamentar a representação, as teorias covariacionistas evitam o problema da abstração com que as teorias da similaridade (como as de Berkeley, Hume ou da escolástica) precisam lidar. Além disso, as teorias covariacionistas resolvem o problema da abstração, isto é, podem lidar com representações abstratas, uma vez que não há necessidade de que o representante seja simultaneamente semelhante a todos os representados (CUMMINS, 1989, p. 38).

Porém, apesar de apresentar algumas vantagens, as teorias covariacionistas apresentam sérias dificuldades para explicar como as deturpações

(erros de representação) são possíveis, pois os casos de deturpação são os casos em que a covariação falha, mas se a covariação faz a representação, os casos de deturpação não podem ser considerados representações genuínas, de modo que o erro parece ser impossível no covariacionismo. Além disso, os meios para lidar com esse problema são limitados, restringindo-se a estipular condições ideais de funcionamento do aparato cognitivo (a deturpação ocorre quando o aparato cognitivo falha) ou circunstâncias ideais de percepção (a deturpação ocorre quando as circunstâncias ambientais impedem a correta percepção por parte do aparato cognitivo).

Um exemplo famoso é o *Quarto de Ames*, que consiste numa sala que, vista de frente, parece ser cúbica, com uma parede de fundo e duas paredes laterais paralelas entre si e perpendiculares ao chão e ao teto. Porém, a sala é, na verdade, trapezoidal, com paredes inclinadas e um dos cantos é mais próximo da frente que o outro. Isso cria uma ilusão de perspectiva que parece ampliar o que está em um canto e encolher o que está no outro, de modo que, uma pessoa que caminhe de um canto a outro da sala parece crescer ou encolher a depender da direção tomada (CUMMINS, 1989, p. 39-40).

1.3 As limitações do covariacionismo

Para Robert Cummins (1989, p. 35-40), as versões contemporâneas das teorias covariacionistas podem ser facilmente compreendidas e criticadas a partir da teoria lockeana acerca do uso classificatório das palavras gerais (T_L), apresentada no livro III do Ensaio (E III.iii.1-20).

Apesar de acreditar que essa teoria protótipo corresponde ao real intento de Locke, Cummins não se compromete de forma alguma com a veracidade histórica de sua atribuição. Ele afirma “[...] eu não me importo. O que eu quero é uma versão simples, clara e honesta, do tipo de teoria que fundamenta a representação na covariação” (1989, p. 35, tradução nossa).²

A ideia de Cummins é que T_L serve como um protótipo para o qual as teorias covariacionistas contemporâneas podem ser reduzidas, de modo que, compreendendo os erros desse protótipo, é possível criticar *qualquer um* de seus

2 “[...] I don't really care. What I want is a clear and fairly simple version of the sort of theory that founds representation on covariance.”

diferentes desenvolvimentos.

T_L pretende resolver a aparente incongruência entre estes dois fatos: que é possível que qualquer símbolo tenha qualquer significado mesmo que as palavras não designem as coisas com uma correspondência exata; e que, apesar disso, é possível usar uma palavra incorreta ou falsamente. Por exemplo, como é possível que a palavra 'gato' designe Tom mesmo que não se encaixe a ele melhor que 'lápiz', 'cadeira' ou qualquer outra? E por que as palavras 'cadeira' e 'lápiz' não servem para designá-lo?

T_L soluciona estes problemas explicando as propriedades semânticas dos símbolos comunicativos em termos de propriedades semânticas das representações mentais. Segundo T_L , ao aprendermos um idioma qualquer, como o português, aprendemos que, na comunidade dos falantes dessa língua, o termo 'gato' é convencionalmente associado com uma ideia abstrata (conceito) que carrega uma relação semântica natural e não-convencional de concordância com todos (e apenas com) os gatos.

É o fato de que uma palavra geral é convencionalmente associada com determinada ideia abstrata que dá a esse termo suas condições de satisfação. Portanto, de acordo com T_L , são os conceitos que possuem uma correspondência exata no mundo, enquanto as palavras significam os conceitos em virtude de uma associação puramente convencional.

T_L pode ser interpretada como uma descrição computacional do uso classificatório de termos gerais. Considere o computador de cartões perfurados LOCKE, que perfura um padrão de furos num cartão (percepto) para cada objeto captado por seu sistema perceptual, e que, além disso, é capaz de comparar perceptos com um conjunto de cartões armazenados (conceitos), cujo verso possui uma palavra escrita.

Uma vez que LOCKE seja confrontado por um objeto, como o gato Tom, um percepto será produzido e LOCKE buscará entre os conceitos um cartão com o mesmo padrão de furos; caso encontre, exibirá o verso deste conceito (que no caso de Tom deverá conter a palavra 'gato'). Em suma, temos que:

(L₁) x representa y em LOCKE =_{df} x é um padrão de furos que ocorre num percepto quando, somente quando e porque LOCKE é confrontado por y .

O principal problema com T_L é a impossibilidade de explicar erros de percepção. Supondo que LOCKE seja confrontado por Tom e produza um percepto

de cão (padrão de furos D), não seria verdade, de acordo com L_1 , que D ocorre quando, somente quando e porque um cão está presente; portanto, D não representaria cães e LOCKE não teria gerado uma percepção de cão, contrariamente à suposição inicial. Desse modo, LOCKE não pode errar porque, dado L_1 , toda representação deturpada é uma noção incoerente; mas como os erros de percepção são possíveis e mesmo comuns em nossa experiência do mundo, é preciso assumir que T_L é falsa.

Ainda de acordo com Cummins (1989, p. 41-46), os teóricos do covariacionismo só têm dois modos de evitar esse problema. Ou estipulando condições ideais de percepção, ou estipulando condições ideais de funcionamento do aparato cognitivo. Ambos, porém, falham em corrigir esse defeito da teoria. Supondo que o erro perceptual de LOCKE se deu por um mal funcionamento, teremos:

(L_2) x representa y em LOCKE =_{df} se LOCKE está funcionando corretamente, então um padrão de furos x ocorre num percepto quando, somente quando, e porque LOCKE é confrontado por y .

L_2 permite deturpações porque introduz uma propriedade modal: LOCKE pode erroneamente reconhecer Tom como um cão, mas se estivesse funcionando devidamente ele produziria o padrão de furos correto. Este acréscimo, contudo, não é suficiente para lidar com situações em que os erros de representação são causados pelo ambiente, como as ilusões de ótica, que fazem os objetos parecerem diferentes do que realmente são, causando ilusões de tamanho, cor, forma, etc. Por isso é ainda necessário acrescentar uma cláusula para lidar com esse problema:

(L_3) x representa y em LOCKE =_{df} se LOCKE está funcionando corretamente e as circunstâncias de percepção são ideais, então x ocorre num percepto quando, somente quando, e porque LOCKE é confrontado por y .

Esta nova definição permite todo tipo de deturpação, mas introduz uma circularidade insolúvel: as condições de percepção são ideais se LOCKE representa corretamente o objeto, mas a definição de *representar* em LOCKE, isto é, L_3 , inclui a noção de condições ideais de percepção. Desse modo, L_3 depende da definição de *condição ideal* e esta definição, por sua vez, depende de L_3 .

Uma vez exposta a “teoria-protótipo” T_L e seus problemas, o principal desafio para Cummins é mostrar que outras teorias covariacionistas mais elaboradas também sofrem dos mesmos defeitos de T_L . Para tanto, abordaremos

adiante apenas um pequeno recorte, dados os limites e o escopo deste trabalho. Por isso nos limitamos a explorar as teorias covariacionistas de John Locke e Jerry Fodor.

1.4 Estrutura e método desta pesquisa

Para fins de investigação científica, partimos da hipótese de que a crítica de Cummins ao covariacionismo é pertinente e que a solução que ele propõe para o problema da representação está correta. Nosso objetivo é demonstrar a pertinência da crítica de Cummins às teorias da representação mental como covariação e como a solução que ele propõe para o problema da representação se mostra adequada.

Esta pesquisa pressupõe dois aspectos metodológicos: um histórico e um teórico. O aspecto histórico se dá pela aplicação de análise bibliográfica, considerando uma bibliografia primária, composta pelos textos de Locke, Fodor e Cummins que são fundamentais para o tema desta pesquisa, e uma bibliografia secundária, composta de textos suplementares. O aspecto teórico se dá pela análise conceitual e argumentação filosófica sobre o material bibliográfico levantado.

Os capítulos seguintes estão estruturados da seguinte maneira: o capítulo segundo visa fornecer uma compreensão geral do Ensaio Acerca do Entendimento Humano³ e da teoria lockeana do uso classificatório das palavras gerais, a fim de verificar se ela corresponde a T_L e identificar possíveis repostas a objeções já adiantadas por Locke. O capítulo terceiro traz uma breve exposição da teoria de Fodor e demonstra se é viável reduzir esta teoria a T_L . O capítulo quarto apresenta a solução proposta por Cummins para o problema da representação e possíveis contra-argumentos, além de uma avaliação crítica do debate na qual se corrige a hipótese inicial adotada, uma vez que as críticas de Cummins nem sempre são adequadas e que sua proposta de solução ao problema da representação mental tampouco encerra o debate de modo satisfatório. Por fim, o capítulo quinto conclui este trabalho retomando os principais pontos onde se critica as teses de Cummins e reiterando nosso posicionamento em relação ao debate.

3 Locke (1996).

2 O COVARIACIONISMO DE LOCKE

Não há consenso entre os comentadores acerca do significado de termos fundamentais que Locke usa no Ensaio (LENNON, 2007, p. 231), de modo que há mais de uma interpretação possível para o conjunto de teses que ele encerra. Por isso o objetivo deste capítulo não é julgar a interpretação de Cummins como certa ou errada, mas mostrar se ela é reconhecida total ou parcialmente pelos comentadores, quais são as suas consequências, e quais as possíveis soluções e respostas de Locke capazes de dissipar a crítica de Cummins.

Para tanto, exporemos nas próximas seções uma visão geral do Ensaio, sua história e suas teses fundamentais. Depois, veremos a interpretação de Cummins, classificando-a, se possível, dentro da tradição dos comentadores de Locke. Por fim, avaliaremos o debate entre as posições de Locke e Cummins.

2.1 O Ensaio Acerca do Entendimento Humano

O Ensaio é uma obra caudatária do giro epistemológico cartesiano que inaugurou a modernidade filosófica através de uma profunda transformação no modo de se fazer teoria do conhecimento, transformação esta que levou ao gradual abandono do *modus operandi* aristotélico e escolástico. De fato, parece ter sido o cartesianismo que primeiramente influenciou Locke a purgar tanto quanto possível de sua filosofia o tradicional aristotelismo e seu apelo ao senso comum (DOWNING, 2016, p. 100-101; SHUURMAN, 2010, *passim*).

Em vista disso, uma das principais preocupações do Ensaio era fornecer um novo fundamento epistemológico capaz de resolver, através da distinção entre os conceitos de ideia e objeto, a discrepância entre duas imagens do mundo: uma imagem manifesta, imediatamente acessível, e uma imagem científica que emergia dos avanços da ciência moderna e explicava aquilo que é de primeira evidência em termos de dimensões, estruturas e movimento (LENNON, 2007, p. 231).

Não há dúvida de que essa nova imagem científica do mundo, de que Locke pretende dar conta, foi moldada em grande parte pelo trabalho do irlandês Robert Boyle (1621–1691) e sua hipótese corpuscular. Vejamos, conforme a exposição de Hill (2010, p. 47-49; 2010, p. 200), como o igualmente antiaristotélico Boyle influenciou na composição do Ensaio.

Boyle criticava os escolásticos principalmente por estes quatro motivos: eles (na visão de Boyle) evitavam a investigação prática, prendiam-se a uma visão aristotélica de natureza (composta de quatro elementos, água, fogo, terra e ar), apelavam a qualidades ocultas (conhecidas apenas por seus efeitos) e ao obscuro conceito de “forma substancial”, isto é, aquilo que para os aristotélicos imbuía a matéria de certas qualidades e poderes.

Por isso Boyle procurou desenvolver e ampliar o uso dos experimentos científicos seguindo o método empírico de Francis Bacon (1561–1626), e substituir os elementos e a forma substancial dos escolásticos por uma hipótese mais simples que não apelava a qualidades ocultas: a hipótese corpuscular, que explicava os fenômenos físicos pela interação mecânica de corpúsculos imperceptíveis, caracterizados apenas por seu movimento, tamanho e forma.

A ontologia da hipótese corpuscular de Boyle teve um impacto profundo no trabalho de Locke, tendo um papel significativo na distinção de conceitos fundamentais do Ensaio, como a distinção entre as essências nominais e reais das substâncias ou a distinção entre as qualidades primárias dos corpos (materialmente presentes e intrínsecas às partículas que compõem a natureza) e as qualidades secundárias dos corpos (derivadas das primárias e sem existência autônoma fora da percepção).

Locke considerava que só a hipótese corpuscular era minimamente inteligível e filosoficamente aceitável, pois partia do princípio de que apenas o contato mecânico, que depende exclusivamente das qualidades primárias dos corpos envolvidos, é suficientemente claro para a inteligência humana, sendo a hipótese corpuscular, assim, a única forma possível de entender o poder causal.

Destarte, parece seguro inferir da exposição acima que foi a influência de Boyle sobre Locke que permitiu a elaboração do conceito geral de ideia que fundamenta o Ensaio e o necessário desvencilhamento do aristotelismo.

Conforme explica Losonsky (2007, p. 286-288), o conceito de *ideia* é o mais importante e fundamental conceito do Ensaio. Seu papel é tornar compreensíveis tanto a natureza e os limites do conhecimento humano quanto desenvolver uma teoria da linguagem, pois, para Locke, a linguagem desempenha um papel fundamental na cognição. Ele afirma: “há uma conexão tão estreita entre ideias e palavras [...] que é impossível falar clara e distintamente de nosso conhecimento [...] sem considerar, primeiro, a natureza, uso, e significação da

linguagem” (E II.xxxiii.19, tradução nossa).⁴ Veremos a seguir como Locke elabora sua teoria das ideias e sua teoria da linguagem.

2.2 A teoria das ideias

Locke considera as ideias como os únicos objetos imediatos da mente, através dos quais se conhece os objetos extramentais, ou seja, nós conhecemos as coisas à medida que sejam representadas pelas ideias (LENNON, 2007, p. 231; NEWMAN, 2007, p. 317; SHUURMAN, 2010, p. 176).

Chappell (1994, p. 27-30) explica, mais especificamente, que as ideias existem na parte intelectual da mente, chamada “entendimento”, e são os objetos das operações mentais de pensar ou perceber algo. Dizer que uma ideia é o objeto de uma percepção ou pensamento é dizer que todo pensamento ou percepção é sobre algo, pois é da natureza do pensamento e da percepção ser direcionado a algo, ter um assunto ou alvo.

Locke usa a palavra “objeto” justamente para se referir a esse alvo ou assunto requerido: o objeto de um pensamento ou percepção é aquilo acerca de que o pensamento ou percepção diz respeito, muito embora não seja necessário que a percepção tenha um objeto não imediato, extramental, pois as próprias ideias podem ser objetos internos da percepção (o que justificaria as alucinações, ou seja, casos em que a percepção ocorre sem uma coisa extramental percebida).

Locke distingue dois tipos principais de ideia: as ideias simples, e as ideias complexas. As ideias simples são aquelas que nos são dadas passiva e diretamente pelos sentidos, como as cores, os sons, a temperatura, etc. Ou seja, são as ideias que a mente capta diretamente dos sentidos sem imprimir nelas qualquer ação. Já as ideias complexas são conjuntos de ideias simples que a mente ativamente compõe e que se distinguem de acordo com sua composição como, por exemplo, as ideias de substância: gato, homem, cera, etc; ou as de relação: igual a, maior que, etc. (BOLTON, 2007, *passim*).

Locke considera que *as ideias são signos dos objetos* porque são causadas por suas qualidades, ou seja, aquilo nos objetos que tem o poder de

4 “*there is so close a connexion between ideas and words [...] that it is impossible to speak clearly and distinctly of our knowledge [...] without considering, first, the nature, use, and signification of language*”.

causar certa ideia no entendimento. Ele escreve: “[...] uma bola de neve tem o poder de produzir em nós as ideias de branco, frio e redondo, os poderes para causar tais ideias em nós, tal como estão na bola de neve, eu chamo qualidades” (E II.viii.8, tradução nossa).⁵

Há dois tipos de qualidades: as primárias e as secundárias. As qualidades primárias não são poderes dos corpos, mas pertencem a eles tal como são, sendo deles inseparáveis e semelhantes. São qualidades primárias a extensão, a forma, o movimento, etc. Já as qualidades secundárias são poderes que os corpos materiais têm para produzir ideias no entendimento e que não se assemelham àquilo que eles são na realidade extramental. São qualidades secundárias a cor, o sabor, etc.

As qualidades secundárias se fundamentam nas primárias, ou seja, os poderes dos corpos são provenientes das propriedades intrínsecas das partículas imperceptíveis, de modo que o odor, por exemplo, seria causado pelo tamanho, forma e movimento dos minúsculos corpúsculos que compõem sua efusão. Por isso só as qualidades primárias são propriedades categóricas dos corpos e o fundamento de seus poderes causais, poderes entre os quais estão as qualidades secundárias (HILL, 2010, p. 199-200).

