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

Guilherme Sanches de Oliveira

Estados Mentais e Atitudes Proposicionais: Abordagens Filosóficas da Psicologia do Senso Comum

São Paulo
2014

Guilherme Sanches de Oliveira

Estados Mentais e Atitudes Proposicionais: Abordagens Filosóficas da Psicologia do Senso Comum

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

São Paulo
2014

Agradecimentos

Seria irônico se uma pesquisa sobre o entendimento interpessoal e a cognição social fosse realizada de maneira isolada e solitária. A verdade é que esta dissertação é fruto de anos de interações com diversas pessoas às quais sou grato:

Ao prof. Osvaldo Pessoa Jr., meu orientador, pelas discussões ainda antes de eu entrar no mestrado, bem como por sua flexibilidade e pela liberdade que me deu para realizar esta pesquisa;

Ao prof. Caetano Ernesto Plastino, cujas aulas ainda na graduação despertaram meu interesse em questões sobre mente, comportamento, racionalidade, conhecimento e ciência na filosofia analítica;

Ao prof. José Lino Oliveira Bueno, pelo acolhimento na Psicobiologia-FFCLRP de um intruso interessado em aprender sobre aprendizagem;

A estes três agradeço também pelos comentários e sugestões no exame de qualificação, que ajudaram a guiar a conclusão deste projeto.

Ao prof. Neil Van Leeuwen, agradeço pelo estágio de pesquisa no Departamento de Filosofia da Georgia State University, e à profa. Nancy Nersessian, por me receber em suas aulas de ciências cognitivas no Georgia Institute of Technology.

Partes desta pesquisa foram apresentadas e discutidas em eventos realizados na UNESP-Marília, no Georgia Institute of Technology, na Vanderbilt University, além de encontros da Sociedad Filosofica del Uruguay e da Central States Philosophical Association, e de um curso de curta duração com o prof. Vittorio Gallese e a profa. Frédérique de Vignemont realizado na Berlin School of Mind and Brain (Humboldt-Universität zu Berlin). Aos participantes desses encontros e conferências agradeço pelos comentários, críticas e sugestões, que contribuíram para o desenvolvimento de minhas ideias.

Pelo apoio financeiro que viabilizou minha dedicação a este projeto, agradeço à CAPES (agosto de 2011 a fevereiro de 2012) e à FAPESP (março de 2012 a julho de 2013).

Agradeço, por fim, à minha mãe, pelo incentivo e sacrifícios que me permitiram chegar até aqui, e à minha esposa pelas memórias dos anos já passados e pelos sonhos para os tantos anos que ainda virão.

RESUMO

Oliveira, G. S. Estados mentais e atitudes proposicionais: Abordagens filosóficas da psicologia do senso comum. 2014. 144 f. Dissertação (Mestrado) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia. Universidade de São Paulo, São Paulo, 2014.

A literatura filosófica sobre a Psicologia do Senso Comum se estende desde a década de 1970, e abrange diversas questões sobre nosso entendimento interpessoal cotidiano, nossa capacidade de interação e coordenação de atividades, o arcabouço conceitual intuitivo que relaciona estados mentais e atitudes proposicionais a comportamentos, e os mecanismos cognitivos de "leitura mental" que nos permitem atribuir estados mentais a outras pessoas. Nesta dissertação eu examino o desenvolvimento histórico desta literatura, identificando dois debates distintos, o primeiro (principalmente entre Paul Churchland e Jerry Fodor dos anos 70 aos anos 90) tendo como foco a relação entre a "teoria da Psicologia do Senso Comum" e teorias científicas (da neurociência e das ciências cognitivas), e o segundo (o debate contemporâneo) tendo como foco os mecanismos cognitivos de leitura mental e o papel das atribuições de estados mentais e atitudes proposicionais nas teorias da cognição corporificada, situada e estendida. Além do exame histórico do que argumento serem dois debates distintos e da transição conceitual entre ambos, também apresento aqui minha crítica à abordagem eliminativista contemporânea de Matthew Ratcliffe e, como alternativa, articulo os princípios de uma abordagem pluralista que combina leitura mental e interpretação contextual situada como fundamentais para a cognição social.

Palavras-chave: psicologia do senso comum, leitura mental, atribuição de estados mentais, atitudes proposicionais, eliminativismo, cognição social, cognição corporificada, cognição situada, cognição estendida.

ABSTRACT

Oliveira, G. S. Mental states and propositional attitudes: philosophical approaches to folk psychology. 2014. 144 p. Thesis (Master's Degree) – Faculdade de Filosofia, Letras e Ciências Humanas. Departamento de Filosofia. Universidade de São Paulo, São Paulo, 2014.

The philosophical literature on Folk Psychology began in the 1970s, and encompasses various questions about our everyday interpersonal understanding, our ability to interact and coordinate activities, the intuitive conceptual framework that relates mental states and propositional attitudes to behaviors, and the cognitive mechanisms of "mindreading" that allow us to attribute mental states to other people. In this thesis I examine the historical development of this literature, identifying two distinct debates, the first (mainly between Paul Churchland and Jerry Fodor from the 70s to the 90s) focusing on the relationship between the "theory of Folk Psychology" and scientific theories (in neuroscience and cognitive science), and the second (the contemporary debate) focusing on the cognitive mechanisms of mindreading and the role played by attributions of mental states and propositional attitudes in theories of embodied, situated and extended cognition. In addition to the historical examination of what I argue are two distinct debates as well as of the conceptual transition between them, here I present my criticism of Matthew Ratcliffe's contemporary eliminativist approach and, as an alternative to it, I articulate the principles of a pluralistic approach that combines both mindreading and situated contextual interpretation as fundamental for social cognition.

Key Words: folk psychology, mindreading, mental state attributions, propositional attitudes, eliminativism, social cognition, embodied cognition, situated cognition, extended cognition.

SUMÁRIO

INTRODUÇÃO 8

CAPÍTULO 1: Contexto Intelectual 17

1.1 Da Psicologia Experimental às Ciências Cognitivas 18

1.2 A Tradição Analítica na Filosofia 25

1.3 Estados Mentais e Atitudes Proposicionais 34

CAPÍTULO 2: O Debate Churchland-Fodor 39

2.1 O Materialismo Eliminativo e a "Defesa" da PSC 40

2.2 Explicações Ordinárias de Comportamentos 43

2.3 A Avaliação da Psicologia do Senso Comum 49

2.3.1 Sucesso Preditivo e Explicativo 49

2.3.2 Dificuldades e Estagnação 52

2.4 Relações Interteóricas e a Unidade da Ciência 56

2.5 Alternativas Hipotéticas à Psicologia do Senso Comum 61

2.6 Cognição, Computação, e a Vindicação da PSC 65

2.7 Conclusão 68

CAPÍTULO 3: Transição Conceitual 71

3.1 Tarefas de Crença Falsa 71

3.2 Mecanismos de Leitura Mental 73

3.2.1 A "Teoria da Teoria" 74

3.2.2 A "Teoria da Simulação" 75

3.3 Stephen Stich e o "Despertar de um Sono Dogmático" 79

3.3.1 Fase Eliminativista 80

3.3.2 A “Teoria” da PSC: Internalismo e Externalismo 81

3.3.3 A Teoria Híbrida 85

3.4 Avanços nas Ciências Cognitivas 89

3.4.1 “Para baixo” e “Para fora” 90

3.4.2 Cognição Corporificada, Situada e Estendida 93

CAPÍTULO 4: O Debate Contemporâneo 99

4.1 O Eliminativismo de Matthew Ratcliffe 100

4.2 Críticas 107

4.2.1 Contraexemplos 108

4.2.2 Racionalidade 114

4.2.3 A Linguagem como Extensão Cognitiva 120

4.3 Princípios de uma Abordagem Pluralista da Cognição Social 124

CONCLUSÃO 132

REFERÊNCIAS 136

INTRODUÇÃO

Em nossa experiência cotidiana nós interagimos com um vasto número de pessoas, de completos estranhos a familiares e amigos íntimos, e em tais encontros intersubjetivos frequentemente coordenamos atividades conjuntas, negociamos projetos conflitantes, influenciamos o comportamento das pessoas ao nosso redor e tomamos decisões com base no que os outros fazem ou dizem, ou até mesmo no que supomos que eles farão ou dirão. Em um grande centro urbano como São Paulo esse aspecto social da vida toma proporções ainda maiores, e nossas atividades ordinariamente coincidem com as de outras centenas, milhares ou milhões de pessoas que encontramos no trânsito, no local de trabalho, em momentos de lazer, etc.

Em cada uma dessas situações estamos cercados de pessoas que participam de nossas atividades direta ou indiretamente, e de maneiras distintas: em um restaurante, por exemplo, além dos amigos que nos acompanham, figuram na cena o garçom que anota nosso pedido, as pessoas sentadas à mesa ao lado, o cozinheiro que prepara a refeição, a banda tocando música ao vivo, e assim por diante. A conversa com nossos amigos pode ser sobre um assunto que nunca tenhamos discutido anteriormente, mas ela flui sem que precisemos ensaiar ou deliberar de maneira explícita sobre qual rumo queremos que a discussão tome, quem defenderá qual opinião, etc. E muito embora nunca antes tenhamos visto e nem tornemos a encontrar as outras pessoas ao nosso redor, situações como esta transcorrem espontaneamente, e com naturalidade conseguimos explicar e até mesmo prever com certo sucesso o comportamento de estranhos: entendemos, por exemplo, que as duas pessoas da mesa vizinha à nossa estão de mãos dadas porque são um casal romântico, percebemos que a mulher gesticulando

energicamente na mesa seguinte está dando uma bronca nos seus filhos que não param quietos (os quais podemos supor que irão chorar em decorrência da repreensão), e por fim assinamos no ar com uma caneta imaginária prevendo que isso levará o garçom a trazer nossa conta.

É notável que a grande maioria de nossos encontros intersubjetivos sejam caracterizados por entendimento mútuo, e isso de maneira aparentemente intuitiva e natural. Isso, porém, não significa dizer que sempre concordamos com o comportamento das pessoas ao nosso redor, o que é claramente falso. Mas, em um sentido mais básico, mesmo o que chamamos de "conflito" envolve um mínimo de compreensão: uma disputa originada em uma diferença de opiniões seria impossível se completamente destituída de inteligibilidade; da mesma forma, em casos de agressão e violência, como num assalto, se nos vemos sem saber como reagir, não é porque não sabemos o que está acontecendo, mas, antes, exatamente porque sabemos – isto é, porque compreendemos a periculosidade da situação. Raras são as instâncias em que as ações dos outros nos parecem de fato completamente estranhas e alheias; pelo contrário, mesmo se julgamos o comportamento de alguém como "incompreensível", muitas vezes o que queremos dizer com isso é que consideramos tal ação reprovável ou injustificável, mas não literalmente enigmática e misteriosa como se fosse de um alienígena.

Mas como explicar esse nosso entendimento interpessoal básico, nossa habilidade de interagir espontaneamente com outras pessoas, coordenar atividades, entender e até mesmo prever o comportamento dos outros, não apenas de conhecidos mas também de completos estranhos? Na segunda metade do século XX, principalmente a partir dos anos 70, surge entre filósofos da tradição analítica uma rica literatura que

persiste até hoje e que trata desta questão como o que convencionou-se chamar de "Psicologia do Senso Comum"[1]. No centro desta discussão estava, já de início, a percepção de que nosso discurso cotidiano está repleto de termos que representam estados mentais como *desejos*, *crenças*, *intenções* e *emoções*: dizemos que temos *vontade* de fazer alguma coisa ou que *desejamos* ou *queremos* algo; dizemos também que alguém tem uma determinada *crença* ou *opinião*, ou ainda que essa pessoa *acredita* ou *acha* algo ou que *pensa* de uma determinada maneira; dizemos, ainda, que temos a *intenção* de fazer alguma coisa, ou que *pretendemos* ou *planejamos* fazer algo; e além de muitas outras palavras como estas, com suas diversas funções sintáticas (e.g., substantivo ou verbo), temos ainda uma variedade de termos para falar de emoções e sentimentos. A partir de observações de nosso uso corriqueiro desta terminologia mentalista, filósofos sugeriram que nossa capacidade de interagir com outras pessoas e explicar e prever seu comportamento baseia-se numa *teoria* ou *arcabouço conceitual* intuitivo que estabelece relações nômicas entre eventos objetivamente observáveis e estados mentais subjetivos que atribuímos aos outros. Providos desta teoria, apontamos para o comportamento visível de um agente e o interpretamos como fruto de um estado mental que não vemos, como as suas crenças, desejos, intenções, ou emoções; ao mesmo tempo, e inversamente, com base no que sabemos sobre os estados mentais daquele agente, esboçamos explicações sobre suas ações passadas e até prevemos alguns de seus comportamentos futuros.