Essa mecânica das qualidades e dos poderes causais dos corpos não se aplica diretamente à produção de ideias complexas, como, por exemplo, as ideias de substância, pois são em parte produto de uma atividade voluntária da mente sobre os dados dos sentidos, sendo constituídas de forma possivelmente variada entre os indivíduos, ainda que as características dessas ideias permitam aferir sua concordância com aquilo que representam, já que são construídas a partir de certos padrões de ideias simples que o entendimento unifica e nomeia. Por exemplo, “cera” é o nome comumente anexado àquela ideia de substância construída a partir do fato de que a moldabilidade, a fusibilidade, a cor amarela e o caráter combustível são encontrados juntos na experiência. (LOSONSKY, 2007, p. 294-295).

2.3 A teoria da linguagem

Segundo Dawson (2010, p. 175), Locke considera que a linguagem serve a dois usos principais: fixar na memória nossos pensamentos (pois são naturalmente

⁵ “[...] a snowball having the power to produce in us the ideas of white, cold, and round, the powers to produce those ideas in us, as they are in the snowball, I call qualities”.

efêmeros), e comunicá-los a nossos semelhantes. As palavras satisfazem esses dois usos da linguagem porque são sensíveis (visíveis e audíveis) e porque atuam tanto como signos mnemônicos de nossas ideias quanto como signos externos para a comunicação.

O uso que os homens dão a essas marcas serve tanto para recordar os próprios pensamentos [...] quanto para colocá-los ante a visão dos outros: as palavras, em sua significação primária ou imediata, não representam nada além das ideias na mente daquele que as usa, tal como imperfeita ou descuidadamente essas ideias são coletadas das coisas que deveriam representar. (E III.ii.2, tradução nossa).⁶

A teoria da linguagem elaborada por Locke pretende dar conta do fato de que as ideias são signos dos objetos e as palavras são signos das ideias. Para ele, é a significação ou relação semântica que distingue a linguagem verdadeira de meros sons articulados, pois a simples capacidade de produzir esses sons, como acontece com algumas espécies de pássaros, não significa capacidade de linguagem, senão que é preciso que esses sons sejam usados como signos de concepções internas, isto é, como signos de ideias (LOSONSKY, 2007, p. 288-292).

Para Locke, as palavras significam ideias “por uma imposição voluntária pela qual uma palavra é arbitrariamente constituída marca de certa ideia” (E III.ii.1). Nesse sentido, a linguagem começaria com certos indivíduos (como Adão, por exemplo) impondo certos sons a certas ideias, criando assim uma conexão entre eles (E III.vi.45). Essas conexões são gradativamente filtradas durante seu uso por um processo no qual a comunidade de falantes decide o que é válido e o que não é: “o uso comum, por um consentimento tácito, apropria certos sons a certas ideias em todas as línguas, o que limita tanto a significação desse som que, a menos que um homem o aplique à mesma ideia, ele não fala corretamente” (E III.ii.8)

Porém, o propósito de Locke ao examinar a linguagem, explica Dawson (2010, p. 232-233) não era simplesmente insistir que palavras significam ideias, pois isso era um lugar comum em seu tempo, mas sim expor as dificuldades inerentes à essa teoria linguística. Para tanto, Locke se concentrou não só no mau uso da linguagem, mas nas deficiências inerentes às próprias palavras que causam problemas de comunicação. Vejamos a exposição de Dawson.

6 *“The use men have of these marks being either to record their own thoughts [...] and lay them before the view of others: words, in their primary or immediate signification, stand for nothing but the ideas in the mind of him that uses them, how imperfectly soever or carelessly those ideas are collected from the things which they are supposed to represent.”*

Ele estava preocupado sobretudo com os impedimentos que a natureza privada e ideacional do significado impunha à comunicação e referência, mais especificamente com três tipos de problema: (a) obscuridade, vacuidade e “realidade” das palavras; (b) instabilidade semântica; e (c) a prioridade prática e o poder destrutivo das palavras.

As palavras significam ideias que são causadas pelas qualidades das coisas, mas as ideias não captam perfeitamente sua essência real, tal como são fora do entendimento. A imperfeição das palavras é exacerbada pela realidade atribuída aos termos que acabam por substituir a real constituição das coisas. Isto se torna abuso de linguagem quando inadvertidamente as pessoas põem as palavras no lugar das coisas que elas significam, raciocinando a partir de termos vagos, obscuros ou sem significado (Cf. E III.ix.12; E III.x.17-19).

O fato dos sons não serem fixados aos significados causa mal-entendidos. Os falantes podem intencionalmente distorcer o significado das palavras (E III.x.5,12,33-34). Se palavras significam ideias simplesmente, então elas só podem significar as ideias que os falantes individuais têm e, diferindo tais ideias, eles não se comunicam. Como as ideias complexas são construídas por nós e podem diferir de acordo com a experiência individual, as pessoas falam “uma linguagem distinta, embora as mesmas palavras” (E III.ii.3; E III.ix.13,22).

Portanto, o uso comum e os meios pelos quais o significado público é estabelecido, reduz-se a uma variedade sobreposta de usos individuais. Por exemplo, as ideias que têm uma criança e um químico em relação à ideia complexa de substância chamada *ouro*: enquanto aquela provavelmente só sabe que se trata de um metal amarelo e precioso, este conhece muitas outras propriedades, como sua densidade, condutividade, solubilidade em certos ácidos, etc.

Embora as palavras devam representar adequadamente as ideias, na prática, elas dominam não apenas na comunicação mas também no pensamento, que guarda muito mais facilmente seu som que sua carga semântica, mais abstrata e fugidia, ocultando a vacuidade e a instabilidade descritas acima e promovendo absurdos e falhas de comunicação.

Locke enfatiza que a força enganosa das palavras não é simplesmente uma função do desejo corrupto de enganar, mas é a consequência inexorável da linguagem e da psicologia humana (E III.ii.4-5; E III.ix.9; E III.x.4,14,22). É porque as palavras não expressam as ideias que representam, mas usurpam-nas, semeando

absurdos e confusão, que Locke dá o passo incomum de inserir um livro inteiro sobre palavras em um ensaio sobre a mente. Ele explica que as palavras se interpõem entre nossos pensamentos e a verdade (E III.ix.21).

Todavia, ainda que reconheça que a linguagem é imperfeita, era absolutamente necessário para Locke esclarecer o modo como se dá a comunicação, isto é, os meios pelos quais o uso correto da linguagem se efetiva. Afinal, se Locke fosse totalmente cético à possibilidade de comunicação, a publicação de seu mesmo Ensaio seria apenas um capricho sem sentido.

Se o que garante a verdadeira linguagem é a expressão de ideias inicialmente privadas, como dito acima, é preciso garantir àquele que comunica o poder de publicar as próprias ideias aos outros e a estes a capacidade de entendê-las. Por isso Locke levanta quatro razões pelas quais se pode afirmar que o ouvinte acessa as ideias do falante.

Em primeiro lugar, a comunicação é possível devido à similaridade fisiológica que torna a percepção do mundo e as operações mentais também similares entre as pessoas (E II.xxiii.12). Assim, mesmo que possam variar, as ideias de diferentes indivíduos sobre um mesmo objeto tendem, por assim dizer, a uma média. Mesmo que as diferenças existam, há um espaço de possibilidades finito que garante características em comum.

Em segundo lugar, os homens compartilham ideias de coisas que não experimentam diretamente, abstraindo ideias particulares por sua semelhança (E III.iii.7), por isso é perfeitamente possível que alguém que nunca viu uma toranja possa classificá-la corretamente como uma fruta cítrica, pois abstraiu de sua experiência com outros vegetais as ideias de fruta e cítrico (mais gerais, abstratas), ou até a capacidade de reconhecer um limão inteiramente novo como sendo mesmo um limão, pois abstraiu de sua experiência com limões as ideias que eles geralmente causam.

Em terceiro lugar, a posse de ideias abstratas compartilhadas permite aos homens concordar sobre quais ideias devem ser atreladas a quais palavras através de definição ostensiva: apontamos para um objeto e o nomeamos (E III.i.5; E III.ix.9; E III.xi.18).

Por último, identificamos e nomeamos ideias sem causas públicas evidentes porque no princípio do processo de formação das palavras tomamos nomes de coisas sensíveis para nomear ideias que não captamos pelos sentidos,

mas que se assemelham a elas, como, por exemplo, a palavra “anjo” cuja origem é “mensageiro”. (E III.i.5)

Assim, afirma Dawson (2010, 176-177), explicando a acessibilidade das ideias advindas diretamente da sensação e da reflexão ativa do entendimento sobre os dados dos sentidos, blocos constituintes de qualquer ideia concebível, Locke pode afirmar ter dado conta de toda a comunicação.

2.3.1 Do uso dos termos gerais

No capítulo terceiro do livro três, Locke se pergunta acerca da relação entre as coisas particulares e os termos gerais. Ele escreve (E III.iii.1-20): tudo quanto existe é particular e as palavras devem se conformar às coisas, mas a maioria dos termos é geral. Como isso é possível?

Em primeiro lugar, dada a enorme quantidade de particulares, não há como cada particular ter um nome distinto, e mesmo que fosse possível, a linguagem seria inútil, pois só haveria comunicação se todas as partes estivessem familiarizadas com os mesmos nomes. É a generalidade das palavras que permite aos homens comunicar-se mesmo que não compartilhem a mesma experiência do mundo.

As palavras gerais são os signos das ideias gerais. E as ideias se tornam gerais separando delas as circunstâncias de tempo e lugar e outras características que as possam determinar para uma existência particular. Assim, por abstração (*abstrahere*, do latim, significa *separar* uma parte de um todo), uma ideia se torna capaz de representar mais de um indivíduo que se conforma a ela.

Isto acontece, por exemplo, quando separando tudo quanto há de diverso nas ideias de homem, cão, peixe etc., chegamos a formar a ideia geral de animal. Por isso, afirma Locke (E III.iii.11), aquilo que é geral ou universal é apenas uma criação do entendimento que só se refere a signos, quer sejam palavras ou ideias, pois os universais não têm existência própria fora do entendimento.

As ideias abstratas funcionam como as essências dos gêneros e da espécie, pois ter a essência de uma espécie é conformar-se com a ideia abstrata a qual o nome da espécie está anexado. O entendimento capta as similaridades das coisas e forma ideias gerais abstratas que classifica sob nomes. Cada ideia abstrata distinta é uma essência distinta. Por isso, diz Locke, não é necessário classificar os

entes sempre em gênero e espécie, como fazia Aristóteles. Embora seja uma forma válida de classificação, a simples enumeração de características e propriedades já é suficiente para identificar corretamente as coisas.

A diferença para o conceito aristotélico de essência é que, para Locke existe uma essência real, que faz parte da constituição interna das coisas, e uma essência nominal, que é a ideia geral abstrata com um nome anexado. Há uma conexão constante entre o nome e a essência nominal: o nome de uma classe de coisas só pode ser atribuído aos seres particulares que têm a essência da classe, ou seja, a ideia abstrata da qual o nome é signo.

A espécie se distingue antes pela essência nominal que pela real, pois a essência real nem sempre é conhecível e usá-la para essa distinção só cria problemas, deixando de fora indivíduos que evidentemente deveriam estar incluídos, como os indivíduos defeituosos, por exemplo. Como a essência real diz respeito à constituição interna das coisas, caso um homem nasça sem uma perna, não pode ser considerado um indivíduo de sua espécie. Mas, se considerado mediante a essência nominal, mais flexível e abstrata, pode ser classificado corretamente. Embora as essências real e nominal coincidam nas ideias simples,⁷ sempre diferem nas ideias complexas de substância, que é o caso do homem.

As essências nominais são também incriáveis e incorruptíveis, pois sendo ideias abstratas consistentes, sem contradições internas em sua estrutura, podem vir a ser conhecidas mesmo que não exista nenhuma instância particular sua, por isso, a essência da sereia ou do unicórnio é tão conhecível quanto a do homem. Por isso, com a elaboração desta teoria sobre o uso de termos gerais, Locke pode desvencilhar-se de vez do essencialismo escolástico e fundamentar sua filosofia no nominalismo.

2.4 De como a ideia representa o objeto

Lennon (2007, p. 247-255), aponta quatro possibilidades para o modo como a ideia simples representa o objeto. São elas: semelhança, causação, referência e convenção. Vejamos a seguir como ele avalia cada uma delas.

⁷ A essência nominal coincide com a essência real também nas ideias de modos, mas não abordaremos aqui essa classe de ideias.

A representação por semelhança é logo descartada, pois, como se vê adiante, um dos motivos pelos quais Locke distingue entre ideias de qualidades primárias e ideias de qualidades secundárias é justamente que estas não podem ter qualquer semelhança com os corpos, isto é, com os objetos extramentais:

[...] as ideias das qualidades primárias dos corpos são semelhanças deles, e seus padrões realmente existem nos corpos mesmos; mas as ideias produzidas em nós por estas qualidades secundárias não têm nenhuma semelhança deles. Não há nada como nossas ideias [de qualidades secundárias] existindo nos corpos mesmos. Elas são, nos corpos, [...] somente um poder de produzir aquelas sensações em nós: e aquilo que é doce, azul, ou quente em ideia é apenas certo volume, figura e movimento das partes insensíveis nos próprios corpos a que chamamos assim (E II.viii.15, tradução nossa).⁸

Assim, se as ideias de qualidades secundárias não são semelhanças, é absurdo supor que Locke tenha em mente uma teoria da representação por semelhança.

Já a representação por causalção se baseia na tese de que só a relação causal constitui a representação, assim, diferentes objetos causam diferentes aparências. Em seus escritos, Locke parece aderir a esta explicação, mas a causalção não é nem necessária nem suficiente para explicar a representação: o calor, por exemplo, causa a ebulição da água, mas a ebulição da água não representa o calor. Sequer é possível explicar a representatividade da ideia estabelecendo “circunstâncias normais” para sua causa, pois a especificação de tais “circunstâncias normais” requer a mesma noção de representação que se pretende explicar.

Por isso alguns comentadores suplementam essa explicação causal com a noção de referência. Porém, a representacionalidade da ideia não pode se reduzir a sua referencialidade, mesmo associada à ideia de causalção, pois Locke diz que é a mente que referencia (E II.xxxi.1), geralmente referenciando ideias a arquétipos dos quais supõe que tais ideias sejam tiradas e com os quais pretende que tais ideias se conformem, de modo que uma ideia é adequada ou inadequada a um objeto se se perfaz ou se falha a conformidade pretendida.

8 “[...] *the ideas of primary qualities of bodies, are resemblances of them, and their patterns do really exist in the bodies themselves; but the ideas, produced in us by these secondary qualities, have no resemblance of them at all. There is nothing like our ideas, existing in the bodies themselves. They are in the bodies [...] only a power to produce those sensations in us: and what is sweet, blue, or warm in idea, is but the certain bulk, figure, and motion of the insensible parts in the bodies themselves, which we call so.*”

Para que uma ideia simples referencie um objeto é preciso um dado adicional que acompanha a percepção do objeto: ou que a ideia seja codificada com alguma informação sobre o objeto ou que haja uma percepção adicional referenciada ao objeto. Como no caso da memória, que, para Locke, é uma percepção adicional que acompanha a percepção atual. Ambas as teses permitem ideias como signos naturais, mas essa informação ou percepção adicional é justamente aquilo que a ideia deveria prover.

Se a ideia precisa de uma percepção adicional para representar um objeto, sempre que a percepção adicional não seja evidente por sua própria estrutura interna, como os objetos ideais da matemática, a própria percepção adicional necessitará de mais outra percepção adicional para ser representada. No caso, para interpretar tal informação ou percepção adicional, a mente precisaria de outra informação ou percepção adicional que provesse a mesma informação ou percepção que se busca. Imaginemos uma bola de neve. Se para interpretar essa percepção fosse necessário que S possuísse uma percepção adicional, de esfera, por exemplo, o encadeamento de percepções acabaria, pois o conceito de esfera, não depende de nada físico, mas apenas de conceitos abstratos. Mas, caso S perceba algo mais complexo, e cujas percepções adicionais necessárias sejam também complexas, pode ser que S necessite de uma árvore de percepções adicionais muito extensa, o que parece não condizer com a velocidade e prontidão de nossas percepções. Este problema só tem solução se a significação natural for abandonada em favor da significação por convenção.

Assim, uma “interpretação definitiva” sobre as ideias simples pode ser obtida a partir daquilo que Locke diz sobre a linguagem usada por elas. Ele parece sugerir que as ideias simples significam, ou representam, por convenção. Ele diz que a referência do termo ‘branco’ é fixada quando o tomamos para nomear a cor que observamos no leite ou na neve. Para fixar a referência de ‘branco’, no entanto, não é preciso saber nada sobre aquilo que causa a cor. O termo ‘branco’ se refere a qualquer ideia de cor que alguém tenha enquanto vê o leite ou a neve, assim como a ideia representada pelo termo ‘branco’ pode ser referenciada a qualquer coisa que está causando aquela ideia de cor. “Enquanto escrevo isto, tenho, pelo papel que afeta meus olhos, aquela ideia produzida em minha mente que, seja qual for o objeto que a cause, chamo branco” (E IV.xi.2).

Isso não significa que precisamos ter previamente a ideia de cor em geral, isto é, algo como “cor é o objeto da visão” ou “cor é uma qualidade atribuída pela mente aos dados da visão que varia em função da frequência de onda da luz”. A ideia de cor, nesse caso, é uma ideia abstrata que inclui várias ideias simples: vermelho, amarelo, branco, etc. Apenas é necessário que a ideia de uma certa cor, ou seja, uma ideia simples, se repita em diferentes objetos de modo que eu possa me referir a certa qualidade em minha mente que aparece sempre quando vejo o leite, a neve, ou algo indeterminado que, distante de mim, não posso determinar se é um saco, um gato, uma caixa, etc. Mas como se repete essa mesma ideia (simples) de cor, posso dar-lhe um nome, seja “white”, “branco”, “Weiß”, ou “abc123”.

Portanto, alguém que enxergasse em espectro invertido poderia referir-se aos mesmos objetos que alguém normal, na medida em que seus espectros sejam estáveis, ou seja, na medida em que a representação covarie com o representado, pois as ideias simples causadas pelas qualidades dos objetos não precisam ser fenomenologicamente idênticas nem entre as pessoas e nem com os próprios objetos extramentais.

As palavras também não são signos naturais. Como escreve Losonsky (2007, p. 288), Locke rejeita a possibilidade de uma linguagem não convencional da natureza, pois considera que a relação de significação entre palavras e ideias é um artefato, produto da atividade voluntária, pois os sons não têm uma conexão natural com nossas ideias:

[...] as palavras [...] passam a ser utilizadas pelos homens como signos de suas ideias; não por nenhuma conexão natural, que haja entre sons articulados particulares e certas ideias, pois então haveria apenas uma linguagem entre todos os homens; mas por uma imposição voluntária, pela qual tal palavra é feita arbitrariamente a marca de tal ideia (E III.ii.1, tradução nossa).⁹

Todavia, ressalta ainda Losonsky (2007, p. 297-301), o conceito de significação do Ensaio não pode ser definido só em termos de indicação convencional, pois se a significação fosse mera indicação convencional, Locke precisaria limitá-la só aos casos em que as palavras são usadas como signos instrumentais, signos que não são eles mesmos um meio de pensar, mas indicam

9 “[...] words [...] come to be made use of by men, as the signs of their ideas; not by any natural connexion, that there is between particular articulate sounds and certain ideas, for then there would be but one language amongst all men; but by a voluntary imposition, whereby such a word is made arbitrarily the mark of such an idea.”

outro objeto. No entanto, ele considera que os nomes às vezes são signos formais, ou seja, são usados para pensar sobre as coisas. “Não é suficiente para a perfeição da linguagem”, afirma, “que os sons possam ser signos de ideias, a menos que esses signos possam ser usados de tal modo que compreendam várias coisas particulares” (E III.i.3, tradução nossa).

Explique-se: a anexação de nomes é especialmente importante para as ideias de “modos mistos”, que são as ideias gerais que combinam arbitrariamente ideias simples que não parecem ocorrer juntas na natureza; é o caso das ideias morais, por exemplo. Os nomes funcionariam como um meio de união e permanência para essas coleções díspares que são os modos mistos. Uma ideia como “justiça” pode significar diversos atos humanos, inclusive atos contrários entre si. Relembrando o célebre exemplo de Platão em *A República*: se guardo uma arma de alguém, é justo que eu a devolva ao proprietário, mas, estando enfurecido esse mesmo proprietário e querendo reaver sua arma para assassinar outrem, não é justo que eu a devolva. Nesse caso, devolver ou se negar a devolver são ideias opostas, mas que, a depender da circunstância, podem ambas representar atos de justiça.

Os nomes também são necessários para as ideias de substância, mas como meio de estabilização, porque os arquétipos, as ideias simples que ocorrem juntas em certos padrões, são muito numerosos e a ideia geral resultante pode divergir de pessoa para pessoa dependendo da qualidade da interação do sujeito com esses padrões, ou seja, dependendo da habilidade, atenção e circunstâncias.

Por isso Locke vincula a classificação à linguagem, pois considera que as ideias que usamos para classificação são instáveis, seja porque a classificação não se baseia em nenhum padrão natural ou arquétipo, seja porque os padrões dados são muito complexos e fluidos. Por isso recorre à estabilidade dos signos externos, especialmente às palavras, para manter as ideias combinadas em ideias gerais. Daí a conclusão supracitada de que “há uma conexão tão estreita entre ideias e palavras”.

Em suma, a teoria de Locke diz que: (1) as ideias são causadas pelas qualidades dos objetos; (2) as ideias simples são representações que covariam com aquilo que representam; (3) as ideias gerais abstratas são elas mesmas as classes de coisas; (4) as palavras são representações convencionais das ideias; (5) as palavras aglutinam e estabilizam as ideias abstratas.

2.5 Como Locke pode responder a Cummins?

Tudo o que LOCKE pode captar através de seu sistema perceptual são ideias simples: cor, movimento, etc. Ao formar um percepto, LOCKE agrupa as ideias simples percebidas numa ideia complexa, o percepto, que é então comparada com as ideias gerais abstratas, que são os cartões de conceito que estão armazenados. O nome escrito no verso do cartão de conceito é o signo da classe de coisas que aquela ideia geral abstrata representa.¹⁰

Assim, diante de Tom, o que LOCKE produz é a ideia complexa que agrupa as características de Tom e então busca a ideia abstrata a qual esta ideia complexa de Tom melhor se encaixa. Como Tom é um gato, o esperado é que a ideia complexa de Tom corresponda à ideia geral abstrata de gato.

Mas LOCKE não pode buscar entre as ideias abstratas o mesmo padrão de furos da ideia complexa formada na primeira percepção de Tom. A ideia abstrata tem necessariamente menos ideias simples em seu conjunto, ou seja, o cartão de conceito tem necessariamente menos furos que o cartão de percepção. Nesse sentido, o conteúdo do cartão de conceito não pode ser idêntico, mas deve estar contido no cartão de percepção. LOCKE precisa encontrar entre os cartões de conceito padrões de furos que se encaixem no cartão de percepção formado quando ele foi confrontado por Tom. Portanto, mais de um cartão de conceito irá se encaixar: gato, animal, vivente, ente, etc.

Quando LOCKE é confrontado por Tom ele não pode formar um cartão de percepção de cão, pois a ideia complexa que acaba de ser formada não é nem ela mesma uma classe e nem possui uma classe. Lembremos que a essência da classe

10 É possível imaginarmos classes confusas, como se o objeto caísse ao mesmo tempo sob duas ou mais classes de mesmo nível hierárquico como, por exemplo, um pato-coelho ou um garfo-colher. Quanto a esse problema há que separar entre dois tipos de objeto: os naturais e os artificiais. Os objetos artificiais se baseiam na arte e conceitos humanos e são mais flexíveis em sua constituição formal. Uma colher com dentes pode facilmente cumprir as funções de garfo e colher, assim como uma mesa baixa pode servir para sentar-se confortavelmente ou uma cadeira alta pode ser usada como uma mesa improvisada, mas ergonomicamente confortável. Quanto aos objetos naturais existe uma dificuldade adicional em se assumir a possibilidade de classes confusas. O que é um pato-coelho? Podemos muito bem reunir numa lista os atributos dos patos e dos coelhos, mas a ordem em que esses conceitos se formam é inversa àquela dos objetos artificiais. Os homens conceberam os objetos artificiais para atender a certas finalidades práticas. Eles surgem na mente e são construídos no mundo. Os objetos naturais surgem no mundo e só então são conhecidos pela mente. Podemos presumir que é impossível classificar um objeto natural como um pato-coelho porque patos-coelhos não existem. Porém, caso pato-coelho se refira a qualquer representação artificial dessa ideia, como uma escultura, um desenho, etc, então, sim, poderemos classificá-la de modo confuso.

é a ideia geral abstrata, e os cartões de percepção são ideias complexas formadas na primeira percepção de um objeto real. Só após a comparação do cartão de percepção com os cartões de conceito LOCKE poderá emitir um juízo: isto é um cão, ou isto é um gato.

Este primeiro passo apenas reúne o que os sentidos captaram: nesta etapa não há valor de verdade nem proposições. Ele equivoca-se ao supor que LOCKE seja confrontado por Tom e produza um percepto de cão (padrão de furos D). A representação linguística mais adequada para esta primeira percepção seria uma oração imperfeita, isto é, sem verbo de cópula. Ao ser confrontado por Tom, LOCKE até poderia formar algo que em nossa percepção corresponda a “isto: um cão”, mas nunca a “isto é um cão” ou “isto não é um cão”, pois a afirmação embutida no verbo de cópula, que no caso de LOCKE corresponde à exibição do verso do cartão e conceito, só aparece após a comparação entre os conceitos “em memória” e que aquilo que foi percebido.

Um “percepto de cão” não faz sentido nesse primeiro momento porque cão é classe e tudo o que LOCKE fez até então foi reunir num cartão os dados captados. Cummins pulou da etapa de percepção para a dos conceitos. Se LOCKE formasse o padrão de furos da classe cão a partir de um único exemplo, Tom, LOCKE cometeria uma indução fraca, pois qualquer coisa real é mais complexa que a ideia abstrata de cão e o único modo de formar esta tal ideia abstrata é depurando a memória de uma vasta coleção de exemplares.

A descrição de LOCKE não funciona porque ela não descreve completamente o processo de percepção dos seres humanos. Ao elaborar essa descrição, Cummins acabou simplificando uma parte importante do processo, e por isso penso que suas conclusões não podem ser extrapoladas de LOCKE para os seres humanos. O cartão de percepção não pode ser um cartão com o padrão de furos de cão, de gato, etc. O cartão de percepção tem o padrão de furos da coisa individual, única no universo, que só após uma conceituação do entendimento sobre ela pode ser compreendida como, por exemplo, um gato chamado Tom.

Há muito mais no processo de percepção e reconhecimento do que LOCKE é capaz de descrever. A memória dessa primeira percepção que captamos é extremamente volátil e frequentemente completamos as partes que somem com informação adicional que contamina o primeiro dado, ou simplesmente esquecemos partes suas. Para notar isto basta tentar repetir uma sequência de números que só

observamos por alguns segundos. Além disso, há as questões como memória, experiência e formação individual, que podem gerar ideias abstratas falsas ou inconsistentes, como vimos acima ao tratar da influência da linguagem e de seus abusos sobre a formação das ideias. E uma percepção malconceituada será sempre, de certa forma, uma deturpação daquilo a que ela corresponde no mundo extramental.

3 O COVARIACIONISMO DE FODOR

Fodor, assim como Locke, defende uma explicação sobre a natureza da representação mental baseada na covariação causal entre representante e representado. Ele assume a Hipótese da Linguagem do Pensamento, a hipótese de que as representações mentais são símbolos linguísticos, e a Teoria Representacional da Mente (doravante TRM), que sustenta que os símbolos da linguagem do pensamento, constituintes das representações mentais, funcionam como: (1) objeto das atitudes proposicionais, ou seja, em “crer que P” ou “desejar que P”, P é um símbolo da linguagem do pensamento; e (2) domínio dos processos mentais, ou seja, o encadeamento causal entre estados mentais corresponde às relações entre símbolos da linguagem do pensamento.¹¹

De acordo com Hermida Lazcano (1993, p. 360), esse conjunto de teses acarreta uma interessante aproximação entre a filosofia da mente e a filosofia da linguagem: ao identificar representações mentais com símbolos linguísticos, Fodor estabelece uma analogia entre os paralelismos ‘intencionalidade e representação’ e ‘significado e símbolo’.

A defesa feita por Fodor da psicologia do senso comum sobre crenças e desejos, prossegue Hermida Lazcano (1993, p. 361), se desdobra numa reivindicação da validade das atitudes proposicionais através da discussão de exemplos intuitivos, testemunhos e experiências, que mostram as virtudes da psicologia do senso comum: sua credibilidade, sua profundidade e sua indispensabilidade.

A fim de sustentar cientificamente essa psicologia do senso comum, isto é, a fim de estabelecer um caminho pelo qual essa psicologia popular possa ser explicada em termos de uma ciência cognitiva rigorosa, Fodor elabora a Teoria Causal Bruta, que descreve o processo de representação através de uma relação semântica entre símbolo e objeto garantida pela conexão causal correspondente. Sinteticamente: se o símbolo ‘A’ é uma representação mental de um sujeito S, então

¹¹ É importante destacar que a TRM enfrenta uma dificuldade em relação a sua necessidade de apelar a símbolos linguísticos, pois, embora a tese de que “representações mentais são símbolos linguísticos” seja plausível para as atitudes proposicionais, ela não o é para a percepção. Os símbolos de uma linguagem só ganham significado quando relacionados a um conjunto de regras gramaticais, e não é nada óbvio que as primeiras percepções táteis de um feto no útero materno, por exemplo, sejam símbolos desse tipo, afinal, como explicar que uma entidade biológica com sistema nervoso imaturo e possivelmente incompleto é capaz de relacionar símbolos de acordo com um conjunto de regras?

existe um A que causa 'A' em S e 'A' denota A para S.

Na primeira seção trataremos da psicologia do senso comum; na segunda, abordaremos a TRM; na terceira seção iremos expor a Teoria Causal Bruta, e em seguida, na quarta seção, trataremos das críticas de Cummins às teses de Fodor. Por fim, na quinta seção, veremos o modo como Fodor poderia rebater essas críticas e minha própria opinião sobre o tema.

3.1 A psicologia do senso comum

Fodor (1987, p. 1-10) defende a *psicologia do senso comum sobre crenças e desejos*, pois acredita que ela explica mais sobre o comportamento que qualquer outra teoria. Tal psicologia se refere ao senso comum sobre a mente e o comportamento humano, sendo uma forma intuitiva de entender e explicar as ações, as emoções e os estados mentais das pessoas no dia a dia.

A psicologia do senso comum envolve o uso de conceitos como crenças, desejos, intenções, percepções e emoções para explicar e prever o comportamento humano, como, por exemplo, quando acreditamos que, porque alguém está com raiva, então essa pessoa tem o desejo de punir ou se vingar de outrem.

Segundo Fodor, a psicologia popular é uma forma primitiva de teoria da mente que as pessoas utilizam para entender o mundo social ao seu redor. Muito embora reconheça que essa abordagem não é uma teoria científica completa, tendo a necessidade de ser complementada por uma psicologia científica mais rigorosa e formalizada, que recorra a conceitos e teorias mais precisos e específicos, ele defende a psicologia do senso comum e dá três razões para sua preferência por ela: ela funciona na maioria dos casos; ela é uma teoria profunda; e ela é indispensável. Vejamos como ele explica cada uma dessas razões.

Em primeiro lugar, Fodor argumenta que a psicologia do senso comum *funciona corretamente na maioria dos casos* em que precisamos inferir as intenções das pessoas a partir de suas declarações, tanto de conhecidos como de estranhos e em situações reais. Mesmo considerando a imprecisão dos enunciados que comumente usamos, como uma promessa, por exemplo, cremos que ela é verdadeira a menos que o falante esteja mentindo, dizendo palavras ininteligíveis, falando enquanto dorme, etc.

Implicitamente consideramos que as promessas incorporam cláusulas

ceteris paribus, isto é, cláusulas que declaram que dada relação se mantém em condições normais, o que equivale a algo como “eu farei tal coisa, a não ser que eu não faça (por isso ou por aquilo, por este ou aquele motivo)”.

A simples existência de exceções não depõe de modo algum contra essa psicologia popular, uma vez que é uma circunstância absolutamente normal em todas as ciências, que se veem obrigadas a recorrer a outras ciências para explicar as exceções às suas generalizações (HERMIDA LAZCANO, 1993, p. 361). Por exemplo, “um rio com meandros erode suas margens a menos que o clima mude e o rio seque” é uma generalização possível na geologia, porém, “clima” não é um termo da geologia (FODOR, 1987, p. 5-6).

Precisamos admitir que a maioria das generalizações da psicologia do senso comum funcionam, pois nos permitem fazer previsões com boa margem de acerto, como “estarei em Fortaleza às quinze horas do dia três de janeiro” ou “hoje à noite meus amigos virão me visitar”, ou ainda, “tenho uma reunião marcada para a tarde da próxima quinta-feira”, etc.

Esse tipo de previsão é trivial e cotidiana, e sabemos por experiência (pelo nosso uso corriqueiro da linguagem) que são extremamente úteis, pois funcionam na maioria das vezes em que precisamos estabelecer compromissos futuros e fazer promessas, e está implícito em seu uso, tanto para o emissor quanto para o receptor da mensagem, que se digo sinceramente que “estarei em Fortaleza às quinze horas do dia três de janeiro”, estarei lá a menos que algo me impeça.

Em segundo lugar, Fodor argumenta que a psicologia do senso comum é uma teoria *profunda*, pois conta com a mesma estrutura dedutiva característica da ciência real, que opera sobre a inobservabilidade de boa parte de suas relações causais. As generalizações da psicologia do senso comum são definidas através de inobserváveis porque nós inferimos o que alguém tem em mente através de seu comportamento, ou seja, as representações mentais, privadas, são os inobserváveis da psicologia do senso comum.