As primeiras questões filosóficas de destaque nesta literatura diziam respeito à relação entre esse arcabouço conceitual mentalista intuitivo e teorias científicas acerca da mente e comportamento humanos. Paul Churchland e Jerry Fodor, dois dos

[1] Comentário sobre a tradução do termo para o português mais adiante (p. 14).

principais filósofos nesta discussão, defendiam interpretações notoriamente divergentes, Churchland favorecendo uma perspectiva materialista que propunha a eliminação de nossa terminologia mentalista ordinária em favor de descrições de padrões de ativamento neural, e Fodor sustentando que o arcabouço conceitual mentalista da Psicologia do Senso Comum é não somente indispensável como também efetivamente compatível às melhores teorias científicas sobre processos cognitivos. A discordância, portanto, abrangia questões metafísicas sobre o status ontológico dos estados mentais a que nos referimos ordinariamente, estendendo-se também a questões epistemológicas e programáticas sobre a relação entre esse arcabouço conceitual intuitivo e o conhecimento científico, bem como sobre a relação dos vários campos das ciências entre si (principalmente neurociência, psicologia e ciências cognitivas).

A partir dos anos 80 e 90, desenvolvimentos conceituais e experimentais na filosofia e nas ciências em geral revelaram limitações importantes nesta discussão sobre a Psicologia do Senso Comum. Um problema central era o que parecia ser uma questão de imprecisão terminológica. Por um lado, o termo "Psicologia do Senso Comum" era utilizado para designar nossa habilidade de interação, coordenação e entendimento interpessoal, ou seja, nossa compreensão social intuitiva (o que pesquisadores nas ciências cognitivas viriam a chamar de "cognição social"); por outro lado, porém, "Psicologia do Senso Comum" também era empregado como designando a própria terminologia mentalista sobre desejos, crenças, intenções e emoções que utilizamos no dia a dia. Como ficou claro neste período, tal ambiguidade no uso do termo "Psicologia do Senso Comum" partia de uma pressuposição dupla, a qual assumia, primeiro, que nossa habilidade de interação, coordenação e entendimento interpessoal (o primeiro sentido de "Psicologia do Senso Comum" listado acima, como compreensão social)

resulta de nossa capacidade de "ler" a mente das outras pessoas e assim atribuir estados mentais a elas, e, segundo, que essa capacidade de "leitura mental" (ou de atribuir estados mentais aos outros) baseia-se no uso de uma teoria implícita que estabelece relações nômicas entre estados mentais e padrões comportamentais (ou seja, o segundo sentido de "Psicologia do Senso Comum" listado acima, como arcabouço conceitual). Embora até então assumidos como ponto de partida na literatura da Psicologia do Senso Comum, o fato é que tais pressupostos eram hipóteses passíveis de corroboração ou não. E a identificação dos mesmos como tal permitiu a distinção dos fenômenos a serem explicados, ampliando assim o leque de questões relevantes na literatura para além daquelas discutidas inicialmente por filósofos como Churchland e Fodor.

O questionamento da segunda hipótese, de que atribuímos estados mentais a outras pessoas por meio do uso de uma *teoria* (hipótese esta convenientemente chamada de "*Teoria da Teoria*"), levou primeiro ao desenvolvimento de uma proposta alternativa segundo a qual atribuímos estados mentais aos outros por meio de processos em que, conscientemente ou não, *simulamos* estar em suas circunstâncias (por isso chamada de "*Teoria da Simulação*"), e posteriormente motivou a formulação de uma proposta conciliatória que combinava aspectos de teoria e de simulação em uma "*Teoria Híbrida*". Mais fundamental ainda, no entanto, é o questionamento do primeiro pressuposto, de que o que explica nossa sociabilidade e entendimento interpessoal básico é nossa suposta capacidade de "leitura mental". Este último questionamento está no centro da discussão contemporânea acerca da Psicologia do Senso Comum, e acompanha a influência de teorias recentes que visam estudar a cognição humana como envolvendo processos corpóreos, culturalmente e fisicamente situados, e funcionalmente estendidos ou distribuídos. Em vez de se ocupar com os mecanismos

que viabilizam a leitura mental, esta nova abordagem põe em cheque a própria ideia de que precisamos atribuir estados mentais a outras pessoas para interagir com elas, coordenar atividades, e entender e prever seu comportamento; antes, como defendem alguns contemporâneos, aspectos da própria situação em que estamos inseridos bastam para modular interações e permitir a compreensão mútua, tornando a "leitura mental" pouco informativa, e por isso desnecessária e incomum ou mesmo inexistente.

Nesta dissertação proponho, por um lado, um exame do desenvolvimento histórico desta literatura, e, por outro lado, uma análise crítica do debate contemporâneo em que considero também direções para o desenvolvimento futuro. A leitura da bibliografia relevante no decorrer desta pesquisa motivou a identificação de dois debates distintos, nos quais usos diferentes do termo "Psicologia do Senso Comum" dotaram de significados diferentes as ideias de "defender" e "rejeitar" a Psicologia do Senso Comum, de forma que "rejeitar" a PSC no debate contemporâneo significa algo completamente diferente do que Churchland propôs no período inicial, assim como "defender" a PSC hoje pode ser inteiramente contraditório às ideias de Fodor.

Os quatro capítulos da dissertação refletem essa divisão temática identificada na literatura. O primeiro capítulo tem como foco o **contexto histórico conceitual** (científico e filosófico) que motiva a discussão e define os problemas centrais e os termos relevantes (como é o caso de "*atitudes proposicionais*", utilizado para designar um subtipo de estados mentais, principalmente crenças e desejos, que podem ser descritos como referentes a uma proposição qualquer X). À luz desse contexto, no segundo capítulo examino o **Debate Churchland-Fodor**, caracterizando as posições de ambos filósofos neste primeiro momento – a saber, o eliminativismo de Churchland e a

defesa da PSC proposta por Fodor. No terceiro capítulo, examino a **Transição Conceitual** provocada por inovações teóricas e experimentais principalmente a partir dos anos 80 e 90, e sua influência na redefinição dos problemas centrais da literatura, justificando assim a caracterização da discussão posterior como mais do que uma continuação das questões anteriores, mas constituindo efetivamente um novo debate. Por fim, o quarto capítulo trata do **Debate Contemporâneo**, no qual analiso a abordagem eliminativista de Matthew Ratcliffe e, em uma crítica à sua proposta, articulo argumentos em defesa da PSC – isto, porém, em direção a uma abordagem pluralista, muito diferente da "defesa" proposta por Fodor. Como o título do trabalho indica, o fio condutor desta dissertação será o papel de estados mentais (e, particularmente, atitudes proposicionais como crenças e desejos) na Psicologia do Senso Comum – ou, mais precisamente, a diversidade de questões que eles suscitam nas várias abordagens filosóficas propostas nos dois debates que caracterizam a literatura.

* * *

Como já indicado, o objeto de estudo desta pesquisa não é simplesmente a Psicologia do Senso Comum como fenômeno da vida social humana, mas principalmente como alvo de sistematização teórica por parte de filósofos da tradição analítica, em uma literatura específica que surge em países de língua inglesa a partir da década de 1970. Dada essa particularidade, esclarecimentos conceituais se fazem necessários já na tradução de termos relevantes.

Um ponto a considerar é a própria opção por "Psicologia do Senso Comum" como tradução do inglês "*Folk Psychology*". O termo "folk psychology" já é de uso

consagrado entre filósofos de língua inglesa, mas tem gerado diferentes traduções em outros idiomas. Uma opção mais literalista seria a tradução "Psicologia Popular", mas esta não parece ser uma boa escolha: primeiro porque o termo poderia ser confundido com teorias psicológicas como as discutidas em obras de divulgação científica ou até em obras de autoajuda, onde "popular" poderia ser interpretado como "pseudocientífico" ou ainda como "popularesco" e "vulgar"; e segundo porque esta opção poderia dar a impressão de que se trata de uma teoria psicológica que possui muitos adeptos, e que, dentre outras teorias, ganhou um concurso de popularidade; nestas duas interpretações de "popular", o termo modifica "psicologia" de um modo que remete a ideias alheias ao sentido pretendido na literatura filosófica.

"Psicologia do Senso Comum", a tradução aqui adotada, também não está imune a críticas, e levanta a questão: o que exatamente é esse "senso comum" – uma propriedade universal, inata, de fato *comum* a todas as pessoas, independentemente de seu contexto cultural, social, etc? Admitir a existência do senso comum nesse sentido estrito, como algo tão robusto e tangível, cria a dificuldade de se explicar não apenas sua natureza mas também sua origem e conteúdo. Ainda assim, esta tradução tem as suas vantagens, e é a que adoto aqui por considerá-la satisfatória se entendermos "senso comum" em sentido lato, simplesmente como denotando uma concepção ordinária, intuitiva, não apurada. Enquanto "pseudocientífico" denota teorias de caráter falsamente científico, "protocientífico" indica esboços de teorias a serem desenvolvidas e confirmadas (ou desconfirmadas) cientificamente. Como ficará claro mais adiante, Churchland e Fodor defendem que a PSC é uma teoria protocientífica. Em contraste com essa ideia, que eu rejeito, a tradução "Psicologia do Senso Comum" deixa esta questão em aberto, permitindo também uma aproximação à noção mais básica de "pré-

científico”, entendido como algo intuitivo e irrefletido, e assim potencialmente diferente do juízo “científico”, mas não por isso sendo “protocientífico” porque não necessariamente visa à corroboração científica, nem “pseudocientífico” ou “popularesco” nos sentidos apontados acima.

CAPÍTULO 1:

Contexto Intelectual

Em uma leitura superficial, o debate filosófico acerca da Psicologia do Senso Comum poderia ser visto como consistindo apenas em variações sobre as abordagens propostas pelo filósofo norte-americano Jerry Fodor e pelo filósofo canadense Paul Churchland, tamanho o impacto que suas obras tiveram. Entretanto, o exame cuidadoso da bibliografia relevante no decorrer desta pesquisa motivaram a caracterização deste como apenas o primeiro debate, em pontos importantes fundamentalmente distinto das discussões mais recentes. O terceiro capítulo desta dissertação examinará as mudanças conceituais que justificam esta divisão, e o quarto capítulo terá como foco as questões centrais do debate contemporâneo. Com o objetivo de permitir a apreciação dessas diferenças, neste primeiro capítulo farei uma breve apresentação do contexto intelectual em que o debate surge, para então no segundo capítulo caracterizar as propostas de Churchland e de Fodor bem como a natureza de sua discordância.

Questões sobre a mente, o comportamento e a sociabilidade humanas têm interessado pensadores desde a antiguidade, mas dados os limites intrínsecos a uma pesquisa de mestrado, me ocuparei aqui apenas com o ambiente intelectual imediatamente anterior ao surgimento do debate nas décadas de 1960 e 70, dando atenção a apenas algumas das ideias diretamente relevantes para a compreensão das abordagens de Churchland e de Fodor. Neste breve exame temático, considerarei desenvolvimentos nas ciências (na psicologia e nas ciências cognitivas) e na filosofia (na tradição analítica), novamente, tendo como foco compreensões teóricas sobre estados mentais, seu status ontológico e sua relação com o comportamento humano.

1.1: Da Psicologia Experimental às Ciências Cognitivas

A segunda metade do século XIX marca o surgimento da psicologia como disciplina independente. Questões que até então eram vistas como pertencendo ao domínio dos filósofos (e sua abordagem especulativa e apriorística) passaram a ser estudadas sob novas perspectivas e metodologias a fim de estabelecer um novo campo de conhecimento científico. Apenas na década de 1870, dois pioneiros iniciam seus próprios laboratórios experimentais, a saber, o alemão Wilhelm Wundt (1832-1920) e o norte-americano William James (1842-1910). Embora desenvolvendo teorias distintas, ambos coincidiam em privilegiar a introspecção como método de investigação da mente. Wundt reconhecia que na "auto-observação" o objeto observado e o sujeito observador são o mesmo, e que o direcionamento de atenção a um fenômeno mental inevitavelmente altera a experiência subjetiva do mesmo; assim, em seus estudos de "psicologia fisiológica" ele propunha que as condições controladas do laboratório podiam superar essa limitação, "*não abandonando a consciência a si mesma, mas forçando-a a responder às questões do experimentador ao colocá-la sob condições reguladas*" (Kim, 2006). Enquanto Wundt utilizava sujeitos treinados como "especialistas" em introspecção, James notoriamente lançava mão, em seus estudos sobre o "fluxo de consciência", de observações que ele próprio fazia de sua experiência subjetiva – o que alguns identificam como antecipando a tradição fenomenológica posterior (cf. Goodman, 2013).

Paralelamente, na Rússia, outro fisiólogo desenvolvia uma linha de pesquisa que viria a mudar o rumo que a psicologia tomava com os estudos de Wundt e James. Inicialmente interessado no estudo do sistema digestório, área pela qual receberia o

prêmio Nobel de fisiologia em 1904, Ivan Pavlov (1849-1936) é mais conhecido pelas descobertas que fez a respeito do que viria a ser chamado de "condicionamento clássico". Em seus famosos experimentos, Pavlov notou que a mera apresentação de comida naturalmente provocava a salivação de cães. Denominando a comida de *estímulo incondicionado* e a salivação de *resposta incondicionada*, Pavlov chamou a relação entre os dois de *reflexo incondicionado* (por ser inato, instintivo). Ele então passou a apresentar esse estímulo incondicionado acompanhado de um *estímulo neutro*, como o soar de um sino. Após algumas repetições desse emparelhamento, Pavlov observou que a apresentação do estímulo que inicialmente era neutro (e.g., o som do sino) bastava para gerar a resposta (salivação), e assim concluiu que estímulo e resposta haviam se tornado *condicionados*, e que o animal adquirira um *reflexo condicionado*, ou seja, aprendido (cf. Moreira e Medeiros, 2007).

O grande trunfo da abordagem de Pavlov era o uso de eventos publicamente *observáveis* e *mensuráveis*, permitindo assim a investigação e formulação de "*leis sobre efeitos de sequenciamento temporal e atraso, efeitos de intensidade de estímulo, e generalização de estímulo (abrindo as portas para a investigação experimental da percepção e discriminação animal)*" (Hauser, 2005). Esse modelo de investigação baseado em estímulos e respostas objetivos parecia tão promissor que foi logo trazido para o ocidente pelo psicólogo americano John Watson (1878–1958). Já em 1913, Watson afirma em um artigo que viria a ser considerado o "manifesto behaviorista":

> A psicologia como o behaviorista a vê é um ramo experimental puramente objetivo das ciências naturais. Seu objetivo teórico é a previsão e o controle do comportamento. A introspecção não constitui parte essencial de seus métodos, nem o valor científico de seus dados depende da facilidade com que eles podem ser interpretados em termos de consciência.
> (Watson, 2008)

Em experimentos inspirados nos de Pavlov, Watson descobriu que reações emocionais em seres humanos obedeciam o mesmo esquema de condicionamento, e, mais do que isso, apresentavam o fenômeno da *generalização respondente*, isto é, a obtenção de uma resposta condicionada até mesmo na apresentação de estímulos apenas *parecidos* com o estímulo condicionado (no caso do famoso experimento com o "pequeno Albert", Watson condicionou o bebê a chorar ao ver um rato branco por meio do emparelhamento deste estímulo neutro com um som estridente, e, surpreendentemente, essa resposta condicionada se estendeu a outros objetos parecidos, e.g. um coelho branco).

Essa abordagem experimental com foco apenas no comportamento observável teve seu auge na obra de B. F. Skinner (1904–90), expoente máximo do behaviorismo, e responsável por muitos desenvolvimentos práticos e conceituais. Para Watson, a psicologia tinha como foco estímulos e respostas, e o objetivo de "*prever, dado um estímulo, qual resposta ocorrerá; ou, dada uma resposta, estabelecer qual situação ou estímulo a causou*" (Watson, 1930, p. 11). Skinner, por outro lado, foi além do paradigma respondente de Watson, e introduziu o conceito de *comportamento operante* como aquele que "*produz consequências (modificações no ambiente) e é afetado por elas*", ou seja, "*comportamentos que são aprendidos em função de suas consequências*" (Moreira e Medeiros, 2007, pp. 47-48). Nessa nova abordagem, Skinner obteve grande sucesso experimental com o estabelecimento de *contingências de reforço*, isto é, por meio de "*alterações no ambiente* [que] *aumentam a probabilidade de o comportamento que as produziu voltar a ocorrer*" (Moreira e Medeiros, 2007, p. 51). Esse sucesso experimental foi em grande parte devido à invenção da caixa de *condicionamento operante* (a caixa experimental popularmente conhecida como "caixa de Skinner"), a

qual fornecia o ambiente controlado necessário para a medição precisa dos efeitos da modelagem do comportamento (e.g., a pressão da barra por um rato).

O filósofo George Graham identifica diferentes formas de behaviorismo na história da psicologia. Por um lado, tanto Watson quanto Skinner defendiam o que pode ser identificado como "behaviorismo metodológico", considerações normativas sobre a psicologia como área de investigação científica, segundo a qual "a *psicologia é a ciência do comportamento*", e não da mente (Graham, 2010). Ao mesmo tempo, ambos participavam do "behaviorismo psicológico", um programa de pesquisa dentro da psicologia que se propõe a "*explicar o comportamento humano e animal em termos de estímulos físicos externos, respostas, histórias de aprendizagem, e (para alguns tipos de comportamento) reforços*"; de acordo com essa posição, o comportamento "*pode ser descrito e explicado sem referências a eventos mentais ou a processos psicológicos internos*" e suas origens são "*externas (no ambiente), não internas (na mente, na cabeça)*" (Graham, 2010). Mas, embora coincidindo com Watson nesses dois aspectos (behaviorismo metodológico e behaviorismo psicológico), Skinner desenvolve no decorrer de sua carreira uma posição ainda mais extrema, que ele mesmo denomina "behaviorismo radical", segundo a qual termos mentais devem ser "*eliminados e substituídos por termos comportamentais, ou podem e devem ser traduzidos ou parafraseados como conceitos comportamentais*" (Graham, 2010).

Embora o behaviorismo tenha sido um programa de pesquisa muito frutífero e que permanece influente até hoje, a década de 1960 marca o início de seu declínio e o surgimento de alternativas. Se por um lado a rejeição da introspecção como método de investigação científica continuou sendo vista como um avanço, por outro lado, a

“revolução cognitiva” pôs em cheque a rejeição de estados mentais internos e a visão da psicologia como apenas ciência do comportamento. Ao menos em parte, a própria posição radical de Skinner pode ser apontada como responsável pelo “início do fim” do behaviorismo. Em seu livro “Verbal Behavior”, de 1957, Skinner deu o passo final na aplicação de seu paradigma operante ao comportamento humano mais comumente apontado como diferenciando seres humanos de outros animais, ou seja, a comunicação verbal. Mas esse projeto de Skinner não teve a mesma recepção que suas obras anteriores, e dois anos mais tarde o linguista Noam Chomsky publica uma resenha criticando severamente o livro, argumentando que o aparato explicativo behaviorista era incapaz de dar conta da “*competência linguística (em geral) e a aquisição da linguagem (em particular)*” (Hauser, 2005), para as quais Chomsky considerava necessário “*postular estruturas representacionais governadas por regras*” (Bechtel & Zawidzky, 1999, p. 753). A análise detalhada de Chomsky concluía que os termos 'estímulo', 'resposta', e 'reforço' são “*vagos*” e “*arbitrários*” (Chomsky, 1959, p. 44), e que “*a compreensão obtida nos laboratórios do teórico de reforçamento, embora genuína, só pode ser aplicada ao comportamento complexo humano de maneira grosseira e superficial*”; mais ainda, para Chomsky “*a magnitude da falha dessa tentativa de explicar o comportamento verbal serve de indicação da importância dos fatores omitidos*” (Chomsky, 1959, p. 28). Tais fatores omitidos, como ele articularia posteriormente, diziam respeito a uma competência linguística inata que explicaria “*como uma criança pode entender e criar sentenças que ela nunca ouviu antes*” (Russell & Norvig, 1995, p. 15), compensando a *pobreza de estímulo* a que somos expostos no curso do desenvolvimento (ou, nos termos skinnerianos, em nossa “história de reforçamento”):

> “as sentenças no ambiente de uma criança fornecem um banco de

> dados pobre demais para que seja crível que processos ordinários de aprendizagem deem conta da competência da criança; uma *Gramática Universal* inata deve guiar as induções que a criança faz a partir dos inputs que recebe" (Bechtel, Abrahamsen, & Graham, 1999, p. 41)

Mas, muito embora o trabalho de Chomsky na linguística tenha sido importante para a "revolução cognitiva", ele não pode ser apontado como único responsável por isso, sendo antes apenas parte de um movimento muito maior que já contava com avanços na neurociência bem como nas áreas da cibernética e inteligência artificial.

Por um lado, paralelamente a esses desenvolvimentos na psicologia, cientistas estavam desenvolvendo métodos cada vez mais precisos no estudo do cérebro. A segunda metade do século XIX marca a retomada da ideia de que o cérebro divide-se em áreas funcionalmente distintas (proposta por Franz Joseph Gall ainda no século XVIII), e também a descoberta de que o tecido nervoso é composto por células individuais (neurônios) separadas por pontos de comunicação (sinapses). Pesquisas como a de Paul Broca motivaram a "*inferência de funções de regiões específicas do cérebro* (…) *a partir de consequências comportamentais ou psicológicas da perda do tecido em questão (seja por causas clínicas ou por intervenção experimental deliberada)*" (Albright & Neville, 1999, p. lii), posteriormente levando à conclusão de que "*déficits neurológicos deveriam ser explicados tanto em termos de lesões na conexão entre áreas distintas quanto em termos de lesões das próprias regiões em questão*" (Arbib, 2003, p. 10). Mas além dessa abordagem a partir de lesões, avanços já em meados do século XX permitem o aperfeiçoamento de estudos baseados na estimulação elétrica, levando o neurocientista Wilder Penfield a descobrir que "*a estimulação de certas partes do córtex em pacientes humanos pode evocar memórias vívidas de experiências do passado*" (Bechtel & Zawidzki,1999, p. 768) e também que por meio da estimulação do córtex motor ele podia levar um paciente a movimentar seus

membros de maneira que o paciente tinha "*a percepção de seu braço se mover, mas* [não tinha] *a experiência de ação voluntária*" (Searle, 2004, p. 142).

Por outro lado, avanços tecnológicos na década de 1930 permitiram o desenvolvimento de máquina baseadas nas calculadoras mecânicas existentes já no século XIX, porém agora equipadas com componentes elétricos "*organizados para ligarem e desligarem uns aos outros e assim realizar operações aritméticas*" (Bechtel, Abrahamsen, & Graham, 1999, p. 9). A utilidade dessas máquinas para fins militares durante a Segunda Guerra Mundial, principalmente na quebra de mensagens criptografadas, motivou investimentos pesados que resultariam em ainda mais avanços em um curto período de tempo. Isso permitiu a implementação do que até então eram apenas possibilidades teóricas, como as operações lógicas sobre proposições, idealizadas por George Boole e por Gottlob Frege ainda na segunda metade do século XIX, ou a *máquina de Turing*, uma máquina abstrata, precursora do computador digital moderno, que demonstrava que "*qualquer operação sensível apenas à sintaxe pode ser duplicada* (...) *mecanicamente*" (Horst, 2009). Esses avanços tecnológicos obtidos nos anos 40 e 50 foram tamanhos que motivaram o desenvolvimento do campo da inteligência artificial, baseado, em sua forma mais extrema, na hipótese de que "*todo comportamento inteligente pode ser capturado por um sistema que raciocina logicamente a partir de um conjunto de fatos e regras descrevendo aquele domínio*" (Russell & Norvig, 1995, p. 827).