Mas, enquanto a meteorologia do senso comum, por exemplo, do tipo “o céu está limpo esta noite, logo não choverá pela manhã”, envolve apenas instanciação e *modus ponens*, e por isso é ineficaz e superficial, a psicologia do senso comum, ao contrário, passa no teste de profundidade assumindo que dado comportamento ocorre após uma cadeia causal arbitrariamente longa e emaranhada de eventos mentais não observáveis. Na psicologia do senso comum essa

inobservabilidade se traduz na inescrutabilidade das cadeias causais que conectam os estados mentais entre si.

Nesse sentido, há uma grande diferença entre “o céu está limpo esta noite, logo não choverá pela manhã” e “dada sua cara de tédio, João vai embora antes de acabar a aula”, pois, enquanto a primeira previsão geralmente não conta com explicações muito elaboradas sobre a relação entre noites de céu nublado e manhãs chuvosas, e a própria natureza caótica do clima dificulta qualquer previsão a longo prazo, a cara de tédio de João evoca uma série de explicações sobre os estados internos de João que permitem prever seu comportamento com boa margem de acerto: sabemos que o movimento dos músculos da face dos seres humanos expressa com fidelidade seus estados emocionais (na maioria dos casos), e que nossos estados emocionais quase sempre influem em nosso comportamento.

Uma vez que a “cara de tédio” (que é certa configuração da tensão muscular facial) expressa o estado interno correspondente ao sentimento de tédio, e que esta emoção está ligada a uma baixa resistência à frustração, à dispersão da atenção e ao impulso de fazer cessar a situação de descontentamento, é muito provável que João, adulto, cansado e com um número de faltas abaixo do limite estabelecido para reprovação, vá embora antes que a aula acabe.

Também se pode afirmar que a psicologia do senso comum é profunda porque as suas generalizações levam a previsões não por instanciação direta, como no caso da meteorologia do senso comum, mas por iteração e interação. Analisemos o seguinte caso: Hérnia, personagem de Shakespeare em *Sonho de Uma Noite de Verão*, conta com uma teoria implícita da interação de crenças, preferências e comportamentos, uma teoria de decisão implícita. Ela acredita que o comportamento de Demétrio é intrincadamente causado por seus estados mentais. Ela não acredita que os rivais no amor necessariamente matam (para todos x e y , se x e y são rivais no amor então x e y matam), antes, ela produz uma generalização mais abrangente que suporta contrafactuais: os rivais no amor preferem a derrota mútua (para todos x e y , se x e y são rivais no amor então x e y preferem a derrota mútua).

Tal princípio leva a previsões comportamentais, mas apenas através de muitas suposições adicionais sobre como as preferências das pessoas podem afetar as ações delas em certas situações. Deste modo, a dedução de Hérnia de que Demétrio possivelmente matou Lisandro, não é consequência apenas do fato de que ambos a amam, mas também de muitas outras condições e causas que interferem

em sua conclusão.

Hércia acredita que é amada por Lisandro. Ela acredita que se Lisandro a ama, então ele quer o seu bem. Portanto, ele não a abandonaria voluntariamente à noite em uma floresta escura. Mas Hércia foi abandonada, logo, não voluntariamente. Portanto, involuntariamente. Logo, é plausível que Lisandro tenha sofrido algum mal. Mas quem o faria mal? Plausivelmente, Demétrio, pois é seu rival, e rivais apaixonados não desejam o bem um ao outro, assim, Hércia deduz que Demétrio acredita que Lisandro vivo é um impedimento para cortejá-la com sucesso.

Em terceiro lugar, a psicologia do senso comum é uma teoria de generalizações com suporte contrafactual *indispensável* para explicar nosso comportamento e suas causas, pois as suas generalizações perfazem a conexão entre as trocas de elocuições e o comportamento subsequente sem a qual não é possível descrever enunciações como palavras ou comportamentos como atos.

De acordo com Hermida Lazcano (1993, p. 361), a argumentação de Fodor trata de abrir espaço para acomodar as atitudes proposicionais dentro do repertório ontológico da ciência cognitiva, de modo que se torna imprescindível perguntar pela natureza dessas atitudes, cujas propriedades essenciais são, para Fodor, *avaliabilidade semântica*, *potencialidade causal* e *generalidade* (com preservação de seu valor semântico).

A avaliabilidade semântica diz respeito à relação entre as atitudes proposicionais e o mundo, ou seja, às condições de satisfação de determinada atitude. O que faz uma crença verdadeira ou falsa, ou um desejo satisfeito ou insatisfeito, é a sua relação com o não psicológico. Afirmar que uma crença é verdadeira ou falsa é avaliá-la em termos de sua relação com o mundo e o mesmo vale para outros tipos de atitudes, como desejos, etc.

As ideias de 'conteúdo do estado psicológico' e 'avaliabilidade semântica' estão intimamente ligadas. Retomando o exemplo da atitude de crença, podemos dizer que, quando se conhece o conteúdo de uma crença, então se sabe aquilo no mundo a que essa crença se refere, o que determina a avaliabilidade semântica dessa mesma crença.

Assim, se Yoda crê que Annakin é perigoso e não deve ser treinado nas artes jedi, a crença de Yoda tem um valor semântico: é uma crença verdadeira, pois corresponde a um fato. É justamente o fato de que Annakin é perigoso que

determina a avaliabilidade semântica da crença de Yoda, não importando quaisquer outros fatos, pois a crença de Yoda expressa a proposição “Annakin é perigoso”.

Em segundo lugar, pela ótica da psicologia do senso comum, existe uma coextensionalidade entre a classe de eventos mentais semanticamente avaliáveis, composta pelos estados mentais dotados de conteúdo, e a classe dos eventos mentais causalmente potentes, composta pelos estados mentais capazes de produzir comportamento e outros estados mentais. Em outras palavras, é característico da psicologia do senso comum atribuir conteúdos e poderes causais às mesmas entidades mentais, o que implica que um evento mental só tem eficácia causal se tiver um conteúdo. É essa conexão entre conteúdo e efetividade que garante a potencialidade causal das atitudes proposicionais (Fodor, 1987. p. 12).

Mais ainda, prossegue Hermida Lazcano (1993, p. 362), as relações causais entre atitudes proposicionais conseguem respeitar suas relações de conteúdo. Esse paralelismo fundamenta uma semelhança entre argumentos e cadeias causais de pensamento, uma harmonia que torna o pensamento um exercício proveitoso e sensato. Por mais surpreendente e misteriosa que seja essa harmonia, nunca será inteligente renunciar à noção de conteúdo em resposta às múltiplas questões que nos são colocadas:

Que tipo de mecanismo poderia ter estados que são semanticamente e causalmente conectados, e cujas conexões causais respeitam as semânticas? [...] Se o paralelismo entre conteúdo e relações causais é, como parece ser, um atributo essencial da mente cognitiva, então, a menos que possamos salvar a noção de conteúdo, há um atributo essencial da mente cognitiva que a psicologia não dará conta (FODOR, 1987, p. 14, tradução nossa).¹²

Finalmente, a terceira propriedade essencial das atitudes proposicionais, a generalidade, relaciona-se à razoabilidade das generalizações e das explicações comportamentais cotidianas, que, como discutido anteriormente, funcionam na maior parte das vezes e nos permitem fazer previsões (HERMIDA LAZCANO, 1993, 362).

Essa argumentação a favor da psicologia do senso comum sobre crenças e desejos, contudo, precisa ser iluminada por uma teoria razoável e ontologicamente

12 *“What sort of mechanism could have states that are both semantically and causally connected, and such that the causal connections respect the semantic ones? It is the intractability of such questions that causes many philosophers to despair of common-sense psychology. But, of course, the argument cuts both ways: if the parallelism between content and causal relations is, as it seems to be, a deep fact about the cognitive mind, then unless we can save the notion of content, there is a deep fact about the cognitive mind that our psychology is going to miss.”*

comprometida com a defesa das atitudes proposicionais exposta anteriormente, que para Fodor, é a TRM. Vejamos a seguir como ele descreve essa teoria.

3.2 A Teoria Representacional da Mente

Conforme supracitado, o principal objetivo de Fodor (1987, p. 6) é argumentar a favor de uma psicologia científica da intencionalidade que conserve as explicações da psicologia do senso comum, uma vez que estas explicações sejam desdobradas em termos de uma ciência mais básica, como a neurociência, por exemplo. Para ele, isso pode ser conseguido através da *Teoria Representacional da Mente* (TRM), cuja tese fundamental é a existência de uma *linguagem do pensamento*, que é um conjunto infinito de representações mentais que funcionam tanto como objetos imediatos das atitudes proposicionais quanto como domínios dos processos mentais.

O termo “atitude proposicional” denota certos eventos psicológicos que correspondem a enunciados de crenças ou desejos de um sujeito S em relação a uma proposição P. Nesse sentido, quando “S crê que P” ou “S deseja que P”, S toma uma atitude (acreditar ou desejar) em relação a um objeto semanticamente avaliável (proposição P) e valorável (P pode ser verdadeira ou falsa).¹³

Fodor (1987, p. xi) acredita que os símbolos são candidatos óbvios para a representação, pois é próprio do símbolo possuir um conteúdo representacional: todo símbolo é sobre algo diferente de si, e, por definição, seria absurdo supor o contrário. Nesse sentido, os estados mentais podem ser explicados em termos simbólicos, assim, desejar ou crer que P seria o mesmo que produzir um símbolo discursivo que é ocorrência (*token*) de uma expressão da linguagem do pensamento.

Por isso, a TRM afirma que ter uma atitude proposicional em relação a uma proposição P é ter um símbolo ou representação mental que significa que P e que cumpre um determinado papel funcional como, por exemplo, ser uma crença, um desejo, etc. Fodor (1987, p. 17), define duas reivindicações desta teoria sobre a natureza das atitudes proposicionais:

¹³ Existem outros tipos de atitudes proposicionais além das crenças e desejos, mas isso não é relevante para a argumentação empregada aqui. Para mais referências, Cf. BRANQUINHO *et al*, 2006, p. 72-78.

Reivindicação 1 — *para todo organismo O e para toda atitude A em relação à proposição P, existe uma relação R (computacional/funcional) e uma representação mental R_M tal que R_M significa que P, e O tem A se, e somente se, O mantém a relação R com R_M .*

Isso significa que, para um sujeito S, se S crê que P, então S possui um símbolo (ou representação mental) de P que cumpre o papel funcional de ser uma crença. Ou seja, se Pedro crê que “Maria é simpática”, então Pedro tem em sua mente a proposição “Maria é simpática” e ele acredita que tal proposição é o caso.

Reivindicação 2 — *os processos mentais são cadeias causais de ocorrências simbólicas de representações mentais.*¹⁴

A TRM sugere um mecanismo plausível para o paralelismo das relações causais entre estados mentais com as relações semânticas entre objetos proposicionais. Por isso é possível usar os computadores como uma analogia para os pensamentos, pois eles nos mostram como conectar propriedades semânticas e causais via sintaxe.

Porém, para que a metáfora computacional funcione, é necessário identificar alguns símbolos que trabalhem ao nível da mente e que sejam dotados tanto de propriedades sintáticas quanto semânticas. Para Fodor, esses símbolos são justamente as representações mentais em que se baseia o nosso “idioma interno” (FODOR, 1987, p. 18-19).

Portanto, é necessário explicar o significado mental: o que significa dizer que um termo primitivo do idioma interno possui um conteúdo? Para esse fim, Fodor elabora uma teoria semântica baseada na causalidade que ele chama de *Teoria Causal Bruta*. Vejamos a seguir em que consiste essa teoria.

3.3 A Teoria Causal Bruta

A intenção de Fodor ao elaborar essa teoria semântica é alcançar uma teoria “naturalizada do significado [...] que articula, em termos não semânticos e não intencionais, condições suficientes para que uma parte do mundo *seja sobre* [...] outra parte” (FODOR, 1987, p. 98, tradução nossa).¹⁵

¹⁴ “*Mental processes are causal sequences of tokenings of mental representations.*”

¹⁵ “*I want a naturalized theory of meaning; a theory that articulates, in nonsemantic and nonintentional term, sufficient conditions for one bit of the world to be about (to express,*

Sua primeira tentativa, a Teoria Causal Bruta, pressupõe uma remissão semântica entre o símbolo e a coisa com suas propriedades, remissão que é garantida por uma conexão causal estreita e férrea (HERMIDA LAZCANO, 1993, p. 367). Ela afirma, de modo geral, que ocorrências simbólicas do idioma interno denotam suas causas e que tipos simbólicos desse mesmo idioma interno expressam a propriedade de confiavelmente causar suas ocorrências quando instanciados.

Segundo essa teoria, um símbolo expressa uma propriedade se for nomologicamente necessário¹⁶ que *todas* as instâncias desta propriedade (e *somente elas*) causem ocorrências do símbolo. Porém, dois problemas logo aparecem: o *problema da sobredeterminação*, nos casos em que alguns *não-x* causam ocorrências simbólicas de *x*; e o *problema da subdeterminação*, nos casos em que nem todos os *x* causam ocorrências simbólicas de *x* (FODOR, 1987, p. 99-100).

Imaginemos um caso de sobredeterminação em que o gato Tom causa a ocorrência simbólica ‘cão’. Seria preciso assumir a imperfeição da dependência entre ‘cão’ e os cães, assim como seria preciso assumir que tanto cães quanto gatos são suficientes para causar ‘cão’. Uma vez que a Teoria Causal Bruta afirma que as ocorrências simbólicas denotam suas causas, segue-se que o que a ocorrência ‘cão’ pode expressar é uma disjunção: a propriedade de ser um “cão-ou-gato”, ou um “cão-ou-Tom”. Isso quer dizer que essa ocorrência simbólica não pode ser reconhecida como uma deturpação, um erro de representação, pois, pelos critérios da Teoria Causal Bruta, cães ou gatos são causas possíveis para ‘cão’ (Fodor, 1987, p. 101-102).

Segundo Fodor (1987, p. 106), o problema da disjunção pode ser evitado introduzindo uma diferença entre as ocorrências simbólicas ‘*x*’ que são causadas por um *x* e aquelas que são causadas por um *não-x*. Essa diferença precisa ser expressa em termos de propriedades não intencionais e não semânticas, já que tais

represent, or be true of) another bit.”

16 Necessidade nomológica tem o mesmo sentido que necessidade física. É um tipo de necessidade que abrange as ciências naturais, como a física, a química ou biologia. De acordo com Branquinho et al (2006, p. 537), P é uma necessidade física/nomológica se, e somente se, as leis da física implicam P. “Nenhum objeto viaja mais depressa do que a luz” é um exemplo de necessidade física. Um corolário dessa definição é que qualquer necessidade lógica é igualmente uma necessidade física, pois as verdades lógicas são “vacuamente” implicadas por qualquer outra proposição – e, portanto, são também implicadas pelas leis da física.

propriedades são tudo aquilo com o que a Teoria Causal Bruta precisa lidar.

Fodor (1987, p. 107), argumenta que as falsidades são ontologicamente dependentes das verdades, enquanto as verdades de modo algum dependem das falsidades. Só podemos ter crenças falsas acerca daquilo sobre o que se pode ter crenças verdadeiras, mas podemos ter crenças verdadeiras sobre qualquer coisa sobre a qual se possa ter uma crença.

Assim, prossegue Fodor (1987, p. 107-108), o fato de confundir o gato Tom com um cão não leva um sujeito S a afirmar “isto é um cão” a menos que haja uma relação semântica entre a palavra “cão” e os cães. Do fato de que a palavra “cão” designe os cães não decorre que S confunda Tom com um cão, mas decorre que, diante de um cão, S afirme “isto é um cão”, independentemente de qualquer relação semântica, ou de qualquer outro tipo de relação, entre a palavra “cão” e os gatos.

Saindo do exemplo semântico acima, podemos imaginar, diz Fodor (1987, p. 108-111), que a conexão causal entre gatos e a ocorrência simbólica ‘cão’ é, do mesmo modo, assimetricamente dependente da conexão causal entre ‘cão’ e os cães, o que estabelece uma condição necessária para que uma ocorrência simbólica ‘x’ causada por um *não-x* seja um caso anômalo: *para todo “A’ causado por B”, “A’ causado por B” é uma deturpação representacional somente se “A’ causado por B” é assimetricamente dependente de “A’ causado por não-B”*. De modo correspondente, *“A’ causado por B” expressa a propriedade disjuntiva A ou B somente se “A’ causado por B” é simetricamente dependente de ‘A’ causado por A*. Assim, a Teoria Causal Bruta diz que ‘A’ expressa a propriedade A se e somente se A causa ‘A’ de modo confiável.

Em resumo, a ocorrência simbólica ‘cão’ não expressa a propriedade de ser um “cão-ou-gato” porque a conexão entre o cão real e a ocorrência simbólica ‘gato’ é assimetricamente dependente da conexão entre o gato real e a ocorrência simbólica ‘gato’, isto é, a conexão cão-‘gato’ depende da conexão gato-‘gato’, mas o contrário não é o caso.