Conjuntamente, "*avanços na compreensão do cérebro contribuíram com o pensamento de pesquisadores a respeito de como conceitos como informação e computação podem fornecer uma base para a compreensão de processos mentais*"

(Bechtel, Abrahamsen, & Graham, 1999, p. 24). Já em 1943, por exemplo, McCulloch e Pitts propõem o primeiro "*modelo matemático relacionando computação e cérebro*" (Jordan & Russell, 1999, p. lxxiii) a partir da ideia de que computações simbólicas são implementadas na rede neuronal (Arbib, 2003, p. 11). No mesmo ano, Craik descreve o sistema nervoso como uma "*'máquina calculadora capaz de modelar* (…) *eventos externos'*", assim propondo que a "*formação de um 'modelo interno'* (…) *é a característica básica do pensamento*" (Arbib, 2003, p. 11). No contexto do declínio do behaviorismo na psicologia, estes avanços na neurociência e na ciência da computação foram fundamentais para o desenvolvimento da ciência cognitiva como área multidisciplinar de investigação a partir da década de 60, com muitos pesquisadores tendo como ponto de partida a hipótese de que "*a cognição humana é uma forma de computação* (…) *e poderia portanto ser emulada por computadores*" (Jordan & Russell, 1999, p. lxxiii), e a expectativa de que "*nós podemos estudar a mente, podemos descobrir como a mente realmente funciona, ao descobrir quais programas são implementados no cérebro*" (Searle, 2004, p. 69). Conforme ficará evidente na discussão sobre as abordagens de Jerry Fodor e de Paul Churchland, tais avanços e expectativas também exerceram enorme influência no debate filosófico acerca da Psicologia do Senso Comum.

1.2: A Tradição Analítica na Filosofia

Além do surgimento e desenvolvimento da psicologia como disciplina independente e de avanços na neurociência e na computação, o ambiente intelectual em que o debate sobre a Psicologia do Senso Comum emerge é profundamente marcado

pela herança recebida da própria tradição analítica na filosofia. Quebrando com as ambições sistematizantes das filosofias kantiana e hegeliana bem como com as práticas hermenêuticas que caracterizavam o estudo da história da filosofia, a filosofia analítica surge entre o fim do século XIX e o início do século XX como uma nova abordagem que favorecia o rigor e a clareza conceitual no tratamento de problemas filosóficos isolados e circunscritos. Foi com essa perspectiva que G. E. Moore (1873-1958) e Bertrand Russell (1872–1970), ambos professores em Cambridge, iniciaram a "virada linguística" na filosofia, "*baseando seus argumentos no 'significado' de termos e proposições*" com o objetivo de esclarecer as ilusões geradas pela linguagem natural (Preston, 2006). Já nessa fase inicial, Moore toma como ponto de partida uma filosofia realista do "senso comum", tendo como foco a análise da linguagem ordinária, enquanto Russell buscava "*re-expressar proposições na linguagem formal ideal da lógica simbólica, assim revelando sua forma lógica real*" (Preston, 2006). Inspirado no trabalho do alemão Gottlob Frege, Russell conta também com a ajuda de seu aluno Ludwig Wittgenstein (1889-1951) no desenvolvimento desse projeto, que viria a ser chamado de *atomismo lógico*. Segundo essa posição, "*o mundo consiste em uma pluralidade de entidades independentes e discretas, as quais, ao se combinarem, formam fatos*"; tais fatos são "atômicos" se não possuem subpartes, como no caso de "*um particular simples exibindo uma qualidade*" (Klement, 2013). Nessa visão, uma proposição corresponde a um estado de coisas atual ou possível, independente da concepção humana (no inglês, "*mind-independent*") ou de formulação linguística; a estrutura de um estado de coisas é isomórfica "*tanto à estrutura da proposição que se refere a* [esse estado de coisas] *quanto à estrutura da sentença que expressa tal proposição – desde que a sentença seja adequadamente formulada na notação da lógica simbólica*" (Preston, 2006).

As abordagens distintas de Moore e de Russell revezariam no papel de protagonistas da filosofia analítica nas décadas seguintes, influenciando dois movimentos que são particularmente relevantes para a compreensão da oposição entre Fodor e Churchland, a saber, o *positivismo lógico* na década de 1930 e a *filosofia da linguagem comum* na década de 50. Ambos movimentos também surgem à sombra da obra de Wittgenstein, sendo o positivismo lógico influenciado pelo "primeiro Wittgenstein" (do *Tractatus Logico-Philosophicus*, de 1921) e a filosofia da linguagem comum pelo "segundo Wittgenstein" (do livro *Investigações Filosóficas*, publicado postumamente, em 1953); mas um exame da herança wittgensteiniana em contraste com a influência de Moore ou de Russell foge do escopo desta dissertação, para a qual será preciso apenas caracterizar em linhas gerais os dois movimentos.

O *positivismo lógico* é geralmente identificado com o grupo de filósofos e cientistas que formavam o Círculo de Viena entre o fim dos anos 20 e início dos anos 30, mas a dispersão desses pensadores por ocasião da ascensão do regime nazista estende sua influência a outros continentes nas décadas posteriores[2]. Embora articulado de maneiras distintas tanto por adeptos quanto por críticos, o positivismo lógico tinha como preocupações centrais questões sobre epistemologia e metodologia na filosofia e nas ciências, adotando uma postura empirista, contrária a discussões metafísicas. Tendo como ponto de partida o formalismo privilegiado por Russell, um princípio típico do movimento (conhecido como "princípio da verificação") propunha que proposições não-tautológicas só são significativas se forem propriamente empíricas, isto é, se "*são 'sobre' o mundo, são 'factuais', possuem 'conteúdo' e seu valor de verdade é*

[2] Uebel (2011) e Creath (2011) fornecem boas descrições da história do Círculo de Viena, listas de seus "membros" e "simpatizantes", bem como referências a outras obras relevantes.

estritamente determinado pelo modo como o mundo é" e como se faz conhecer experimentalmente (Parker-Ryan, 2012); proposições que não podem ser reduzidas a proposições lógicas e que também não são passíveis de confirmação ou verificação a partir de observações empíricas não são *falsas*, mas simplesmente destituídas de significado.

Essa preocupação com a relação lógica entre proposições e a rejeição do não-empírico fica clara também em outra questão central para o positivismo lógico, a saber, a ideia da *unidade da ciência*, ou ainda, a "*unidade de método e linguagem que incluía todas as ciências, naturais e sociais*" (Cat, 2013). Essa unidade era caracterizada pela demarcação do científico (como é evidente no princípio da verificação), mas mesmo adeptos do positivismo lógico divergiam no que diz respeito a relações interteóricas e o requerimento de redução entre os diversos campos da ciência. Uma posição mais extrema, defendida por Rudolf Carnap, tinha a física como base da hierarquia das ciências, e propunha que sentenças de redução estabeleceriam ligações entre conceitos e leis de outras ciências com as leis genuínas e fundamentais da física; outra posição menos extrema, defendida por Otto Neurath, enfatizava o "*fisicalismo*" apenas como *linguagem do material*, e não simplesmente como o vocabulário da física (Cat, 2013). A diferença entre essas abordagens pode ser caracterizada como uma distinção entre "unidade interna", que diz respeito às "*relações em uma ciência ou entre ciências diferentes*", e "unidade externa", que diz respeito à distinção entre a ciência e o que não é empírico (e que por isso não é significativo) (Cat, 2013). Dada essa distinção, Carnap parece exigir ambos os tipos de unidade, enquanto Neurath tem o fisicalismo como base da unidade externa, sem requerer a redução das leis das ciências entre si.

O outro movimento na filosofia analítica que antecede aspectos importantes da discussão a respeito da Psicologia do Senso Comum é a *filosofia da linguagem comum*, muito identificada com filósofos de Oxford no fim da década de 1940 até os anos 60, e caracterizada pela retomada da análise linguística inaugurada por Moore, sem a ambição formalista russelliana presente no positivismo lógico. Essa recuperação do foco em problemas filosóficos a partir da análise da linguagem cotidiana deu-se em parte pela percepção de que o positivismo lógico era estrito demais, e que o princípio da verificação parecia contradizer a si próprio: uma vez que não estabelecia uma tautologia e nem era suscetível a confirmação por meio da observação empírica, o princípio mesmo não deve ser significativo; tentativas de resguardar o princípio dessa crítica inevitavelmente resultavam no enfraquecimento do mesmo, tal que ele não mais sucedia na rejeição categórica da metafísica (cf., por exemplo, discussões em Creath 2011, Uebel 2011, Preston 2006). Assim, para um grupo distinto de pensadores, a preocupação com a elucidação de problemas filosóficos por meio da dissolução de confusões linguísticas toma uma forma diferente da abordagem formalista prevalente até então, baseando-se antes na ideia de que "*nossa percepção ordinária do mundo, e nosso uso ordinário da linguagem para falar sobre ela, são tudo o que precisamos para dissipar perplexidades filosóficas*" (Parker-Ryan, 2012).

Uma obra de destaque nesse movimento é o livro *Intention*, publicado em 1957 pela filósofa britânica G. E. M. Anscombe (1919-2001), discípula do "segundo" Wittgenstein. Neste livro Anscombe propõe uma análise dos conceitos de voluntariedade, intenção, e ação intencional, articulando uma distinção entre *causas* e *razões* para ação, as quais são privilegiadas em diferentes descrições de um mesmo evento: descrições de uma ação como "intencional" ou "proposital" baseiam-se em

razões que podem ser dadas para tal ação, enquanto descrições da mesma ação como "não-intencional" ou "acidental" têm como foco causas mas não motivos que a pessoa poderia ter para realizá-la. Anscombe discute ainda dificuldades na descrição de eventos e de comportamentos cujas consequências se estendem temporalmente, bem como outras descrições que privilegiam enunciados de voluntariedade baseados em previsões prospectivas ou a partir de introspecções retrospectivas. É notável que, embora este seja considerado um clássico da teoria da ação, a abordagem da filosofia da linguagem comum permite a Anscombe se ocupar com essa elucidação conceitual pontual sem porém se preocupar em formular um sistema que aborde questões de outras áreas da filosofia.

Outra obra central neste período da filosofia da linguagem comum (e que exerce influência mais ampla que o livro de Anscombe) é *The Concept of Mind*, publicado pelo filósofo Gilbert Ryle em 1949. Neste livro Ryle critica o dualismo de substância cartesiano, que ele descreve como "o dogma do fantasma na máquina" por estabelecer que corpo e mente têm naturezas distintas, o primeiro sendo físico, material, e habitado pelo segundo, imaterial. Ryle denomina o dualismo cartesiano de "Doutrina Oficial" por causa da influência dessa ideia não apenas entre filósofos como também no pensamento popular. Seguindo a estratégia típica da filosofia da linguagem comum, sua rejeição da "Doutrina Oficial" não se dá pela construção de um complexo sistema metafísico alternativo ao cartesiano; antes, ele argumenta que uma análise linguística revela que o problema que o dualismo tenta resolver não passa de uma confusão, um erro de categoria pelo qual nossa linguagem nos leva a conceber operações e disposições como entidades. Daniel Dennett (2002) explica esse erro com uma analogia simples: ao ser notificada que o saldo de sua conta bancária está negativo, uma pessoa ingênua poderia

imaginar que o banco tem um cofre para cada um de seus clientes e que cada cofre contém notas de reais ao lado de notas de "anti-reais" (notas de valor negativo), e que naquele momento o cofre que corresponde à sua conta tem mais "anti-reais" do que reais. Mas conceber o saldo negativo como uma entidade substancial é um engano já que não existe dinheiro negativo; antes, saldos bancários constituem obrigações entre cliente e banco, que podem ser entendidas como disposições: o banco "*está disposto a honrar seus compromissos monetários até um certo limite, disposto a cobrar uma determinada taxa pela condição atual de sua conta,* (…) *etc., num grande sistema de condições* se-então *entrelaçadas*" (Dennett, 2002).

Analogamente a esse exemplo, Ryle examina o uso de palavras que "*utilizamos comumente para descrever e explicar o comportamento das pessoas*" e conclui que muitas delas "*significam disposições e não episódios*" (Ryle, 2009, p. 100), como é o caso de 'saber', 'acreditar', 'aspirar' e outras palavras que "*significam habilidades, tendências ou inclinações de fazer* (…) *coisas de vários tipos distintos*" (Ryle, 2009, p. 102). Uma pessoa que acredita que o gelo está fino, tende a avisar os outros disso, a patinar com cuidado, e a fornecer essa informação quando alguém lhe pergunta se é seguro patinar naquele lago (Ryle, 2009, p. 117); assim, para Ryle, dizer que essa pessoa acredita que o gelo está fino indica apenas o conjunto destas propensões comportamentais, e não alguma entidade (uma *crença*) na substância mental, o "fantasma na máquina". Da mesma maneira Ryle rejeita a ideia convencional de *conhecimento* como um estado mental monolítico, propondo antes uma distinção entre o conhecimento prático de como realizar certas atividades (em inglês, *knowledge how*) e o conhecimento abstrato ou conceitual de fatos (em inglês, *knowledge that*): uma pessoa que "sabe" andar de bicicleta simplesmente possui uma habilidade, é capaz de fazer

algo (andar de bicicleta); já uma pessoa que "sabe" que "*pedras são duras e esponjas são macias*" (Ryle, 2009, p. 48) tem uma série de inclinações comportamentais que podem incluir (mas não requerem) a formulação de tal afirmação de maneira explícita, linguística. Em ambos os casos, para Ryle, é um erro de categorização afirmar que qualquer dessas formas de "saber" constitui "conhecimento" como uma entidade abstrata armazenada mentalmente, sendo antes melhor entendidas como conjuntos de capacidades e disposições de agir de determinadas maneiras.

Ryle é comumente apontado como defendendo um tipo de *behaviorismo filosófico* ou *analítico* (cf. e.g. Graham 2010, Hauser 2005), mas vale notar que ele critica tanto a "*'redução' do mundo material a estados e processos mentais*" quanto a "*'redução' de estados e processos mentais a estados e processos físicos*" (Ryle, 2009, p. 12), propondo antes a dissipação do contraste entre Mente e Matéria que define tanto o dualismo quanto posições materialistas. Assim, Ryle não se compromete com a ideia de que atividades mentais não *existem* – "*fazer uma divisão longa é um processo mental assim como fazer uma piada*" (Ryle, 2009, p. 11) – mas apenas sugere que quando dizemos que "*existem mentes e corpos*" a noção de 'existência' pode implicar categorias distintas e que fogem da dicotomia pressuposta pela Doutrina Oficial, assim como é o caso quando dizemos que existem cães, macieiras, água, e o sol, em contraste com quando dizemos que "*existem números primos e quartas-feiras e opiniões públicas e marinhas*" (Ryle, 2009, p. 12).

O legado de Ryle se faz presente no debate contemporâneo acerca da Psicologia do Senso Comum principalmente na ênfase que alguns pensadores dão hoje à "*conexão intrínseca entre pensamento e ação*" e no "*anti-representacionismo, a ideia de que o*

pensamento não é o processamento de representações" (Chemero, 2002), mas esses pontos ficarão claros oportunamente (Capítulos 3 e 4). Por ora, é importante mencionar duas decorrências da abordagem de Ryle que influenciariam o debate entre Churchland e Fodor, a saber, a teoria da identidade e o funcionalismo.

Alguns filósofos e cientistas da época ficaram insatisfeitos com a interpretação ryleana de que predicados mentais designam propensões comportamentais, julgando-a incapaz de dar conta de alguns tipos de estados mentais como sensações, as quais podemos experimentar sem necessariamente envolver quaisquer reações comportamentais. A partir dessa insatisfação, a *teoria da identidade* é formulada no fim dos anos 50 pelo filósofo J. C. Smart e o psicólogo U. T. Place, propondo que sensações como a dor são idênticas a estados cerebrais; a influência dessa teoria entre filósofos materialistas leva a desenvolvimentos posteriores visando generalizar a identidade mente-cérebro a também outros estados mentais como crenças (Bechtel, Abrahamsen, & Graham, 1999, p. 62), e gerando a formulação de que "*cada tipo de estado mental é idêntico a um tipo particular de estado neural*" (Levin, 2013).

O funcionalismo surge nos anos 60, formulado pelo filósofo Hilary Putnam como uma reação à teoria da identidade. A teoria da identidade tinha como implicação que "*se estados mentais são idênticos a estados cerebrais, então estados mentais particulares limitam-se a organismos com cérebros similares*", o que ia de encontro à ideia popular no nascente campo das ciências cognitivas de que "*o mesmo tipo de estado ou condição cognitiva poderia ocorrer em cérebros ou sistemas nervosos distintos*", incluindo outros animais bem como possíveis mentes artificiais (Bechtel, Abrahamsen, & Graham, 1999, p. 62). Assim, o funcionalismo propõe que "*o que faz de*

algo um estado mental de um tipo particular não depende de sua constituição interna, mas antes da maneira como ele funciona, ou o papel que ele exerce, no sistema do qual faz parte", ou em outras palavras, que "*a identidade de um estado mental deve ser determinada por suas relações causais a estímulos sensoriais, outros estados mentais, e o comportamento*" (Levin, 2013). Embora possa parecer um retorno ao dualismo, o funcionalismo na verdade reconhece que estados mentais são realizados pelo cérebro, mas apenas defende que a identidade de um estado mental qualquer não é *definida* por fatos sobre essa implementação particular; assim, continua sendo possível que o mesmo estado mental (incluindo pensamentos, desejos, dores) seja implementado de maneiras estruturalmente distintas (Bechtel, Abrahamsen, & Graham, 1999, p. 63).

Essas duas teorias representam já a transição, na filosofia analítica, da filosofia da linguagem comum ao período de ecletismo que marca o fim do século XX, no qual diversas subáreas se desenvolvem, ampliando o leque de questões relevantes (e até revitalizando o interesse em questões metafísicas), mas preservando um certo estilo de argumentação típico dessa tradição, como o foco a questões pontuais (em contraste com a construção de grandes sistemas) e o privilégio do rigor e clareza na apresentação de conceitos (Preston, 2006).

1.3: Estados Mentais e Atitudes Proposicionais

Desta herança intelectual, tanto científica quanto filosófica, os termos "estado mental" e "atitude proposicional" são centrais para a discussão acerca da Psicologia do Senso Comum e merecem atenção especial. Embora em certos contextos específicos

“estado mental” possa ser utilizado como dizendo respeito ao estado de saúde mental de um indivíduo ou seu “estado de ânimo”, nesta literatura “estados mentais” compreendem de maneira abrangente processos cognitivos e predicados mentais, incluindo várias formas de pensamentos, sensações e emoções. Em contraste com esse uso amplo, “atitudes proposicionais” designam um tipo particular de estados mentais. A referência clássica como originando o termo é de Russell em seu artigo “The Philosophy of Logical Atomism” de 1918/1919:

> Que tipo de nome devemos dar a verbos como “acreditar” e “desejar” e assim por diante? Minha inclinação é chamá-los “verbos proposicionais”. Esse é apenas um nome sugerido por conveniência, porque eles são verbos que têm a forma de relacionar um objeto a uma proposição. Como eu tenho explicado, isto não é o que eles realmente fazem, mas é conveniente chamá-los verbos proposicionais. É claro, você pode chamá-los de “atitudes”, mas eu não gosto disso porque este é um termo psicológico, e embora todas as instâncias de nossa experiência sejam psicológicas, não há razão para supor que todos os verbos de que estou tratando sejam psicológicos.
> (Russell, 1971, p. 227)

Apesar dessa ressalva de Russell, filósofos na tradição analítica acabaram adotando o termo “atitude proposicional” para se referirem à classe de estados mentais que (1) são intencionais, e (2) têm a forma lógica proposicional.

A noção de intencionalidade empregada nesse contexto é diferente daquela do uso ordinário. No dia a dia falamos em “intenções” como relacionadas a propósitos ou objetivos, de forma que podemos perguntar se determinada ação foi “intencional”, isto é, se foi realizada “de propósito” (nesse sentido, por exemplo, em uma partida de futebol o juiz pode punir mais rigorosamente uma falta que ele julga “proposital” do que outra que não pareceu “intencional”), e da mesma maneira podemos perguntar qual é a “intenção” de uma pessoa, como o pai conservador que questiona a um pretendente “*quais são as suas intenções com a minha filha?*”, querendo saber, portanto, se ele tem

o objetivo de se casar ou não; são desses usos ordinários do termo "intenção" que trata a análise linguística de G. E. M. Anscombe em sua obra clássica *Intention* (1957). Mas em vez desse sentido cotidiano, no contexto desta dissertação a noção de intencionalidade é aquela introduzida por Franz Brentano ainda no fim do século XIX e retomada por John Searle na década de 1980 (cf. Searle, 1983), que diz respeito à característica de estados mentais serem *sobre alguma coisa*. Se, por exemplo, eu me lembro da última viagem que fiz quando saí de férias, minha lembrança é *sobre* algum aspecto da minha experiência, talvez um determinado local que visitei, pessoas que encontrei, algo diferente que comi, ou qualquer situação específica; se leio no jornal sobre uma injustiça que foi cometida e que os responsáveis não foram punidos, fico indignado, e esse estado emocional tem como *objeto* a situação descrita na notícia. Embora estados mentais sejam tipicamente intencionais nesse sentido de terem um *objeto* e serem *sobre algo*, há claras exceções: uma pessoa pode ficar triste sobre determinado acontecimento, e esse estado é intencional, mas se essa pessoa entra em depressão, ela pode experimentar estados de tristeza que não possuem um objeto definido, não são sobre nada em particular, mas constituem apenas uma indisposição generalizada; outro exemplo é a fome, que podemos experimentar como "vontade de comer determinado alimento" (nitidamente intencional), mas que também pode ocorrer pura e simplesmente sem objeto, apenas como experiência de um tipo de dor ou desprazer; e finalmente, se eu sinto dor de cabeça, o estado de experimentar a dor e ter consciência da minha experiência tem como *objeto* a dor que eu sinto, mas não dizemos que a dor em si seja "sobre" a *cabeça*, sendo antes apenas uma sensação não-intencional.

Mas além de serem intencionais no sentido de terem um *objeto*, alguns estados

mentais também seguem uma forma lógica proposicional, isto é, tendo como seu *conteúdo* uma proposição. É o caso dos verbos que Russell chamou de "verbos proposicionais", como crer/acreditar e desejar/querer, que tomam a forma de Sujeito + Verbo + "que" + Proposição. Assim, se digo que "Yuri Gagarin acredita que a terra é azul", nesse caso *Yuri Gagarin* é o sujeito, *acreditar* é o verbo, e "*a terra é azul*" indica a proposição que é o conteúdo de sua crença; outra pessoa poderia afirmar em inglês "*Yuri Gagarin believes that the Earth is blue*", e embora se trate de outro idioma, a mesma proposição é referida. De maneira geral, portanto, "*uma atitude proposicional* (…) *é o estado mental de ter uma atitude, postura, perspectiva, ou opinião sobre uma proposição ou sobre o estado de coisas potencial em que tal proposição é verdadeira*", de tal maneira que "*o que uma pessoa duvida ou espera, outra pode temer, ou acreditar, ou desejar, ou pretender – atitudes diferentes, todas à mesma proposição*" (Schwitzgebel, 2010).