Porém, para admitir casos de deturpação seria necessário reformular a definição da Teoria Causal Bruta, excluindo os termos “todas” e “só”, pois esses termos impõem uma ligação inequívoca e exclusiva entre a ocorrência simbólica e a propriedade representada. Por isso, Fodor elabora uma versão mais branda de sua teoria, a *Teoria Causal Ligeiramente Menos Bruta*, que retira essas exigências

da definição anterior. Assim, nesta nova versão, *um símbolo expressa uma propriedade se for nomologicamente necessário que as instâncias desta propriedade causem ocorrências do símbolo.*

Essa nova definição, portanto, abarca os casos de deturpação, ou seja, quando um 'x' é causado por um *não-x*. Portanto, de acordo com a nova definição, uma ocorrência 'cão' causada por um gato constitui um caso de deturpação previsto na teoria que não expressa uma disjunção de propriedades (dada a situação de dependência assimétrica explicada anteriormente), enquanto que um 'gato' causado por um gato é um caso normal de representação.

3.4 Como Cummins critica a teoria de Fodor

Cummins (1989, p. 61) critica a tese da dependência assimétrica como solução ao problema da disjunção. Ele diz que a assimetria requerida depende da verdade do contrafactual C_1 : "se gatos não causam 'gato', então cães não causam 'gato'" e da falsidade do contrafactual C_2 : "se cães não causam 'gato', então gatos não causam 'gato'". Ele argumenta, porém, que só podemos asseverar ao mesmo tempo a verdade de C_1 e a falsidade de C_2 se aplicarmos regras de avaliação diferentes para cada um deles. Examinemos o argumento.

Assumindo que "se cães causassem 'gato' seria porque eles se parecem com gatos", analisemos primeiramente C_2 . Se cães não causassem 'gato', seria porque: (1) a aparência de gato não causa 'gato', ou (2) cães não causam a aparência de gato. Se cães não causassem 'gato' porque 1, gatos também não causariam 'gato', e o segundo contrafactual seria verdadeiro. Portanto, precisamos assumir 2. Neste caso, como os gatos continuariam causando sua aparência, eles ainda causariam 'gato' e C_2 seria falso, como requerido.

Quanto a C_1 , também temos duas maneiras de avaliar a questão: começamos considerando que, se gatos não causam 'gato', seria porque: (3) a aparência de gato não causa 'gato', ou (4) gatos não causam a aparência de gato. Se a aparência de gato não causasse 'gato', então cães também não causariam 'gato' – uma vez que, por hipótese, os cães causem erros de representação por se parecerem com gatos. Desse modo, C_1 seria verdadeira, como requerido. Todavia, esta maneira de fazer C_1 verdadeira também faz C_2 verdadeira, como demonstrado acima. Então precisamos assumir 4: gatos não causam aparência de gato. Mas

como isso não afetará a conexão cão-‘gato’, C_1 é falsa, ao contrário dos requisitos.

Como os mundos possíveis em que C_1 é verdadeira e C_2 é falsa não são os mesmos, então não existe uma única interpretação que faça C_1 verdadeiro e C_2 falso. Portanto, C_1 e C_2 não expressam em conjunto algo sobre a conexão entre gatos e cães e entre gatos e ‘gato’.

Quanto ao problema da subdeterminação, Cummins (1989, p. 62-63) argumenta que, se alguns gatos não causam a ocorrência simbólica ‘gato’, então a extensão de ‘gato’ deve ser o subconjunto dos gatos que causam ‘gato’. Mas como a Teoria Causal Bruta precisa evitar a possibilidade de que ‘gato’ denote apenas alguns gatos (uma vez que uma teoria da representação não pode deixar de fora gatos gordos e peludos, por exemplo, mas deve ser capaz de explicar como todos os tipos de gato podem ser representados), a única maneira de contornar esse problema é insistindo na covariação genuína: todos os gatos causam ‘gato’, o que acaba por comprometer Fodor com a ideia de que sistemas cognitivos são inerrantes.

Fodor rejeita como um absurdo a ideia de que sistemas cognitivos possam ser inerrantes, mas, para evitar essa consequência, é forçoso evitar a exigência de que só gatos causem ‘gato’. É preciso lidar satisfatoriamente com a possibilidade de que outras coisas também causem ‘gato’, ou que alguns gatos não causem ‘gato’.

Segundo Cummins (1989, p. 63), se alguns gatos não causam ‘gato’, então a extensão de ‘gato’ é o subconjunto dos gatos que causam ‘gato’. Contudo, é preciso descartar a possibilidade de que ‘gato’ expresse uma propriedade particular, como ser preto ou gordo, e a única maneira de contornar isso é insistir na covariação genuína: todos os gatos causam ‘gato’.

Cummins (1989, p. 63-65), argumenta que, de acordo com a Teoria Causal Bruta, há possibilidade de um gato causar um ‘gato’ somente se o sistema estiver preparado para prestar a devida atenção e fazer as inferências corretas com base no conhecimento correto. É preciso apenas que haja alguma garantia de que o mecanismo causal existe (circunstâncias em que gatos são suficientes para causar ‘gato’) e uma maneira sem petições de princípio de escolher esse mecanismo. Tudo o que a teoria de Fodor requer é uma maneira “naturalista” de selecionar tal mecanismo, isto é, sem o uso explícito de termos intencionais ou semânticos.

Talvez possamos escolher o mecanismo em termos “naturalistas”,

prossegue Cummins, mas a Teoria Causal Bruta sustenta que não podemos entendê-los ou descrevê-los sem algum jargão representacional. Isso implica que há uma circularidade na teoria de Fodor, pois a representação se baseia em sua covariação com o representado, mas a covariação se baseia num mecanismo que, nas devidas condições, produzirá um 'gato' a partir de um gato. Fodor, porém, argumenta que é possível para um computacionalista especificar condições causalmente suficientes para um gato causar um 'gato' sem contrabandear noções intencionais ou semânticas.

Mas, de acordo com a Teoria Causal Bruta, prossegue Cummins (1989, p. 65), o mecanismo em questão só pode ser entendido apelando-se para representações internas, pois ele é inferido a partir do conhecimento armazenado. Logo, para compreender o mecanismo que a Teoria Causal Bruta invoca para explicar a covariância entre gatos e 'gato', já devemos entender a representação e o papel explicativo que ela desempenha nos mecanismos mentais. Isso é o suficiente para minar a capacidade das teorias da covariância de explicar a natureza da representação na Teoria Causal Bruta, pois, para explicar a representação em termos de covariação, é preciso explicar o que é covariação, mas, ao tentar explicar o que é covariação, faz-se necessário dizer o que é representação.

Por isso, prossegue Cummins, não basta evitar o vocabulário intencional ou semântico, antes, é preciso fazê-lo de modo a explicar o que é representação. O problema era dizer em que condições os gatos são suficientes para causar 'gato', e fazê-lo em um vocabulário naturalista, mas o que Fodor faz é encontrar uma ocasião real O_R em que um gato causa um 'gato' através do mecanismo M_T que "faz o truque", não importa como, e constrói o contrafactual desejado:

Se M_T opera num gato em circunstâncias como as obtidas em O_R então resulta um 'gato'.

Porém, o covariacionista nos diz que há representação porque há covariância. E a Teoria Causal Bruta nos diz que há covariância porque há representação. Mas não é possível ter as duas coisas sem prejudicar o poder explicativo de uma delas. E uma vez que precisamos explicar a representação de uma forma que subscreva (e não enfraqueça) seu papel explicativo na Teoria Causal Bruta, é a doutrina da covariação que deve ser posta de lado (CUMMINS, 1989, p. 65-66).

Para o covariacionista, permitir o erro implica a incapacidade de escolher

um mecanismo que produzirá um 'gato' em resposta a um gato arbitrário. Portanto, a única maneira de fazer isso é por referência a alguma(s) instância(s) específica(s) em que um gato produz um 'gato', de modo que "para todo sujeito S, se S estivesse em uma tal situação, um gato produziria um 'gato'" (CUMMINS, 1989, p. 66).

A sensação de que não temos mais uma explicação da representação pode ser atribuída ao demonstrativo. O relato é essencialmente ostensivo. "Representação", diz o covariacionista, "é quando você tem um caso como esse." E dá um exemplo de como seria: "Como quando você pensa que há um gato ali porque de fato há." É impossível especificar o C de "Em C, qualquer gato causaria um 'gato' em S", portanto, o covariacionista deve, enfim, recorrer à ostensão e esperar que ninguém perceba que não há como generalizar de modo fundamentado sobre o exemplo (CUMMINS, 1989, p. 66).

3.5 Como Fodor poderia responder a Cummins?

Retomemos a argumentação de Cummins exposta anteriormente na qual ele critica a solução de Fodor ao problema da disjunção. Ele afirma que o sucesso da solução em questão depende da verdade do contrafactual C_1 e da falsidade do contrafactual C_2 . Porém, Cummins afirma que não há um mundo possível em que C_1 seja verdadeiro e C_2 falso, de modo que essa solução é impossível e argumentação de Fodor é inválida. Relembremos os contrafactuais em questão:

C_1 : *se não é o caso que gatos causam 'gato', então não é o caso que cães causam 'gato'* ($\neg p \rightarrow \neg q$).

C_2 : *se não é o caso que cães causam 'gato', então não é o caso que gatos causam 'gato'* ($\neg q \rightarrow \neg p$).

Cummins pede ao leitor que assuma que "se cães causam 'gato' então cães se parecem com gatos" e parte para a análise de C_2 .

Ele afirma: "se cães não causam 'gato', é porque: (1) a aparência de gato não causa 'gato', ou (2) cães não causam a aparência de gato."

Contudo, *essas opções não são exaustivas*. Levando em consideração a alteração que Fodor faz em sua teoria (ou seja, a Teoria Causal Ligeiramente Menos Bruta), uma terceira alternativa válida seria: *(5) cães não causam aparência de gato em situações funcionais e não limitadas*, ou, o que resulta no mesmo, *cães causam*

aparência de gato em situações de percepção disfuncional ou limitada.

Prossegue Cummins: “se cães não causam ‘gato’ porque 1, gatos também não causariam ‘gato’, e C_2 seria verdadeiro. Portanto, precisamos assumir 2”. Mas, como vimos, poderíamos também assumir 5.

Assumindo 2 ou 5, os gatos continuariam causando sua aparência, portanto, continuariam causando ‘gato’ e C_2 seria falso, como requerido.

Há também duas maneiras de avaliar C_1 : se gatos não causam ‘gato’, seria porque: (3) a aparência de gato não causa ‘gato’, ou (4) gatos não causam a aparência de gato. Porém, poderíamos incluir: (6) *gatos não causam a aparência de gato em situações de percepção disfuncional ou limitada.*

Se a aparência de gato não causasse ‘gato’, então cães também não causariam ‘gato’ – uma vez que, por hipótese, os cães causem erros de representação por se parecerem com gatos. Desse modo, C_1 seria verdadeira, como requerido. Todavia, esta maneira de fazer C_1 verdadeira também faz C_2 verdadeira, como demonstrado acima. Então precisamos assumir 4 ou 6.

Assumindo 4, temos que gatos não causam aparência de gato. Mas como isso não afetará a conexão cão-‘gato’, C_1 é falsa, ao contrário dos requisitos.

De fato, a conexão cão-‘gato’ não é afetada, mas isso não implica a falsidade de C_1 , pois a verdade de “gatos não causam a aparência de gato”, não falseia nem confirma a afirmação “cães causam ‘gato’”. E para que C_1 seja falsa, deve ser o caso que “cães causam ‘gato’”.

Assumindo 6, temos que gatos não causam a aparência de gato (e nem ‘gato’, portanto) em situações de percepção disfuncional ou limitada. Ora, a restrição inserida pela ideia de percepção disfuncional / limitada permite que os casos de deturpação possam ser teoricamente descritos sem afetar a verdade de C_1 .

Portanto, existe um mundo possível tal que C_1 é verdadeira e C_2 é falsa. Logo, seguindo os critérios de análise do próprio Cummins, a solução de Fodor ao problema da disjunção é perfeitamente aplicável, pois não viola nenhuma das condições propostas por Cummins para a validade da assimetria.

Em segundo lugar, penso que é preciso considerar que a asseveração de um sujeito S “isto é cão” ou “isto é gato” não necessariamente corresponde ao que de fato foi captado pelo sistema perceptual de S. Como discutido acima, ao tratarmos de Locke, não é possível simplificar o processo da percepção e representação humana para funcionar tal e qual o sistema LOCKE.

Dados confusos não podem produzir respostas exatas, e se é necessário responder com exatidão, pode ser que se escolha uma ou outra asseveração (“isto é cão”, “isto é gato”) por uma necessidade prática, e não por fidelidade àquilo que foi representado na mente. Pode ser, por exemplo, que por qualquer razão de sociabilidade e convivência não queiramos deixar sem resposta a indagação de um interlocutor sobre algo que não tivemos condições de perceber adequadamente. Assim, se uma criança pergunta a seu pai: “que foi aquilo que passou correndo ali na frente?” pode ser que a resposta seja algo como “foi uma raposa” simplesmente porque se quer encerrar o assunto em vez de suscitar mais perguntas, mesmo que esse pai não tenha representado em sua mente algo como uma raposa.

Outra possibilidade é uma falha cognitiva. Caso exista um defeito no aparato cognitivo de S que o torne incapaz de perceber rostos, não seria correto dizer que o borrão que S vê ao por-se diante da pintura de Monalisa é causado pelo rosto de Monalisa, pois não faz diferença qual rosto está diante de S, qualquer um será visto por S como um borrão. No entanto, caso S tivesse uma cognição normal, o rosto que S veria diante da pintura de Monalisa seria causado pelo rosto de Monalisa. O que faz com que o rosto de Monalisa seja a causa de sua própria representação mental é que há uma relação causal entre essas duas coisas quando as condições de funcionamento cognitivo são apropriadas, ao passo que o que faz com que o rosto de Monalisa não seja a causa da representação mental de borrão é que *não há* uma relação causal entre essas duas coisas quando as condições de funcionamento cognitivo são apropriadas.

Minha opinião é de que a afirmação que assevera o conteúdo de uma representação não é a representação mesma, mas um juízo sobre a representação, uma crença, uma proposição que pode ser verdadeira ou falsa.

Os objetos afetam nossos sentidos, e estes enviam dados ao cérebro. Esse momento ocorre antes de qualquer acesso mental de seu conteúdo, antes da reunião dos dados sensíveis que captamos numa representação mental.

Esses dados são unificados numa mesma representação (se assim não fosse não reconheceríamos gatos brancos miando, mas a figura do gato, a cor do gato, o som do gato, etc, como coisas separadas), e esse momento de unificação é ainda não proposicional, pois não se afirma ou nega nada. Numa comparação gramatical, aqui é como se tivéssemos apenas o sujeito da frase.

Essa representação do indivíduo é retida na memória pelo menos por

algum tempo (posso me lembrar perfeitamente da maçã que comi ontem, pode ser que me lembre dela para o resto da vida, mas pode ser que me esqueça dela como já esqueci de tantas outras). Não entrarei na discussão sobre os fatores que geram memórias mais ou menos duradouras.

Essa representação da maçã individual passa a compor (também a modo de memória) uma representação mais geral e abstrata de maçã, uma espécie de “maçã mental” que serve como base de comparação para o reconhecimento de novos indivíduos do mesmo tipo. Essa etapa é ainda não proposicional. A comparação gramatical que mais se aproxima dessa representação abstrata seria uma oração imperfeita (sem verbo de cópula), algo como “isto: maçã”.

Peço ao leitor que imagine o sabor de maçã, a imagem de uma maçã, seu peso, sua textura, sua temperatura. Pergunto: o leitor imaginou sabor, imagem, peso, textura, temperatura de qual maçã? Pode ser que tenha se lembrado de uma maçã em especial, mas acredito que na maior parte das vezes os atributos imaginados pertencerão a uma espécie de “maçã mental” cujos atributos são uma seleção aleatória a partir de uma coleção dos atributos das maçãs que já comemos, tocamos, cheiramos, etc. Acredito que a rememoração da própria experiência através de nossa capacidade imaginativa é prova suficiente e dispensa argumentações mais elaboradas em favor da “maçã mental”.

É também por causa dessa representação abstrata que quem vê uma perfeita maçã de dois metros de altura tem todo o direito de duvidar que seja mesmo uma maçã, pois foge ao padrão formado pela experiência, ou seja, podemos pensar que não há no “espaço amostral” da memória casos em número suficiente de maçãs gigantes para assegurar a especificação “maçã” ou talvez que a “maçã mental” não seja assim tão grande.

É a partir dessa “maçã mental” que podemos inferir aquilo que as maçãs são em essência, através de um conceito de maçã, tal que maçã = “um ente material animado vivo vegetal parte (isto é, parte da macieira)”. Obviamente esse é apenas um exemplo de conceito. Como esse conceito deriva da experiência individual e nem sempre é verbalizado, varia de indivíduo para indivíduo, apesar de sua funcionalidade prática.

Quando afirmamos “isto é uma maçã” estamos asseverando uma proposição que se reporta ao conceito mais ou menos adequado que temos de maçã. É a partir desse momento, quando somos capazes de dizer “isto é maçã”,

“isto não é maçã” é que nosso pensamento pode se tornar inferencial, pois podemos conectar diferentes proposições num raciocínio e obter conclusões.