No trecho citado anteriormente, Russell menciona "acreditar" e "desejar", e filósofos analíticos têm também tradicionalmente dedicado maior atenção aos estados *crença* e *desejo*. Tal preferência não é acidental, mas estratégica: faz sentido afirmar que crenças e desejos são os tipos básicos de "atitude" ou postura epistêmica que podemos tomar com relação a qualquer representação do mundo. Dada determinada proposição *p*, a "*crença que p"* corresponde à *atitude* de que *p* é uma descrição adequada de como o mundo ou alguma parte dele realmente *é*, enquanto o "*desejo que p*" corresponde à *atitude* de que *p* é uma descrição de como o mundo ou alguma parte dele *deveria ser*; atitudes como *achar* ou *recear* podem ser consideradas variações de crenças, e mesmo *pretender* e "*ter a intenção de ..."* parecem ser subtipos de combinações de crença e desejo. É com base nessas considerações que a literatura filosófica sobre a Psicologia do

Senso Comum dá tanta ênfase a, por um lado, estados mentais em geral, e por outro lado, mais especificamente a atitudes proposicionais (e particularmente crenças e desejos). Como veremos, questões centrais nesta literatura dizem respeito a se nosso entendimento interpessoal ordinário de fato baseia-se na atribuição de estados mentais e/ou de atitudes proposicionais a outras pessoas, e quais são os mecanismos que nos dão acesso a esses fatos sobre a "vida mental" de outras pessoas.

CAPÍTULO 2:

O Debate Churchland-Fodor

O debate entre Paul Churchland e Jerry Fodor acerca da Psicologia do Senso Comum é fortemente influenciado pelo ambiente intelectual apresentado no Capítulo 1. Do contexto científico, predominam temas herdados das discussões sobre o behaviorismo na psicologia e dos avanços na neurociência e na ciência da computação: questões sobre a adequação científica da terminologia mentalista, sobre a relação entre cérebro e comportamento e entre cérebro e mente, e sobre a natureza da cognição como processamento de representações simbólicas análogo ao funcionamento de computadores. Do contexto filosófico, predomina o legado da tradição analítica, não apenas na abordagem visando o esclarecimento conceitual em tópicos pontuais, mas também em temas recebidos de discussões anteriores e simultâneas ao debate: preocupações com relações interteóricas e a unidade da ciência, e sobre a natureza de predicados mentais e sua implementação física.

O objetivo desta dissertação é em parte criticar uma abordagem contemporânea à Psicologia do Senso Comum (Capítulo 4), e em parte examinar as diferenças entre o debate contemporâneo e este inicial entre Churchland e Fodor; assim, antes de discutir a transição conceitual entre os debates distintos (Capítulo 3), no presente capítulo buscarei caracterizar as abordagens conflitantes de Churchland e de Fodor, bem como os limites de sua discordância. Como filósofos prolíficos que continuam ativos, em cerca de 50 anos de carreira tanto Churchland quanto Fodor têm discutido um amplo leque de questões em diversas áreas da filosofia, mas dado o foco desta dissertação, aqui privilegiarei apenas sua contribuição ao debate filosófico acerca da Psicologia do Senso

Comum.

2.1: O Materialismo Eliminativo e a "Defesa" da PSC

A abordagem proposta por Paul Churchland nesta literatura é conhecida como "*materialismo eliminativo*", ou apenas "*eliminativismo*", e consiste basicamente na ideia de que a Psicologia do Senso Comum é tão inadequada que deverá ser eliminada (daí "eliminativo") em favor de uma descrição "mais científica" do que se passa no cérebro (daí "materialismo"). Tal ideia geral é articulada com mais precisão no artigo "*Eliminative Materialism and the Propositional Attitudes*", publicado em 1981:

> O materialismo eliminativo é a tese de que nossa concepção ordinária de fenômenos psicológicos constitui uma teoria radicalmente falsa, uma teoria tão fundamentalmente defeituosa que tanto os princípios quanto a ontologia desta teoria serão finalmente substituídos, em vez de suavemente reduzidos, pela neurociência completa.
> (Churchland, 1981, p. 67)

Churchland reconhece que a Psicologia do Senso Comum é responsável pelo modo como pensamos e agimos no dia a dia, guiando tanto nosso entendimento interpessoal quanto nossa compreensão de nós mesmos, mas ele argumenta que essa "concepção ordinária de fenômenos psicológicos" é equivocada e incompatível com as melhores teorias científicas, e que portanto ela deve dar lugar a explicações do comportamento a partir de suas bases neurológicas.

Em contraste com a posição crítica de Churchland, Jerry Fodor é considerado por muitos o grande defensor da Psicologia do Senso Comum (a qual ele também denomina "*psicologia crença-desejo*", ou "*psicologia de atitudes proposicionais*", ou ainda "*psicologia intencional*"[3]). Até certo ponto, essa fama é justificada; uma de suas

[3] Como discutido anteriormente, o termo "intencional" pode ser entendido no sentido técnico como *dotado de conteúdo* ou *sobre algo*, ou ainda no sentido comum como relativo a *intenções*, *propósitos* ou

frases mais citadas nesta literatura deixa claro como ele avalia a possibilidade de a Psicologia do Senso Comum ser, como quer Churchland, falsa: "*se a psicologia intencional realmente entrasse em colapso, isso seria, sem comparação, a maior catástrofe intelectual na história de nossa espécie*" (Fodor, 1987, p. xii); para Fodor, explicações intencionais nos termos de crenças e desejos são não apenas cruciais para nossas interações com outras pessoas e entendimento delas e de nós mesmos como agentes, mas elas também são compatíveis com as explicações do comportamento oferecidas pelas ciências cognitivas.

Como fica claro nesta breve introdução das abordagens de Churchland e de Fodor, uma questão central neste debate diz respeito à adequação científica da Psicologia do Senso Comum e sua ontologia mentalista, de maneira que poderíamos imaginar a seguinte resolução: se estabelecido que nossa concepção intencional ordinária é compatível com a ciência, Fodor está certo; caso contrário, é Churchland quem ganha o debate. Mas essa é obviamente uma perspectiva simplista demais, já que não é possível falar de compatibilidade com "a ciência" em geral; antes, o ponto em questão envolve discussões sobre a relação entre a PSC e abordagens científicas específicas (da neurociência e das ciências cognitivas), bem como sobre a unidade do conhecimento científico e a relação entre as diversas áreas da ciência entre si.

O que talvez não fique tão claro até aqui, mas que se torna evidente na transição conceitual (examinada no Capítulo 3) que informa o debate contemporâneo (discutido no Capítulo 4), é que o que motiva tanto Churchland quanto Fodor a argumentarem

---

objetivos. Aqui, a Psicologia do Senso Comum é uma "psicologia intencional" em ambos sentidos: estados mentais podem ser *sobre algo* (*intencionais* no sentido técnico), incluindo não apenas objetos, conceitos e suas relações, mas também possíveis estados de coisas futuros que uma pessoa pretende realizar (*intencionais* no sentido comum).

contra ou a favor da adequação científica da Psicologia do Senso Comum é precisamente o compromisso que eles têm com a ideia de que a PSC é uma *teoria*. Ambos assumem que em nossas interações com outras pessoas no dia a dia nós utilizamos uma teoria intuitiva por meio da qual entendemos o comportamento dos outros e de nós mesmos como originando de estados mentais em geral, e de atitudes proposicionais como crenças e desejos em particular; assim, para ambos autores, a avaliação da Psicologia do Senso Comum como uma *teoria protocientífica* – isto é, como esboço de teoria científica, que visa a corroboração empírica – deve ser feita à luz de outras teorias, a saber, teorias científicas; precisamente *quais* teorias científicas, no entanto, é uma questão central na discordância entre Fodor e Churchland. Notando-se essa particularidade do debate, a fama de Fodor como "defensor da Psicologia do Senso Comum" deixa de ser justificada: sua "defesa" não é teoricamente neutra, sendo mais correto dizer que ele defende uma compreensão específica da PSC; como o restante desta dissertação visa mostrar, outras formulações são possíveis, de forma que o eliminativismo de Churchland e a "defesa" de Fodor não são as únicas opções.

Feita essa breve apresentação das abordagens de Churchland e de Fodor, no restante deste capítulo buscarei elucidar alguns dos pontos principais do debate, começando pela estrutura proposta para as explicações ordinárias de comportamentos (ponto em que os dois concordam), e em seguida tratando de suas divergências no que diz respeito ao status da Psicologia do Senso Comum, a adequação de sua terminologia mentalista e a questão de relações interteóricas, por fim tratando da imagem da mente como análoga a um computador.

2.2: Explicações Ordinárias de Comportamentos

Já em seu primeiro artigo publicado, "*The Logical Character of Action-Explanations*" (1970), Paul Churchland escreve a respeito da Psicologia do Senso Comum, embora aqui ele ainda não utilize o termo "*folk psychology*", em vez disso se referindo às "*explicações ordinárias de ações humanas – explicações em termos de desejos, crenças, etc*"[4]. Churchland se propõe a comparar tais explicações ordinárias aos tipos de explicações oferecidas na ciência, e conclui que ambas têm as mesmas características lógicas e estruturais; mais especificamente, ele defende que os dois tipos de explicações constituem *esboços de argumentos dedutivo-nomológicos*, conforme o modelo de condições iniciais e hipóteses universais proposto por Carl Hempel (cf. p. ex. Hempel 1942). Partindo de uma explicação ordinária como "*X fez A porque desejava Ø*" em que a variável Ø pode ser substituída por uma indicação proposicional do tipo "*que P (seja o caso)*", Churchland argumenta que seu caráter dedutivo-nomológico fica evidente quando identificamos objeções que poderíamos apresentar a tal explicação ordinária. Por exemplo, se alguém respondesse "mas X não desejava Ø!" ou ainda "X não sabia que fazer A causaria Ø!", a explicação "*X fez A porque desejava Ø*" seria automaticamente rejeitada. Após discutir uma série de objeções como estas, Churchland propõe que nossas explicações ordinárias baseiam-se em *esboços de leis* como a seguinte:

> L1: (X) (Ø) (A) (Se [1] X deseja Ø, e
> [2] X acredita que fazer A é uma maneira de realizar Ø naquelas circunstâncias, e
> [3] não existe nenhuma outra ação que X acredita ser uma maneira de realizar Ø, naquelas circunstâncias, que X julga preferível a fazer A, e

[4] Churchland escreve "*garden-variety explanations of human actions*". A expressão idiomática "*garden-variety*" não possui tradução direta para o português (ao pé da letra, "*garden-variety explanations*" seria algo como "explicações de jardim"), mas a ideia é simplesmente de algo "ordinário" ou "comum".

> [4] X não tem nenhum outro desejo (ou conjunto de desejos) que, naquelas circunstâncias, sobrepuje seu desejo de Ø, e
> [5] X sabe como fazer A, e
> [6] X é capaz de fazer A,
> então [7] X faz A).
> (Churchland, 1970, p. 221-222) [5]

Churchland reconhece que a maioria das pessoas não seria capaz de compor uma lista como esta, mas defende que o fato de L1 ser "*conhecimento comum*" (embora tácito, implícito) é que faz com que a negação de uma condição similar àquelas listadas de [1] a [6] sejam objeções reais à explicação "*X fez A porque desejava Ø*"[6]. Ele ainda explica que não está propondo que L1 *é uma lei verdadeira*, isto é, que sempre que suas condições são atendidas o resultado é que a pessoa *X* em questão de fato faz *A*; antes, o que ele propõe é que *é verdade* que o esboço de lei L1 é "*pressuposto por nossas práticas explicativas como o esboço de uma lei verdadeira*" (1970, p. 223). Ou seja, mesmo se em algum cenário todas as condições de [1] a [6] são satisfeitas mas [7] não ocorre e *X* não realiza ação *A* (mostrando assim que L1 é uma lei falsa), isso não tornaria falsa a ideia de que as pessoas implicitamente raciocinam de acordo com L1 em suas explicações comportamentais ordinárias. É precisamente esse último ponto que Churchland defende, e ele esclarece sua proposta:

> Considera-se que predicados psicológicos recebem uma definição implícita (parcial) pelo conjunto de princípios nômicos de que fazem parte, e que a justificativa para essa "Teoria Humana sobre Pessoas"

---

[5] Neste esboço de lei L1, as variáveis (X), (Ø), e (A) são interpretadas da seguinte maneira: (X) corresponde a uma pessoa, (Ø) corresponde a uma proposição que representa o que esta pessoa deseja realizar, e (A) indica a ação considerada como um meio de realizar tal desejo. Pode-se, portanto, ler o esboço de lei substituindo (X) por um nome próprio (e.g. João), (Ø) por um objetivo (e.g. manter-se seco, não se molhar), e (A) pela estratégia adotada para realizar tal objetivo (e.g. abrir o guarda-chuva): "se [1] João deseja manter-se seco, e [2] João acredita que abrir o guarda-chuva é uma maneira de manter-se seco naquelas circunstâncias, e [3] não existe nenhuma outra ação que João acredita ser uma maneira de manter-se seco, naquelas circunstâncias, que João julga preferível a abrir o guarda-chuva, e [4] João não tem nenhum outro desejo ou conjunto de desejos que, naquelas circunstâncias, sobrepuje seu desejo de manter-se seco, e [5] João sabe como abrir o guarda-chuva, e [6] João é capaz de abrir o guarda-chuva, então [7] João abre o guarda-chuva".

[6] Voltando ao exemplo da nota anterior, a explicação "*João abriu o guarda-chuva porque desejava manter-se seco*" seria rejeitada se qualquer uma das condições de [1] a [6] fosse negada: por exemplo, a negação de [5], "*João não sabe como abrir o guarda-chuva*", ou então de [6], "*João não é capaz de abrir o guarda-chuva*".

> não é que as pessoas concluem introspectivamente que a teoria é verdadeira a respeito de seu próprio caso e então observam que o comportamento de outros humanos é similar a seu próprio comportamento, mas sim que a teoria é explicativamente bem-sucedida com relação ao objeto a que se aplica – especificamente, o comportamento de humanos em geral. Portanto, uma exposição completa de nosso arcabouço conceitual sobre pessoas – o arcabouço que compreende todos os predicados psicológicos – conteria um vasto número de princípios ("verdades conceituais") que são de caráter nômico em vez de "meramente analítico".
> (Churchland, 1970, p. 225-226).

A conclusão de Churchland nesse primeiro artigo é que essa "Teoria Humana sobre Pessoas" está tão arraigada em nosso raciocínio ordinário que sequer notamos que a utilizamos corriqueiramente; a despeito disso, ele defende, nossas explicações ordinárias só nos parecem de fato *explicativas* porque elas obedecem à lógica dos esboços de leis no molde dedutivo-nomológico (como o exemplo L1) que tornam predicados psicológicos inteligíveis.

Em um artigo posterior, talvez o mais importante no desenvolvimento da abordagem eliminativista de Churchland, "*Eliminative Materialism and the Propositional Attitudes*" (1981), Churchland reitera que são leis ou esboços de leis que conectam comportamentos a condições explicativas e assim garantem que nossas explicações ordinárias pareçam explicar alguma coisa; mas, indo além do que já propusera em 1970, aqui Churchland estende essa função das leis como também nos permitindo fazer previsões e considerações contrafactuais, e, de maneira geral, compreender as pessoas ao nosso redor:

> Cada um de nós só entende os outros tão bem assim porque compartilhamos o domínio implícito de um corpo integrado de conhecimentos tradicionais[7] a respeito das relações nômicas entre circunstâncias externas, estados internos, e comportamentos observáveis.

[7] Aqui traduzo como "conhecimento tradicional" o termo original de Churchland, "*lore*", que em português significa literalmente "saber" (substantivo), e do qual deriva, por exemplo, "*folklore*", ao pé da letra "o saber do povo".

(Churchland, 1981, p. 69)

Churchland não entra em detalhes para explicar porque, em sua perspectiva, esse *corpo integrado de conhecimentos tradicionais* deve ser considerado uma *teoria empírica*; antes, o principal (se não único) critério para tal classificação parece ser as *relações nômicas* que ele supõe fundamentarem nosso entendimento ordinário. Mesmo assim, é interessante notar que para Churchland essa teoria é responsável por nossa compreensão tanto do comportamento de outras pessoas quanto os nossos próprios. Ele defende que nossas atribuições de estados mentais a outras pessoas não se dão por *dedução* a partir de seus comportamentos e nem por *analogia indutiva* a partir de nosso próprio, mas sim pela formação de *hipóteses explicativas* a partir das leis ou esboços de leis que ele supõe comporem nosso entendimento ordinário. O mesmo se aplica ao nosso conhecimento introspectivo, o qual ele concebe como "*apenas uma instância de um hábito adquirido de resposta conceitual aos nossos estados internos*" cuja integridade é "*sempre contingente à integridade do arcabouço conceitual (teoria) em que se origina*" (Churchland, 1981, p. 70). Estes dois últimos pontos são complementares: como, na visão de Churchland, entendemos mentes e comportamentos humanos por meio de hipóteses explicativas baseadas nas relações nômicas de uma teoria empírica, tanto nosso conhecimento de outras mentes quanto nosso conhecimento de nossa própria mente não passam de instâncias de nossa aplicação dessa teoria; nesta perspectiva, portanto, não existe nenhuma diferença significativa entre a atribuição de estados mentais a outras pessoas ou a nós mesmos (introspecção).

Em comparação com o artigo de 1970, aqui Churchland dá maior ênfase a como essas supostas leis da Psicologia do Senso Comum quantificam sobre *proposições* e estabelecem relações entre comportamentos e *atitudes proposicionais*, isto é, estados

mentais representados no modo indicativo e cujo conteúdo é expresso proposicionalmente, como "a crença que *p*", em que a proposição *p* pode ser expressa de maneiras linguisticamente variadas desde que tenham o mesmo significado (como "está chovendo", em português, ou "*it is raining*", em inglês, etc.). Churchland oferece mais exemplos das relações nômicas que ele considera fundamentarem a Psicologia do Senso Comum:

(4) (x) (p) (q) [((*x* acredita que *p*) & (*x* acredita que (se *p* então *q*)) ⊃ (excluindo confusão, distração, etc., *x* acredita que *q*)]

(5) (x) (p) (q) [((*x* deseja que *p*) & (*x* acredita que (se *q* então *p*)) & (*x* é capaz de fazer com que seja o caso que *q*)) ⊃ (excluindo desejos conflitantes ou estratégias preferidas, *x* faz com que seja o caso que *q*)]

(Churchland, 1981, p. 71)

Postulando que nossa compreensão ordinária do comportamento humano, tanto introspectivamente quanto no caso de outras mentes, fundamenta-se em generalizações como estas, Churchland novamente conclui que a Psicologia do Senso Comum espelha a estrutura de teorias e leis científicas, como por exemplo a relação nômica entre massa, força e aceleração: "*(x) (f) (m) [((x tem massa m) & (x sofre força f) ⊃ (x acelera a f/m)*" (Churchland, 1981, p. 70).

Jerry Fodor entende de maneira semelhante a estrutura das explicações do comportamento que utilizamos no dia a dia. Ele reconhece, em sua típica veia crítica, que seria errôneo pensar na Psicologia do Senso Comum como "*meramente um punhado de truísmos do tipo que se aprende no colo da avó: que a criança que se queima tem medo de fogo,* (...) *que o dinheiro não compra felicidade, que reforço afeta a frequência de respostas*[8]*, e que o caminho para o coração de um homem é seu*

[8] Referência óbvia ao behaviorismo como trivial e superficial. Na sequência Fodor afirma que nenhum desses truísmos (incluindo o behaviorista) merece ser preservado.

estômago"; antes, para Fodor, a forma típica de explicações psicológicas do senso comum não é a de "subsunção de platitudes", mas sim:

> quando tais explicações são tornadas explícitas, elas frequentemente exibem a 'estrutura dedutiva' que é tão característica de explicações nas ciências reais. Isso tem duas partes: as generalizações que subjazem a teoria são definidas sobre inobserváveis, e elas resultam em predições por meio de iterações e interações em vez de serem diretamente instanciadas.
> (Fodor, 1987, p. 6-7)

Para Fodor, nossas explicações ordinárias assumem que comportamentos observáveis surgem "*ao fim de uma cadeia causal cujos elos são eventos mentais – e por isso, inobserváveis*" (Fodor, 1987, p. 7); mais ainda, tais generalizações sobre as interações entre comportamentos observáveis e estados mentais inobserváveis têm uma estrutura que "*suporta* [considerações] *contrafactuais em uma maneira familiar a explicações causais em geral*" (Fodor, 1990, p. 4), sendo assim similar à estrutura de argumentos em geral (Fodor, 1987, p. 13), e à do silogismo prático em particular (Fodor, 1990, p. 5), no qual a conclusão das premissas tem caráter imperativo, "*governando inferências que terminam em ação*" (Pakaluk, 2005, p. 256) e substanciando a "*formação e execução de intenções*" (McCann, 1999, p. 729). Dadas as premissas "*Eu desejo P*" e "*P apenas se Q*", a 'conclusão' (ou seja, a realização de Q) só é obtida "*em virtude da relação entre os conteúdos de crenças e desejos*" (Fodor, 2008, p. 15), ou seja, em virtude da relação entre *o desejo que P* e *a crença que P apenas se Q*; da mesma forma que essa relação nomológica motiva uma pessoa à ação, ela também informa nossa compreensão do nosso próprio comportamento e o das pessoas ao nosso redor.

Em suma, para Fodor "*a estrutura das explicações psicológicas do senso comum é dedutiva-nomológica*" e "*as generalizações empregadas são generalizações sobre inobserváveis dotados de poderes causais*" (Davies & Stone, 1995, p. 11); tais

generalizações são nitidamente similares ao tipo proposto por Churchland: "*pessoas que desejam que P, e acreditam que Q seria suficiente para causar P, e que não possuem desejos conflitantes ou estratégias preferidas, tentarão realizar Q*" (Churchland, 1988, p. 59).

2.3: A Avaliação da Psicologia do Senso Comum

Churchland e Fodor concordam no que diz respeito à estrutura lógica de nossas explicações ordinárias de comportamentos, de maneira que ambos poderiam afirmar ser "*incontroverso que as pessoas regularmente dão conta de seu comportamento voluntário citando crenças e desejos que elas têm*" (Fodor, 1990, p. 4), o mesmo valendo para explicações e previsões de comportamentos alheios, em todos os casos a partir de generalizações de caráter nômico entre comportamentos e estados mentais. Da mesma maneira, os dois também concordam no que diz respeito à natureza da Psicologia do Senso Comum, assumindo que ela é tal teoria implícita composta por generalizações ordinárias. A essência do debate entre Churchland e Fodor, portanto, reside na sua discordância sobre a *avaliação* desta teoria, ou seja, sobre a qualidade da Psicologia do Senso Comum e seu status em relação a teorias científicas (o que também traz à tona sua divergência na questão da relação de teorias científicas entre si – ponto a ser discutido mais adiante, na seção 2.4).