Assim, ‘gato’ denota, para um sujeito S, qualquer gato que possa ser devidamente classificado pela experiência de gatos que S guarda na memória. Portanto há, pelo menos, mais uma alternativa que explica os casos de gatos que não causam ‘gato’, que pode ser: ruídos no processo de captação dos dados (circunstâncias não ideais de percepção); insuficiência de dados (circunstâncias não ideais do ambiente); ou a inexperiência de S (insuficiência de dados em memória). Afinal, não parece sensato cobrar de alguém que não tenha nenhuma experiência com gatos que conheça sobre seus hábitos, como gostar de leite ou levantar a cauda quando estão felizes, atributos próprios dos gatos.

Desse modo, podemos dizer que o processo de representação deve causar ‘gato’ a partir de um gato se e somente se S tem condições ambientais de percepção suficientes para gerar os dados necessários, se os dados disponíveis são captados em número suficiente e se o conhecimento de S é capaz de classificar esses dados. Nesse sentido, o processo de representação representa porque covaria com o objeto, mas o que é representado não é a classe do objeto (a representação não é o símbolo da classe), mas a reunião dos dados fornecidos por ele, afinal, pelos sentidos nós não captamos o gato, mas, *do* gato, certa cor (objeto da visão), certo som (objeto da audição), certa textura (objeto do tato), certo odor (objeto do olfato) e, quem sabe, certo gosto (objeto da gustação).

Assim, a classe gato covaria com certo gato só indiretamente, enquanto covaria com a representação que reúne os dados que dele captamos. A classe é símbolo de uma representação classificável, enquanto a representação é símbolo do objeto captado na percepção, que pode corresponder mais ou menos ao objeto real de acordo com as condições já discutidas.

4 A SEMÂNTICA INTERPRETATIVA DE CUMMINS

Como vimos anteriormente, Cummins (1989, p. 87-111) considera que a suposição de que as representações mentais covariam com os estados do mundo não é compatível com a TCM, e propõe uma outra solução que ele considera compatível. Trata-se da *semântica interpretativa*, que, diferentemente da maioria dos posicionamentos dentro da discussão sobre a representação, se concentra no conteúdo das representações de um sistema computacional e não no conteúdo (intencional) dos pensamentos.

A tese da semântica interpretativa é que a TCM requer representações simuladas para fundamentar seus apelos explicativos à representação. A semântica interpretativa é ontologicamente equivalente a um tipo de semântica de papéis funcionais¹⁷, mas, como parte de motivações diferentes, acaba por direcionar a atenção para outras questões, gerando uma discussão diferente daquela de teorias tradicionais sobre papéis funcionais (CUMMINS, 1989, p. 87-88).

A semântica de papéis funcionais aplicada ao problema da representação mental é uma aplicação específica do funcionalismo às representações mentais, em vez de ser aplicado aos estados e eventos mentais de um modo geral. De acordo com o funcionalismo, estados mentais são funcionalmente individuados, mas, alguns estados mentais, como as representações, são individuados por seu conteúdo. Por isso o funcionalismo implica que conteúdos mentais são funcionalmente individuados (CUMMINS, 1989, p. 114). Ou seja, pessoas são sistemas pensantes em virtude da estrutura computacional do cérebro. Pensamentos e estados com conteúdo podem ser o que são em virtude de seu lugar nessa estrutura. (CUMMINS, 1989, p. 123).

É uma tese absolutamente central da TCM que a representação em

17 “o papel funcional de uma atitude [proposicional] é a rede característica de conexões causais em que ela entra, a maneira como ela interatua causalmente com dados provenientes do meio ambiente, com outros estados mentais e com o comportamento. Considere-se, p. ex., a crença que eu tenho de que daqui a pouco vai chover. Grosso modo, o papel funcional dessa crença seria especificado por meio da consideração de fatos do seguinte gênero: a) o fato de a crença ser normalmente causada por certo tipo de *input* [ou entrada] sensorial (p. ex., a minha percepção visual de nuvens cinzentas no céu); b) o fato de a crença ser usualmente uma causa de, bem como um efeito de, certos outros estados mentais (p. ex., um efeito da crença de que nuvens cinzentas no céu prenunciam chuva); e c) o fato de a crença, em interação com outros estados mentais (em particular, certos desejos), dar normalmente origem a certo comportamento: tomada em conjunção com o desejo de não me molhar (e com outros estados mentais), ela pode levar-me a ir buscar um guarda-chuva.” (BRANQUINHO *et al*, 2006, p. 78).

sistemas cognitivos é exatamente a mesma coisa que a representação em sistemas computacionais de um modo geral. Calculadoras representam do mesmo modo que pessoas, muito embora não sejam sistemas de crença-desejo (CUMMINS, 1989, p. 118). A TCM deve explicar as capacidades cognitivas recorrendo à maneira das calculadoras aritméticas, à representação e à computação. Por isso, Cummins (1989, p. 89-91) explica a modo de introdução, como máquinas de somar são capazes de executar uma função de soma, exemplo que servirá de modelo para explicar as representações simuladas requeridas pela TCM.

De acordo com essa explicação, uma máquina de somar não executa a função $f(a, b) = a + b$ diretamente, pois números não podem ser estados, processos ou eventos de nenhum sistema físico. Porém, máquinas de somar podem ter estados que são numerais, ou seja, representações de números, e podem, desse modo, executar uma instância da função f .

Seja g uma instância de f que relaciona uma sequência de pressionamento de teclas a uma exibição em tela, tal que $g(N_1, +, N_2, =) = E$.

Uma máquina de somar M satisfaz g se os argumentos e valores de g são estados físicos de M , de modo que máquinas de somar podem executar f indiretamente ao instanciar g , cujos argumentos e valores representam os argumentos e valores de f , ou seja, g tem esses argumentos e valores como interpretações. Resumidamente, para que um dispositivo físico adicione basta que ele satisfaça uma função que pode ser interpretada como adição.

Os argumentos de g que funcionam como representações dos argumentos de f não necessitam ser símbolos com significado fora do sistema que os computa, sendo necessário apenas que uma função de interpretação estabeleça um mapeamento entre g e f . Nesse sentido, a representação é a relação induzida pelo mapeamento interpretativo entre os elementos de g e os elementos de f (CUMMINS, 1989, p. 92)

Seja i uma função de interpretação que estabelece um mapeamento entre os elementos de f e os elementos de g , temos que $i(e_1 \dots e_n)$ corresponde a $\langle a, b \rangle$ e a $i(E)$ corresponde $a+b$, de modo que $g(e_1 \dots e_n) = E \leftrightarrow f(i(e_1 \dots e_n)) = i(E)$.

A TCM considera que um sistema é cognitivo em virtude das restrições epistemológicas apropriadas ao seu domínio de tarefas (o inferir), e propõe uma compreensão dos sistemas cognitivos como sendo *sistemas que computacionalmente instanciam funções cognitivas*. Porém, só podemos considerar

um sistema como cognitivo se pudermos descrever suas ações semanticamente, pois restrições epistemológicas só são definidas para proposições ou coisas com conteúdo proposicional e valor de verdade (CUMMINS, 1989, p. 108).

Para a TCM, ter capacidade cognitiva é instanciar uma função que relaciona conteúdos proposicionais inferencialmente, de modo que, explicar a cognição é explicar como um sistema pode ser inferencial. Nesse caso, o sistema requerido pela TCM pode ser caracterizado como inferencial porque calcula símbolos que representam conclusões a partir de símbolos que representam as premissas correspondentes, tal que, para o sistema, mover-se disciplinadamente símbolo a símbolo (computar), equivale a, dada a interpretação adequada, mover-se disciplinadamente conteúdo a conteúdo (CUMMINS, 1989, p. 109-110).

Por isso, dentro do arcabouço teórico da semântica interpretativa, dizer que um mecanismo é cognitivo é o mesmo que dizer que o comportamento deste mecanismo pode ser descrito por uma interpretação que o revela como racional, uma vez que a computação de símbolos instancie uma função inferencial. Além disso, o que resulta da semântica interpretativa é que as representações são simulações, já que as relações entre símbolos espelham (via interpretação) as relações entre conteúdos, isto é, no mesmo sentido em que uma função de onda senoidal pode representar a corrente alternada ou que a área de um retângulo pode representar a carga da bateria de um *laptop*. Nesse sentido,

Para a TCM, a representação é apenas um aspecto da simulação. No caso especial das funções cognitivas isto toma um significado especial para a filosofia [...]. Significados são apenas aquilo que a interpretação prevê. A interpretação é a noção fundamental; significado é só um nome para o [seu] produto (CUMMINS, 1989, p. 125, tradução nossa).¹⁸

Por isso, Cummins usa o termo *representação simulada* para denotar o tipo de representação requerido pela semântica interpretativa.

O sucesso ou falha das representações simuladas não depende só dos símbolos e da interpretação, mas também, e de modo essencial, dos processos que acessam esses símbolos, pois um erro de lógica no programa de um tal processo

18 “For the CTC, representation is simply an aspect of simulation. In the special case of cognitive functions, this takes on a special significance for philosophy, though it is nothing special to the CTC, because when the simulated function is cognitive, meaning maps a mechanical process onto a rational one, revealing mind where there was only mechanism. But the concept of meaning has a life in the CTC only as the child of interpretation. Meanings just are what interpretations provide. Interpretation is the fundamental notion here; meaning is just a name for the product.”

poderia gerar erros de representação, o que poderia ser corrigido alterando as estruturas de dados do programa ou o próprio algoritmo, ou seja, a lógica interna desse programa (CUMMINS, 1989, p. 98).

Além disso, podemos imaginar como uma função interpretativa pode ser mais ou menos adequada. Caso uma interpretação mapeie entre uma função de soma $f(a, b) = a+b$ e uma simulação de produto $h(N, \times, M, =) = E$, isso não faria diferença no caso de $a=b=0$, por exemplo, embora a interpretação mais adequada, que funcionaria para quaisquer valores de a e b , devesse mapear entre h e uma função de produto $j(a, b) = a \times b$. Isto evidencia como uma interpretação pode ser adequada só em alguns casos, poucos ou muitos, a depender de seu grau de adequação em relação ao alvo da simulação, ou seja, o conteúdo que se pretende representar (CUMMINS, 1989, p. 101).

O fato de uma determinada função de interpretação garantir uma simulação perfeita, totalmente adequada, não implica na impossibilidade de erros de representação, pois, como dito anteriormente, uma representação simulada não depende só de símbolos e interpretação, mas também dos processos em execução. Do mesmo modo, interpretações que geram simulações perfeitas em determinados casos podem se tornar erros de representação se aplicadas a alvos de simulação inadequados. Como no exemplo acima: a representação simulada h é perfeitamente adequada à função de produto j , mas gera erros de representação se aplicada à função de soma f em todos os casos em que $a+b \neq a \times b$ (CUMMINS, 1989, p. 102-103).

Na semântica interpretativa, ou há representação ou não há representação. Apenas quando há representação, é que pode haver erros de representação. Para que haja erro de representação é necessário que exista: (1) uma representação e (2) uma função de interpretação que mapeia de modo imperfeito os elementos da representação simulada nos elementos do objeto, ou um erro de execução nos processos que acessam os dados da representação simulada (CUMMINS, 1989, p. 100).

Esse tipo de explicação para os casos de deturpação é impossível nas teorias covariacionistas porque elas não são capazes de explicar o que é representação de um modo compatível com a TCM (CUMMINS, 1989, p. 102), virtude que a Semântica Interpretativa alcança por fazer descrições de baixo nível a respeito da cognição, ou seja, sem usar um linguajar intencional.

4.1 Objeções à Semântica Interpretativa

Cummins (1989, p. 126-127), adianta algumas objeções que podem ser feitas à Semântica Interpretativa e as classifica em dois tipos: o primeiro está relacionado à natureza do conteúdo mental, tal como ele é explanado em sua teoria, e o segundo trata da relação entre representação e intencionalidade no âmbito dessa teoria.

No primeiro tipo, encontram-se duas objeções. A primeira objeção acusa a Semântica Interpretativa de ser incompatível com certos requisitos do *naturalismo*. Isso ocorreria porque ela seria incapaz de dar uma explicação do conteúdo mental em termos dos conceitos não-intencionais disponíveis nas ciências naturais, apelando, ao invés disso, ao conceito de *interpretação*. A segunda objeção tem relação com essa, mas se concentra na inadequação da explicação da Semântica Interpretativa para o *status causal do conteúdo mental*. Nesse caso, o problema alegado é que uma explicação da potencialidade causal do conteúdo mental que não pode ser assimilada na ordem causal natural deixa de fora um pressuposto básico da ciência cognitiva.

Por outro lado, a objeção característica do segundo tipo afirma que o conteúdo representacional construído pela Semântica Interpretativa é relativo, e isso acarretaria uma desconexão entre representação e *intencionalidade*, ou entre as representações simuladas e as atitudes proposicionais.

Vejamos a seguir cada uma dessas críticas de forma mais detalhada.

4.1.1 Naturalismo

Pode parecer, afirma Cummins (1989, p. 128-129), que a referência essencial que a Semântica Interpretativa faz à noção de interpretação está fadada a violar a restrição contra o emprego de noções semânticas ou intencionais na explicação das propriedades semânticas dos sistemas cognitivos.

Com efeito, a sugestão interpretacionista é esta:

R representa x em um sistema S se e somente se existe uma função g , uma função f e uma interpretação I tal que S satisfaça g , g simule f sob I e $I(R) = x$.

Não obstante, se bem entendida, fica claro que essa sugestão recorre a uma noção de interpretação sem “intérprete”, vale dizer, sem um agente intencional ou cognitivo que executa a interpretação. Na verdade, a Semântica interpretativa invoca apenas uma função de interpretação, e, no caso em consideração, uma função de interpretação é tão somente uma função do conjunto dos R no conjunto dos x, uma relação matemática que não depende de intérprete.

Com esse simples esclarecimento, Cummins rebate a acusação de que a Semântica Interpretativa é incompatível com o naturalismo por recorrer a noções ou agentes intencionais. Mas há outra acusação de antinaturalismo que ainda pode ser levantada contra a teoria de Cummins, e rechaça-la demandará um maior esforço argumentativo.

4.1.2 O status causal do conteúdo de representações simuladas

Conforme explica Cummins, (1989, p. 129-130), é bastante difundida a ideia de que os estados de um sistema que têm conteúdo têm poderes causais, e o apelo a esses poderes é, de algum modo, central para a explicação psicológica. Por outro lado, se os conteúdos das representações mentais não fazem parte da ordem causal, se nada acontece por causa deles, então o apelo a eles não tem um papel explicativo a desempenhar na ciência séria.

Também é geralmente aceito que esses estados são causalmente potentes não em virtude de terem os conteúdos que possuem, mas em virtude de algo mais, como, por exemplo, seus potenciais eletroquímicos ou seus níveis de ativação e conexões. Essa visão (que Cummins chama de “Visão Tradicional”) torna o conteúdo causalmente inerte, mas isso implicaria que o conteúdo não tem um papel explicativo sério?

Há dois aspectos distintos sobre essa questão: (1) que tipo de papel causal (se é que há) é desempenhado pelo conteúdo (tal como elaborado pela Semântica Interpretativa) na TCM? (2) O papel explicativo do conteúdo na TCM identifica-se com um papel causal do conteúdo?

4.1.2.1 Conteúdo e causalidade

Cummins (1989, p. 130-131) considera duas afirmações quanto ao

primeiro aspecto da questão sobre o papel explicativo do conteúdo mental:

(A₁) Existem declarações causais de atribuição de conteúdo verdadeiras nas quais uma causa é identificada por meio de seu conteúdo.

(A₂) Em uma declaração causal de atribuição de conteúdo verdadeira, ter o conteúdo atribuído é (pelo menos algumas vezes) o fator causalmente relevante.

A₁ parece ser ponto pacífico, mas a Visão Tradicional nega A₂. Por isso a argumentação de Cummins se concentra nesta última. Ele convida o leitor a supor que durante a execução de um programa, o computador execute a seguinte instrução: “se o conteúdo do acumulador for igual a zero, carregue o conteúdo do endereço de memória EdM₁ no contador do programa”, sendo que o endereço EdM₁ contém a instrução “adicione *um* ao conteúdo do acumulador” e o conteúdo do acumulador é igual a zero quando a primeira instrução é executada. Então:

(E₁) O conteúdo zero do acumulador causa o conteúdo *um*.

E₁ certamente serve para verificar A₁, pois se trata de uma declaração causal (o primeiro conteúdo do acumulador causa o segundo) e verdadeira (pois, seguindo a lógica de execução do programa, é exatamente isso que deve acontecer). Mas, quanto a A₂, é difícil afirmar se o fator causalmente relevante é o conteúdo ou se são as propriedades eletrônicas dos estados aos quais os conteúdos são atribuídos.

Para Cummins, há apenas dois modos de pensar sobre A₂: a *Via da Generalização* e a *Via dos Contrafactuais*. Vejamos a seguir como ele (1989, p. 131-133) explica a Via da Generalização.

Uma razão para rejeitar A₂ e que fundamenta a Visão Tradicional é que quando generalizamos sobre declarações causais de atribuição de conteúdo, descobrimos que os meios não acidentais de perfazer a generalização são estabelecidos não pelo conteúdo, mas por coisas como propriedades eletrônicas, padrões de ativação ou forma abstrata. Não descobrimos que “para todo *x*, se *x* tem o conteúdo C então *x* é tal e tal”, mas sim que “para todo *x*, se *x* tem uma certa propriedade eletrônica ou padrão de ativação ou forma abstrata, então *x* é tal e tal”.

Pode-se argumentar que: (1) a melhor evidência de que F é um fator

causal em uma conexão causal singular é que a generalização correspondente (sobre F) seja observada e que ela não seja equivalente a nenhuma generalização que não se refira a F, e (2) que a melhor evidência para rejeitar F como um fator causal é a ausência das generalizações correspondentes.

A Via da Generalização nos instiga a aceitar A_2 apenas quando as afirmações causais que atribuem conteúdo singular se generalizam através do conteúdo, caso contrário A_2 deve ser rejeitada. A Visão Tradicional diz que A_2 deve ser rejeitada justamente por entender que as generalizações dessas afirmações causais singulares não dependem essencialmente do conteúdo atribuído por elas, mas de propriedades físicas.

Retornando ao exemplo anterior, explica Cummins (1989, p. 132), não é verdade, em geral, que o acumulador ter o conteúdo *zero* faz com que ele tenha o conteúdo *um*, ou seja, a generalização de E_1 não parece depender essencialmente dos conteúdos em questão. Entretanto, com alguns acréscimos, é possível obter a generalização relevante:

(E_2) Sempre que o acumulador representa *zero* e o contador do programa representa certo endereço de memória contendo uma instrução para incrementar o acumulador, o acumulador (posteriormente) representa *um*.

Tudo o que E_2 requer são coisas que funcionem como endereços de memória, cujos estados sejam capazes de representar instruções; algo que funcione como um acumulador, cujos estados sejam capazes de representar *zero* e *um*; e algo que funcione como um contador de programa, cujos estados sejam capazes de representar endereços de memória.

Essa generalidade só pode ser alcançada especificando os estados em termos de seus conteúdos representacionais e, correlativamente, especificando unidades funcionais em termos de suas capacidades representacionais.

Como a TCM considera a representação como uma característica tanto de sistemas computacionais cognitivos quanto não cognitivos, o argumento também poderia ser apresentado em termos de uma classe mais ampla de sistemas computacionais que instanciam ou simulam outras funções. Portanto, não há nada de especial sobre as funções cognitivas nessa conexão e, portanto, o *status* causal do conteúdo não se baseia em nenhuma tese especial sobre a cognição sustentada

pela TCM.

Segundo Cummins (1989, p. 134-135), outra abordagem para avaliar A_2 é a *Via dos Contrafactuais*, que pergunta se declarações causais de atribuição de conteúdo suportam contrafactuais de atribuição de conteúdo. Seguindo o exemplo anterior, temos o seguinte:

(E₃) Se o acumulador não tivesse representado *zero* no tempo t , então ele não teria representado *um* no tempo t' .

Também temos o seguinte:

(E₄) Se o acumulador não tivesse a propriedade eletrônica PE₁ no tempo t , então ele não teria a propriedade eletrônica PE₂ no tempo t' .

Ambos, E₃ e E₄, são claramente verdadeiros. Caso E₃ seja apenas um reflexo de E₄, então A_2 é falsa e a Visão Tradicional é verdadeira. Para defender A_2 seria preciso apresentar um contrafactual que seja equivalente a E₃ mas que não tenha um análogo verdadeiro formulado em termos de propriedades eletrônicas subjacentes. O problema é que contrafactuais do tipo apresentado não são sensíveis à *forma* como os eventos são identificados, eles apenas relacionam os estados de coisas que mencionam; eles são “*de re*” nesse aspecto e, portanto, não fornecem nenhuma vantagem para distinguir entre E₃ e E₄.

Para contornar esse problema, seria preciso considerar contrafactuais que relacionam *tipos* (isto é, contrafactuais que relacionam uma determinada classe de objetos em vez de instâncias dessa mesma classe), como o incremento do acumulador ou algo semelhante. Mas a evidência para contrafactuais dessa sorte será apenas uma evidência relacionada a *tipos*, ou seja, uma evidência para generalizações. Desse modo, seguir a Via dos Contrafactuais não leva a nenhuma evidência em favor de A_2 além daquilo que a Via da Generalização já proporciona.

Portanto, resta admitir que A_2 repousa na possibilidade de confirmar generalizações que atribuem papel causal a conteúdos. Uma vez que tais generalizações são pressupostas pela TCM, os defensores dessa teoria precisam aceitar A_2 e conceder ao conteúdo qualquer “realidade” ou pertencimento à “ordem natural” que decorra do seu papel causal.

Ainda assim, afirma Cummins (1989, p. 135), permanece a sensação de que é a sintaxe (ou a eletrônica) que exerce o poder causal, e não o conteúdo. Seja como for, a dificuldade pode ser anulada uma vez que seja esclarecida a questão de se, na TCM, o papel explicativo do conteúdo pode ser identificado com seu papel causal.

4.1.2.2 Causalidade do conteúdo e explicação pelo conteúdo

Para que possamos dar uma resposta satisfatória à questão levantada no final da seção anterior, Cummins (1989, p. 136) acredita que é preciso, em primeiro lugar, lembrar que a TCM assume que os sistemas cognitivos podem ser descritos por um conjunto de generalizações de atribuição de conteúdo que são autônomas no sentido de que não se reduzem a generalizações que não fazem atribuição de conteúdo. Além disso, é perfeitamente compatível com a TCM a suposição de que estados de sistemas cognitivos tenham efeitos causais em virtude de seus conteúdos.

No entanto, nada implica que o papel explicativo do conteúdo na TCM tenha relação com seu *status* causal ou com o fato de existirem generalizações importantes e autônomas que atribuem papel causal a conteúdos. A descoberta de tais generalizações é, obviamente, crucial para a ciência cognitiva, conforme a TCM a vê, pois descobri-las é o mesmo que especificar os mecanismos causais responsáveis pelas capacidades cognitivas dos sistemas em estudo. No entanto, identificar o papel explicativo do conteúdo na TCM com a subordinação de eventos particulares a generalizações que atribuem papel causal a conteúdos seria um duplo equívoco.

Primeiro, porque isso distorceria o que a TCM quer explicar sobre os sistemas cognitivos, a saber, as operações formais que os processos cognitivos realizam ao empregar e manipular representações mentais. Em segundo lugar, perder-se-ia a característica mais distintiva da explicação pelo conteúdo na TCM, ou seja, que, sob uma interpretação, um mero processo computacional e, em última instância, um mero processo físico, é revelado como uma simulação ou uma instância de outro processo.

A interpretação relaciona análises em duas dimensões diferentes, dimensões que não precisam estar causalmente relacionadas, portanto, quando

vemos um processo em termos de sua interpretação, ou seja, em termos de seus conteúdos de representação simulada, vemos isso como algo diferente e independente dos processos físicos subjacentes. Rotações de engrenagens, por exemplo, não participariam de um processo computacional como engrenagens, mas poderiam participar se interpretadas como etapas em um algoritmo de divisão.

4.1.3 Intencionalidade

De acordo com Cummins (1989, p. 137), a representação, como a Semântica Interpretativa a entende, é algo muito simples: o sistema S representa (de forma simulada) os argumentos e valores de qualquer função que simula. Contudo, não se segue daí que S represente qualquer coisa que se queira, segue-se apenas que os conteúdos representacionais são radicalmente variáveis. De fato, embora a TCM possa viver com a ideia de que máquinas de somar podem simular outras funções além de +, parece inverossímil que quando alguém pensa na América Central também está inevitavelmente pensando em um número, afinal, números e crenças são muito diferentes em si.

Certamente existem inúmeras proposições numéricas que nunca foram e nunca serão o conteúdo de nenhum dos meus pensamentos. E, no entanto, minhas estruturas de dados (assumindo, com a TCM, que eu as tenho) certamente representam números se eles representam qualquer coisa, já que qualquer coisa com qualquer interpretação sistemática está fadada a ter uma interpretação numérica sistemática.

Por isso Cummins se pergunta como reconciliar a não singularidade do conteúdo representacional com o fato óbvio de que pensamentos sobre a América Central não são pensamentos sobre números. Vejamos como ele soluciona essa questão.

4.1.3.1 Representações simuladas e intencionalidade

Cummins (1989, p. 137-139) abre sua discussão sobre a relação entre representações simuladas e intencionalidade recordando que um sistema com intencionalidade é um sistema com crenças, desejos e (talvez) outros estados psicológicos da psicologia do senso comum individualizados por seus conteúdos.

Assim, ter uma crença com o conteúdo de que a política dos EUA na América Central é perigosa é estar em um estado intencional, e o mesmo pode ser dito do desejo de encontrar a fonte da juventude. Porém, o simples fato de abrigar uma estrutura de dados com o conteúdo “a política dos EUA na América Central é perigosa” ainda não significa estar em um estado intencional, pois há uma distinção entre a mera representação simulada e a intencionalidade.

Os conteúdos dos pensamentos típicos de humanos adultos são surpreendentemente únicos e determinados. Quando penso que a atual política dos EUA na América Central é imprudente, a proposição “a atual política dos EUA na América Central é imprudente” é o conteúdo intencional do meu pensamento. A proposição “a atual política dos EUA na África do Sul é imprudente” não é o conteúdo intencional desse pensamento (embora pode ser algo em que acredito e, portanto, o conteúdo intencional de um pensamento diferente).

É importante notar que o conteúdo de um pensamento específico se distinguirá até mesmo de um conteúdo que expresse uma proposição verofuncionalmente equivalente.¹⁹ Com efeito, é lugar-comum afirmar que, dadas as proposições logicamente equivalentes p e p' , um sujeito S pode ter um pensamento com o conteúdo intencional de p sem que tenha qualquer pensamento com o conteúdo de p' . Por exemplo, é possível que S acredite que *está chovendo ou não está chovendo* sem que acredite que *a lua é de queijo ou o céu é vermelho se a lua é de queijo*²⁰. Uma análise adequada desse fenômeno e da dificuldade que ele coloca para a Semântica Interpretativa e para a TCM depende de considerarmos algumas questões que remontam à TRM.

De acordo com a TRM, os estados intencionais herdam suas propriedades intencionais das propriedades semânticas de sua representação constituinte. Um pensamento é sobre a América Central e não sobre um número (o número de Gödel da “América Central”, por exemplo) porque a estrutura de dados que abriga essa crença representa a América Central e não um número. É óbvio que esta teoria nunca funcionará se alguém aceitar a semântica interpretativa, pois a semântica interpretativa implica que essa estrutura de dados representa um número.

19 Verofuncionalidade é a propriedade de representar uma função de verdade (funções que tomam valores de verdade como argumentos ou valores). Cf. Branquinho *et al* (2006, p. 127).

20 Como ambas as declarações são tautológicas, elas são logicamente equivalentes.

Tampouco podemos alinhar o conteúdo intencional com o conteúdo representacional relativizando o primeiro da maneira como se relativizou este último. Simplesmente não é verdade que meus pensamentos sobre a América Central são sobre a América Central apenas em relação a alguma escolha de domínio (função a ser simulada) e interpretação. Uma estrutura de dados representa coisas diferentes sob diferentes interpretações, mas os pensamentos não representam nada desse modo. Eles apenas são sobre o que são. A semântica interpretativa deve rejeitar a TRM. Assim, se Cummins está certo sobre o compromisso da TCM com a Semântica Interpretativa, a TCM também deve rejeitar a TRM.

Essa incompatibilidade entre TCM e TRM ajuda a explicar por que o problema da representação tem parecido tão difícil. A TRM é bastante popular entre os simpatizantes da TCM. A TRM exige que a representação seja discriminatória: como posso pensar em camundongos sem pensar em musaranhos, mesmo que eu não saiba a diferença? A TRM exigirá que eu seja capaz de representar um sem, portanto, representar o outro.

Este é o problema em que incorrem a teoria de covariância de Locke e a teoria do papel adaptativo. Talvez, em condições ideais, a representação 'camundongo' apareceria em minhas percepções apenas no caso de eu ser confrontado por um camundongo, ao passo que um musaranho não causaria um 'camundongo'. Ou talvez o mecanismo que "interpreta" gerando um 'camundongo' em mim foi replicado em mim por causa de minha experiência prévia de camundongos mas não de musaranhos.

Se se assume a TRM, ou qualquer outra coisa que alinhe conteúdos representacionais com conteúdos intencionais, o problema da representação se parecerá muito com o problema da intencionalidade; a questão central será conceber uma explicação do conteúdo representacional que os individue de forma tão sutil e única quanto os conteúdos intencionais são individualizados, ou seja, tão refinada e exclusiva quanto a forma pela qual as crenças são individualizadas.

Se, no entanto, não se adere à TRM ou a qualquer coisa comparável, então é possível distinguir fortemente o problema do conteúdo intencional do problema do conteúdo representacional e, portanto, adotar uma explicação da representação que simplifica a explicação do conteúdo representacional. Só rejeitando a TRM, é possível aderir à TCM e à Semântica Interpretativa que a acompanha.

4.1.3.2 Se a Teoria Computacional da Mente pode ignorar a intencionalidade

Para Cummins (1989, p. 139-141), se a cognição for entendida como uma questão de entrar em estados intencionais como crenças, por exemplo, então a TCM não pode explicar a cognição, pois explicar como alguém entra no estado intencional correto é, pelo menos em parte, explicar como esse alguém entra em um estado com o conteúdo correto. Porém, o que a TCM pode fazer é explicar como alguém entra em um estado de representação simulada tendo esse mesmo conteúdo.

Mas o simples “entrar em um estado de representação simulada” não explica por que o estado intencional tem o conteúdo que tem e não outro, enquanto o estado de representação simulada é obrigado a ter outros conteúdos. Se as funções cognitivas forem especificadas intencionalmente, então a TCM deverá explicar funções especificadas intencionalmente, o que exige mais do que representações simuladas para ser feito.

Contudo, não é óbvio que as funções cognitivas devam ser, ou mesmo possam ser intencionalmente especificadas. A TCM busca uma psicologia individualista que se concentra nas capacidades cognitivas que podem ser exercidas em ambientes radicalmente diferentes. Se a posição anti-individualista em relação à intencionalidade está certa, isto é, se as crenças e os desejos não podem ser especificados de forma independente do ambiente, então os *explananda*²¹ de uma psicologia individualista não podem ser especificados intencionalmente.

Portanto, a TCM não deve se preocupar com *explananda* intencionalmente especificados. O que os argumentos anti-individualistas provam do ponto de vista da TCM é que crenças e desejos não são estados psicológicos no sentido específico que interessa à TCM.

A TCM não se limita à explicação das funções cognitivas, pois, ao contrário da típica filosofia da mente, ela começa não com crenças e desejos, mas com a simulação computacional de várias funções especificadas. A proposta sobre a cognição é simplesmente uma extensão dessa estratégia, extensão essa que deve ser efetuada considerando funções que incorporam restrições epistemológicas apropriadas para alguma tarefa.

²¹ *Explananda* é um termo que se refere aos fatos, fenômenos ou eventos que precisam ser explicados. Em outras palavras, refere-se aos objetos de estudo de determinada disciplina. Na prática científica, por exemplo, é aquilo que os cientistas buscam compreender e explicar por meio de teorias, leis e modelos.

As funções que respeitam as condições de cogência não precisam, em geral, ser especificadas em termos intencionais (i.e., em termos de crença e desejo), embora precisem ser especificadas em termos semânticos. A função que especifica a validade dedutiva é um exemplo óbvio. Recusar-se a chamar cognitivas às operações que respeitam a cogência simplesmente porque seus argumentos e valores não são crenças ou outros estados intencionais é certamente um absurdo.

Porém, a TCM deve fornecer alguma explicação da intencionalidade. Ou melhor, deve ser possível fornecer uma explicação da intencionalidade que seja consistente com a TCM (caso ela seja viável). Mas, se essa explicação não é a TRM, qual é ela então? Vejamos a seguir como Cummins introduz uma explicação alternativa da intencionalidade.

4.1.3.3 *Introduzindo a intencionalidade*

Cummins (1989, p. 141-145) introduz uma explicação da intencionalidade compatível com a Semântica Interpretativa. É possível combinar um relato causal ou adaptativo da intencionalidade com a TCM, de modo que a representação simulada somada a uma *restrição adicional* externa à TCM (como, por exemplo, contexto causal, adaptativo, social, etc.) resulte na intencionalidade. Desse modo: *conteúdo da representação simulada + restrição adicional = conteúdo intencional*.

Uma possibilidade é pensar na restrição adicional como um filtro que seleciona o conteúdo intencional de uma estrutura de dados dentre seus conteúdos de representações simuladas. Como as explicações da TCM não são intencionalmente especificadas e, portanto, não são especificadas em termos sensíveis à restrição adicional, a TCM não impõe restrições substantivas à escolha da restrição adicional.

A fim de ilustrar essa ideia de filtro, Cummins (1989, p. 141-142) toma emprestado o exemplo da Terra Gêmea, de Putnam²², e pede que o leitor imagine um sistema S que simule uma série de funções H₂O (funções tendo H₂O como argumento ou valor). S-gêmeo, o gêmeo de S na Terra Gêmea, simulará apenas funções XYZ isomórficas às funções H₂O de S (caso contrário, não serão gêmeos).

Segue-se que qualquer simulação das funções H₂O de S será uma simulação das funções XYZ de S-gêmeo. Portanto, S simulará as funções XYZ de S-

²² Cf. Putnam (1975).

gêmeo e, portanto, representará XYZ (de forma simulada) tão bem quanto S-gêmeo. Mas S, ao contrário de S-gêmeo, tem relações de restrição adicional com H₂O e não com XYZ, por isso as crenças de S são sobre H₂O e não sobre XYZ. S intenciona H₂O, mas não XYZ, enquanto S-gêmeo intenciona XYZ, mas não H₂O.

Essa estratégia corresponde a uma teoria representacionista da intencionalidade na medida em que, como a TRM, implica que cada conteúdo intencional atribuível a S é um conteúdo de alguma representação de S ou S-gêmeo. O conteúdo de cada crença, por exemplo, é tido como um conteúdo de representação simulada privilegiado, ou seja, aquele dentre tantos outros numa determinada estrutura de dados capaz de satisfazer à restrição adicional.

Nesta visão, a estrutura de dados subjacente a uma crença não é uma crença. Na verdade, ela é uma espécie de contraparte que se torna uma crença quando se aplica sobre ela uma restrição adicional. Todavia, esclarece Cummins, não há razão *a priori* para supor que existam tais contrapartes. Os proponentes da TCM devem considerar a possibilidade de que o componente psicológico de uma crença particular não seja um estado “local” do sistema, mas um estado “global”.

Pode ser que uma crença esteja para a TCM como um ponto de vista está para um editorial: pode não haver nenhuma parte específica do editorial que expresse o seu ponto de vista; na verdade, o editorial como um todo é que faz isso. Além disso, inúmeros outros editoriais, diferentes nos detalhes e até mesmo no assunto, podem expressar o mesmo ponto de vista. E o contrário também é possível, isto é, mesmo que o estado psicológico de S como um todo suporte cada uma de suas crenças e desejos, um estado psicológico completamente diferente poderia suportar exatamente as mesmas crenças e desejos.

Enfim, segundo Cummins (1989, p. 144-145), a crítica à TCM quanto à questão da intencionalidade, ou seja, o problema de vincular os conteúdos das crenças às representações, está errada de, pelo menos, três maneiras.

Primeiro, porque se assume sem argumentação que algo como a TRM é verdadeiro, ou seja, que ter crenças envolve abrigar representações cujos conteúdos são equivalentes a essas crenças.

Segundo, porque se assume que o papel explicativo do conteúdo nas teorias empíricas da cognição (como é o caso da TCM) é apenas o papel explicativo do conteúdo mental na psicologia do senso comum sobre crenças e desejos, o que não é verdade.

Terceiro, porque se vincula de modo essencial a representação à cognição, impedindo uma das principais virtudes da TCM, ou seja, sua afirmação de que a representação tem o mesmo papel em explicar a cognição que tem em explicar as capacidades dos sistemas computacionais em geral.

Para Cummins, a solução é abordar o problema da representação não como um filósofo da mente, mas como um filósofo da ciência: primeiro, determinar qual papel explicativo a representação desempenha em alguma teoria científica específica ou arcabouço teórico que invoca a representação. Depois, perguntar o que a representação deve ser e como ela deve ser explicada para que possa desempenhar esse papel.

4.2 Avaliando a Semântica Interpretativa

A Semântica Interpretativa de Robert Cummins é sem dúvida uma abordagem interessante sobre a questão da representação e do conteúdo mental, mas, em minha opinião, ela apresenta alguns problemas que enfraquecem seu poder argumentativo.

Representações mentais são símbolos? Há duas possibilidades: (1) se representações mentais não são símbolos, então a TCM está errada, pois ela lida com símbolos; (2) se representações mentais são símbolos, então a TCM *pode* estar certa. Essa segunda possibilidade equivale à afirmação de que a verdade da TCM permanece sem garantia mesmo que representações mentais sejam de fato símbolos. A garantia não existe porque Cummins não prova que o cérebro é um computador". Se acreditamos que representações mentais são símbolos, por que deveríamos aceitar imediatamente a TCM sem antes excluir outras possibilidades? Para aceitar a TCM de modo imediato precisaríamos ou provar que a mente é um computador ou que não existem outras coisas que manipulam símbolos. Mas Cummins parte da TCM como um pressuposto.

Cummins tenta explicar a representação mental, mas não justifica sua explicação. É claro que sua explicação é competente e sólida, mas o que se quer explicar no fim das contas é como a representação mental funciona, e ao que me consta, as únicas mentes de que temos notícias são as mentes humanas. Pensemos por um instante na física. Do que é composta a matéria? De partículas ou cordas? Embora o modelo físico das cordas seja possível, atualmente temos mais

evidências em favor do modelo físico de partículas. Muito bem, quais (e quantas) evidências nós temos em favor do “modelo” computacional da mente? Talvez a TCM seja um modelo possível, mas Cummins não deixa claro o porquê ele é o mais provável.

Outro ponto a se considerar é que, mesmo que muitos filósofos tentem fugir do “psicologismo”, a investigação sobre a mente é uma questão complicada: onde estão as mentes quando observamos o funcionamento de cérebros em laboratório? As mentes têm um caráter privado que cria algumas dificuldades para o filósofo da mente. Ora, toda ciência tem um objeto de investigação. Portanto a filosofia da mente, enquanto pretende ser um conhecimento rigoroso, científico, tem um objeto de investigação: a mente. Mas, se os filósofos da mente investigam a mente através de ciência cognitiva, computação, etc., seu objeto de estudo não está presente. E ausente o objeto de estudo, a ciência ou conhecimento rigoroso que se quer construir se torna impossível.

Cummins afirma que sua Semântica Interpretativa é sobre a representação de um modo geral e não se confunde com a intencionalidade. Mas, por que ela é a única, a melhor ou a mais provável possibilidade sobre a representação? E mesmo que seja a melhor possibilidade disponível para explicar a representação de um modo geral, o que garante que a representação mental funciona dessa maneira?

Seria preciso mostrar o porquê ela é a melhor possibilidade de explicação e mostrar também que mentes representam exatamente assim. Mas, para mostrar algo que as mentes fazem, não ajuda muito mostrar coisas que computadores fazem. Penso ser mais provável que a intencionalidade fundamente a representação, e não o contrário, já que encontramos mais dificuldades quando tentamos explicar a representação sem a intencionalidade do que quando a incluímos no processo.

Conforme supracitado, Cummins afirma que a associação entre intencionalidade e representação deriva de uma assunção não argumentada de que a TRM é verdadeira (ou seja, ter crenças envolve abrigar representações cujos conteúdos são equivalentes a essas crenças). Mas isso não faz sentido, pois quando tenho uma crença, evidentemente, creio em algo e este algo está expresso no conteúdo do estado mental “crença”. Afinal, toda crença é sobre algo.

Sobre a função de interpretação necessária à representação simulada da

Semântica Interpretativa, vimos que ela pode variar e pode até funcionar corretamente para *alguns* casos, mesmo quando a função escolhida é inadequada para *todos* os casos. Lembremos o exemplo da multiplicação e da adição. Uma função de interpretação que mapeia para uma função de adição quando deveria mapear para uma função de multiplicação funcionará para todos os casos em que $x+y = x \times y$. Pois bem, qual é o critério que determina qual função interpretativa será instanciada quando S representa x? Parece-me que Cummins não deixa isso claro.

Afinal, tal função é estabelecida por um observador externo? Como uma função pode ser independente de alguém que determina o que aponta para o quê? Em que medida esta perspectiva não sofre das dificuldades apontadas por Searle (1980, p. 417-420) no *argumento do quarto chinês*? Examinemos essas questões.

O *argumento do quarto chinês* é um experimento mental projetado para questionar a capacidade da inteligência artificial de entender e criar significado. O argumento parte da premissa de que um computador pode ser programado para responder a perguntas em chinês de maneira tão convincente que pareça entender a língua, mesmo que na verdade não compreenda nada do idioma. Para ilustrar o experimento, Searle pede ao leitor que imagine o seguinte cenário:

Suponha que alguém esteja dentro de um quarto, e essa pessoa não fala uma palavra de chinês. No entanto, ela possui um conjunto de regras detalhadas em sua língua nativa (por exemplo, inglês) que lhe permitem receber perguntas em chinês e responder a elas corretamente usando o mesmo conjunto de regras.

Do lado de fora do quarto, há um indivíduo chinês que passa perguntas em chinês para o ocupante do quarto. Esse ocupante, seguindo as regras que possui, examina os símbolos e manipula-os de acordo com as instruções, devolvendo outros símbolos para quem está do lado de fora sem realmente entender o significado dos símbolos ou palavras em chinês.

A pessoa que faz as perguntas do lado de fora do quarto não tem conhecimento do fato de que o ocupante do quarto não compreende chinês e, portanto, acredita que a pessoa dentro do quarto realmente entende e responde às perguntas em chinês de forma inteligente, pois os símbolos que ela recebe são as respostas para as suas perguntas.

O argumento de Searle é que, apesar de o ocupante do quarto ser capaz de responder a perguntas em chinês de maneira indistinguível de uma pessoa que realmente compreende o idioma, não há compreensão real por parte do indivíduo

dentro do quarto, pois este apenas segue regras e manipula símbolos sem sequer entender o idioma chinês.

Aplicando o argumento de Searle à Semântica Interpretativa, temos o agente cognitivo (S) no papel do indivíduo dentro do quarto chinês, e a função de interpretação (F_i) como o conjunto de regras, procedimentos, programas, etc., que S precisa executar para responder a uma pergunta. As entradas (perguntas) seriam os dados da percepção e as saídas (respostas) seriam as representações simuladas (R_s).

Seguindo o argumento temos como consequência: (1) S não compreende R_s ; (2) S precisa compreender o procedimento a ser executado sobre determinada entrada para produzir uma saída adequada, o que, *mutatis mutandis*, significa que S precisa entender F_i . A consequência direta de 1 é que se S tem uma R_s acerca de um objeto O, então R_s não pode significar O para S, pois S não é capaz de entender o significado de R_s . Quanto a 2, temos como consequência que S não pode gerar uma R_s , pois, isto depende da execução de F_i , que por sua vez depende da compreensão adequada do processo descrito por F_i , compreensão essa que S não tem (de acordo com a própria Semântica Interpretativa).

Além disso, se S instancia uma interpretação que mapeia entre “redondo e vermelho” e “maçã”, e S está em um ambiente cujas únicas coisas redondas e vermelhas são maçãs, S não poderá errar sobre as maçãs. Contudo, é óbvio que S não representa adequadamente uma maçã, pois redondo e vermelho é muito vago para determinar a espécie “maçã”. Portanto, se S não intenciona as maçãs, se a proposição “esta coisa redonda e vermelha é uma maçã” não é uma atitude proposicional de S, não poderíamos dizer que S representa maçãs, pois há muito mais coisas redondas e vermelhas além das maçãs.

Assim, parece-me que, embora a teoria de Cummins seja interessante e levante muitos questionamentos e ideias frutuosas sobre a representação, ela deixa em aberto algumas questões fundamentais sobre os conteúdos mentais e tangencia a explicação da natureza das representações mentais através de concepções não mentais da representação, como a computação e a simulação, sem com isso estabelecer de modo inequívoco a ligação entre esses dois campos.

5 CONCLUSÃO

A tese de Cummins, como dito antes, é que todas as teorias da representação mental por covariação podem ser criticadas a partir da teoria protótipo T_L . Porém, minha argumentação aponta para uma inadequação da descrição computacional LOCKE em virtude de sua incapacidade de descrever completamente os processos de percepção e de representação dos seres humanos. Nessa descrição computacional, Cummins simplifica demais o processo, abstraindo passos importantes para a correta compreensão da percepção e formação de representações e por isso suas conclusões não podem ser extrapoladas de LOCKE para os seres humanos.

Também argumentei contra a crítica de Cummins às ideias de Jerry Fodor quanto à representação mental. Em primeiro lugar, demonstramos que a análise dos contrafactuais C_1 e C_2 contém erros e que esses contrafactuais não são exaustivos, atributo que era essencial para a argumentação de Cummins. Em segundo lugar, explorei algumas possibilidades de explicação para os erros de representação ou deturpações: a insuficiência de dados e falhas cognitivas podem explicar as deturpações sem comprometer as teses de Fodor sobre a representação.

Por fim, argumentei sobre a falta de critério objetivo de alguns aspectos da Teoria Computacional da Mente e da Semântica Interpretativa que se pretendem explicações da representação mental e apontamos para o fato de que Cummins não foi capaz de resolver de modo satisfatório a questão da intencionalidade em sua própria teoria da representação.

Por isso, corrijo aqui minha hipótese inicial de trabalho, a saber, que a crítica de Cummins às teorias da representação por covariação são pertinentes e que sua proposta de solução para a questão da representação mental está correta. Concluo, portanto, que as críticas de Cummins não são pertinentes e que sua teoria sobre a representação mental, apesar de bem construída, deixa ainda algumas questões em aberto, principalmente acerca da objetividade da interpretação e da questão da intencionalidade. Assim, espero que este trabalho, apesar de suas limitações de escopo, possa contribuir positivamente para o debate acerca da representação mental.

REFERÊNCIAS

- ALLEN, K. Ideas. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 159-165.
- BENNETT, J. Locke's philosophy of mind. *In*: CHAPPELL, V. C. **The Cambridge companion to Locke**. Cambridge ; New York: Cambridge University Press, 1994. p. 89-114.
- BRANQUINHO, J; MURCHO, D; GOMES, N. **Enciclopédia de termos lógico-filosóficos**. São Paulo: Martins Fontes, 2006.
- CHAPPELL, V. C. Locke's theory of ideas. *In*: CHAPPELL, V. C. **The Cambridge companion to Locke**. Cambridge ; New York: Cambridge University Press, 1994. p. 26-55.
- CUMMINS, R. **Meaning and mental representation**. Cambridge: MIT Press, 1989.
- DAWSON, H. Meaning and signification. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 175-177.
- DAWSON, H. Names. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 184-186.
- DAWSON, H. Words, their imperfections and abuses. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 232-233.
- DENNETT, D. **Brainstorms**. Cambridge: MIT Press; A bradford Book. 1978.
- DOWNING, L. Locke and Descartes. *In*: STUART, M. **A companion to Locke**. Chichester, West Sussex, UK: Wiley, Blackwell, 2016. p. 100-120.
- FODOR, J. **Psychosemantics: the problem of meaning in the philosophy of mind**. Cambridge: MIT Press, 1987.
- GUYER, P. Locke's philosophy of language. *In*: CHAPPELL, V. C. **The Cambridge companion to Locke**. Cambridge ; New York: Cambridge University Press, 1994. p. 115-145.
- HERMIDA LAZCANO, P. La psicosemántica de Jerry A. Fodor. **ENDOXA**, [S. l.], v. 1, n. 1, p. 359–372, 1993.
- HILL, J. Boyle, Robert. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York:

Continuum, 2010. p. 47-51.

HILL, J. Primary and Secondary Qualities. *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 199-201.

LENNON, T. M. Locke on Ideas and Representation. *In*: NEWMAN, L. **The Cambridge Companion to Locke's "Essay Concerning Human Understanding"**. New York: Cambridge University Press, 2007. p. 231-257.

LEVIN, J. Functionalism. **The Stanford Encyclopedia of Philosophy (Winter 2021 Edition)**. Stanford: Metaphysics Research Lab, Stanford University. 2021. Disponível em: <https://plato.stanford.edu/archives/win2021/entries/functionalism>. Acesso em 13 out. 2022.

LOCKE, J. **An Essay Concerning Human Understanding**. Cambridge: Hackett, 1996.

LOSONSKY, M. Language, Meaning and Mind in Locke's Essay. *In*: NEWMAN, L. **The Cambridge Companion to Locke's "Essay Concerning Human Understanding"**. New York: Cambridge University Press, 2007. p. 286-312.

LOWE, E. J. Language and Meaning. *In*: STUART, M. **A companion to Locke**. Chichester, West Sussex, UK: Wiley, Blackwell, 2016. p. 297-295.

NEWMAN, L. Locke on Knowledge. *In*: NEWMAN, L. **The Cambridge Companion to Locke's "Essay Concerning Human Understanding"**. New York: Cambridge University Press, 2007. p. 313-351.

PUTNAM, H. The Meaning of Meaning. *In*: **Language, Mind and Knowledge**. Minneapolis: University of Minnesota Press. 1975.

SCHUURMAN, P. Descartes, René (1596–1650). *In*: SAVONIUS-WROTH, S.-J.; WALMSLEY, J.; SCHUURMAN, P. **The Continuum companion to Locke**. London ; New York: Continuum, 2010. p. 53-57.

SEARLE, J. Minds, Brains and Programs. **Behavioral and Brain Sciences**, v 3. n. 3, p. 417-457, 1980.

SOLES, D. The Theory of Ideas. *In*: STUART, M. **A companion to Locke**. Chichester, West Sussex, UK: Wiley, Blackwell, 2016. p. 140-155.

WILSON, R. A. Primary and Secondary Qualities. *In*: STUART, M. **A companion to Locke**. Chichester, West Sussex, UK: Wiley, Blackwell, 2016. p. 193-211.

WOOLHOUSE, R. Locke's theory of knowledge. *In*: CHAPPELL, V. C. **The Cambridge companion to Locke**. Cambridge ; New York: Cambridge University Press, 1994. p. 146-171.