2.3.1: Sucesso Preditivo e Explicativo

Fodor reconhece que "*é possível que a psicologia 'crença/desejo' do senso comum precise ser corrigida. Ou cancelada*" (Fodor, 1990, p. 5), mas ele parece

considerar esta como apenas uma possibilidade lógica; antes, sua opinião é que "*a adequação preditiva da psicologia do senso comum está além de disputa racional*":

> Se você quer saber onde meu corpo físico estará na próxima quinta-feira, a mecânica – nossa melhor ciência de objetos de tamanho médio (…) – será absolutamente inútil. A melhor maneira de descobrir (na prática, geralmente, a única maneira de descobrir) é: me pergunte!
> (Fodor, 1987, p. 6)

Ou seja, o conhecimento de fatos cientificamente estabelecidos sobre o movimento de corpos de tamanho médio é muito pouco informativo em uma situação como essa; antes, a melhor (ou única) maneira para uma pessoa prever onde Fodor estará na semana que vem envolverá o conhecimento dos estados mentais dele: seus desejos, intenções, crenças (como, por exemplo, seu desejo de descansar, sua intenção de ir à praia em um dia de folga, e sua crença de que a próxima quinta-feira é ponto facultativo na universidade em que trabalha), bem como o conhecimento das relações nomológicas entre estados mentais como estes e comportamentos relevantes. E a fonte desses conhecimentos sobre estados mentais e suas relações com comportamentos é óbvia:

> A teoria da qual obtemos tamanho poder preditivo é a boa e velha psicologia crença/desejo do senso comum. É ela que nos diz, por exemplo, como inferir a intenção das pessoas a partir dos sons que elas emitem (…), e como inferir o comportamento das pessoas a partir de suas intenções (…). E tudo isso funciona não apenas com pessoas cuja psicologia você conhece intimamente (…) [mas também] funciona com completos estranhos.
> (Fodor, 1987, p. 3)

Mas não é apenas em *previsões* de comportamentos futuros que vemos o sucesso da Psicologia do Senso Comum. Naturalmente, o caráter dedutivo-nomológico das generalizações sobre relações entre estados mentais e comportamentos é responsável pelo seu sucesso também em *explicações* de comportamentos já realizados: uma lei pode ser aplicada em ambas as direções, ou seja, a partir das condições iniciais para a obtenção de uma previsão dos efeitos, ou a partir dos efeitos obtidos para a geração de

uma explicação das condições iniciais. Assim, quando buscamos compreender nosso próprio comportamento ou o de outras pessoas, nossas *explicações* invariavelmente recorrem às categorias intencionais características da PSC: "*Na prática, nós não temos nenhuma alternativa ao vocabulário das explicações psicológicas do senso comum*" (Fodor, 1987, p. 8), e mais ainda, "*Nós não temos a menor ideia de como explicar nós mesmos a nós mesmos exceto com um vocabulário que é saturado de psicologia crença/desejo*" (Fodor, 1987, p. 9).

Fodor afirma duvidar que possamos de fato "desistir" da Psicologia do Senso Comum como propõe o materialismo eliminativo de Churchland, dizendo não acreditar que "*nosso intelecto é constituído de tal maneira que a eliminação* [da PSC] (…) *seja biologicamente viável*" (Fodor, 1987, p. xii). Ao contrário, Fodor defende que nossa compreensão intuitiva nos termos de estados mentais e atitudes proposicionais são (muito provavelmente) universais e inatas. Por um lado, ele diz, "*não existe, até onde eu saiba, nenhum grupo humano que não explica comportamos imputando crenças e desejos ao comportamento*", e a esse comentário ele acrescenta que "*se um antropólogo afirmasse ter encontrado tal grupo, eu não acreditaria nele*" (Fodor, 1987, p. 132). Por outro lado, ele afirma que no desenvolvimento ontogenético humano "*muito do aparato de explicações mentalistas parece estar em operação bem cedo*", tal que "*sofisticação cognitiva*" já aparece em formas rudimentares nos primeiros anos de vida de uma criança (Fodor, 1987, p. 132).

Em suma, para Fodor, tanto em previsões de comportamentos futuros quanto na explicação de comportamentos passados, e no que diz respeito tanto aos nossos próprios comportamentos quanto aos de outras pessoas (conhecidas ou não), "*A psicologia do*

senso comum funciona tão bem que ela desaparece" (Fodor, 1987, p. 3). Seu otimismo com relação à qualidade da Psicologia do Senso Comum é de fato extremo:

> Eu não tenho dúvidas de que essa teoria (que eu chamo de "psicologia ordinária crença-desejo") é muito próxima de ser verdadeira. Minha razão para acreditar nisso (...) é que a psicologia ordinária crença-desejo explica vastamente mais fatos sobre comportamentos do que qualquer outra teoria disponível. E isso não poderia ser diferente: não *existem* outras teorias disponíveis.
> (Fodor 1987, p. x)

Enquanto em *Psychosemantics* (1987) Fodor faz sua famosa declaração de que o colapso da PSC seria *a maior catástrofe intelectual da história*, em um livro posterior, *A Theory of Content and Other Essays* (1990), ele manifesta novamente o mesmo sentimento:

> Se não for literalmente verdadeiro que o meu desejo é causalmente responsável por eu esticar o meu braço, (…) e que minha crença é causalmente responsável por eu dizer [alguma coisa], se nada disso é literalmente verdadeiro, então praticamente tudo o que acredito sobre qualquer coisa é falso e este é o fim do mundo.
> (Fodor 1990, p. 156)

2.3.2: Dificuldades e Estagnação

Churchland obviamente não compartilha desse otimismo, e aponta alguns problemas que ele considera indicações centrais da inadequação da Psicologia do Senso Comum. Ele argumenta que, embora à primeira vista a PSC pareça explicativamente e preditivamente bem-sucedida em nosso uso cotidiano, nossa compreensão de comportamentos nos termos de estados mentais e atitudes proposicionais só se aplica a condições e contextos muito limitados, não dando conta de fenômenos como "*doenças mentais, sono, criatividade, memória, diferenças de inteligência e formas distintas de aprendizado*" (Churchland, 1996, p. 8). Em uma exposição mais longa, Churchland lista ainda outras questões que ele considera problemáticas para a Psicologia do Senso

Comum, ou seja, questões nas quais ela *não* é bem-sucedida:

> Considere a natureza e a dinâmica de doenças mentais, a faculdade da imaginação criativa, ou o fundamento das diferenças de inteligência entre indivíduos. Considere nossa completa ignorância da natureza e funções psicológicas do sono, este estado curioso em que uma pessoa passa um terço da vida. Reflita sobre a habilidade comum de pegar uma bola no ar enquanto se está correndo, ou atingir com uma bola de neve um carro em movimento. Considere a construção interna de uma imagem visual em 3D a partir de diferenças sutis no conjunto de estímulos 2D em nossas retinas. Considere a rica variedade de ilusões da percepção visual ou de outros tipos. Ou considere o milagre da memória, com sua capacidade instantânea de recordação relevante. (Churchland, 1981, p. 73)

Todos esses fenômenos, Churchland afirma, não são facilmente explicáveis com base na concepção ordinária de que a mente opera por meio da *manipulação e armazenamento de proposições*; nossa teoria intuitiva com suas relações nômicas entre comportamentos e atitudes proposicionais não esclarece esses aspectos de nossa experiência cotidiana, principalmente aqueles em que a linguagem não é diretamente envolvida (como é o caso de percepções e ilusões da percepção) ou ainda aqueles em que estritamente não pode existir linguagem (como no caso da aprendizagem em bebês recém-nascidos e do comportamento adaptativo de animais não-humanos). Embora essas dificuldades não provem que a Psicologia do Senso Comum é equivocada, para Churchland, a "*incapacidade constitutiva de ao menos começar a examinar esses mistérios básicos*" de nossa experiência comum indica que a PSC é, na melhor das hipóteses, "*uma teoria altamente superficial, uma interpretação parcial e limitada de uma realidade mais profunda e complexa*" (Churchland, 1981, p. 74).

Se estas dificuldades não decretam de maneira definitiva a falsidade da Psicologia do Senso Comum, Churchland sugere que podemos analisar seu potencial de desenvolvimento futuro com base em seu crescimento histórico. Ele argumenta que a

história da Psicologia do Senso Comum é uma história de *retração*, *infertilidade* e *degeneração*. Ele afirma que "*explicações intencionais têm sido rejeitadas de um domínio após outro – dos corpos celestes, do vento e dos mares, de uma pletora de deuses e espíritos, da visitação de doenças, e assim por diante*" (Churchland, 1996, p. 8) – e a tendência é que também desapareçam de nossas descrições de pessoas, incluindo nós mesmos. Em culturas tradicionais, os elementos da natureza (como ventos, rios ou o sol) podiam irar-se ou ser generosos, por exemplo, e isso não era visto metaforicamente já que as pessoas de fato faziam sacrifícios para apaziguar essas entidades; apenas nos últimos dois ou três mil anos é que o domínio da PSC gradativamente se retraiu aos animais superiores, e ainda assim "*o conteúdo e o sucesso da PSC não avançou sensivelmente*", ao ponto de que "*hoje somos insignificantemente melhores em explicar o comportamento humano* [nos termos da PSC] *do que era Sófocles*" e os gregos antigos (Churchland, 1981, p. 74). Churchland identifica isso como um período longo demais de *estagnação e infertilidade*, e afirma, nos termos de Imre Lakatos, que há milênios a Psicologia do Senso Comum tem sido um *programa de pesquisa degenerativo*.

Em resposta a essas duas críticas, apesar de seu otimismo extremo, Fodor parece concordar ser possível que existam problemas com a Psicologia do Senso Comum, e que se ela não é uma teoria completa, ela pode ser corrigida. O ponto para ele é que a psicologia científica completa deverá ter semelhanças importantes com o arcabouço conceitual intencional que guia nossa compreensão de comportamentos: "*existem certos elementos centrais da psicologia do senso comum que serão compartilhados por uma psicologia científica madura*", a saber, "*uma psicologia madura será intencional do princípio ao fim, e* (…) *os estados intencionais que ela postula serão causalmente*

implicados em explicações nomológicas do comportamento humano" (Rives, 2010). Portanto, se de fato existem tantos fenômenos psicológicos cotidianos dos quais a PSC não dá conta, isso não contradiz a posição de Fodor, já que ele não defende a *integridade* absoluta da PSC em seu estado atual, mas sim sua *integrabilidade* com ciências maduras:

> Preservar as atitudes [proposicionais] – vindicar a psicologia do senso comum – significa mostrar como poderíamos ter (ou, ao menos, mostrar *que* poderíamos ter) uma ciência respeitável cuja ontologia explicitamente reconhece estados que exibem os tipos de propriedades que o senso comum atribui às atitudes [proposicionais].
> (Fodor, 1987, p. 10)

Já no que diz respeito à suposta estagnação da Psicologia do Senso Comum, Fodor responde que "*essa acusação é até certo ponto justificada*", mas acrescenta que "*a psicologia do senso comum pode ser uma ciência implícita, mas ela não é* (…) *uma ciência implícita de pesquisa*" (Fodor, 1987, p. 20); isto é, para ele esta aparência de esterilidade não decorre da *falsidade* da teoria, mas se deve ao fato de que a PSC não é uma ciência em que são realizadas pesquisas que possam gerar progresso. Nesse sentido, e em consonância com a resposta do parágrafo anterior, Fodor afirma:

> Se, em suma, você quer avaliar progresso, você deve examinar não a teoria ordinária implícita mas sim o melhor candidato para sua vindicação explícita. E aqui o progresso tem sido enorme. (…) A verdadeira proeza é que estamos (talvez) prestes a resolver um grande mistério sobre a mente: *Como podem seus processos causais serem semanticamente coerentes?* Ou, se você prefere com mais pompa: *Como é mecanicamente possível a racionalidade?* Note que esse tipo de problema não pode ser exposto, e muito menos resolvido, sem que suponhamos – como quer a psicologia crença/desejo do senso comum – que existem estados mentais com conteúdos semânticos e papéis causais. Uma boa teoria é uma que nos leva a fazer perguntas que tem respostas.
> (Fodor, 1987, p. 20)

2.4: Relações Interteóricas e a Unidade da Ciência

Dada a avaliação negativa que Churchland faz da capacidade explicativa e preditiva da PSC, para ele a questão central passa a ser determinar como "*a ontologia de uma teoria (a Psicologia do Senso Comum) será ou não relacionada à ontologia de uma outra teoria (a neurociência completa)*" (Churchland, 1981, p. 72). Se por um lado, em sua avaliação, o retrospecto da Psicologia do Senso Comum não indica uma tendência de progresso e desenvolvimento, mas sim de diminuição de seu domínio de aplicação, por outro lado Churchland acredita que a situação é ainda pior no que diz respeito à compatibilidade da PSC com o restante de nosso conhecimento científico.

Churchland aponta que uma grande síntese teórica tem aproximado áreas como "*a física de partículas, a teoria molecular e atômica, a química orgânica, a teoria da evolução, a biologia, a fisiologia, e a neurociência materialista*"; entretanto, ele argumenta, não há indicações de que as categorias intencionais da Psicologia do Senso Comum poderão formar um todo coerente com essas "*descrições e explicações dos inputs sensoriais, atividade neural, e controle motor humanos*" (Churchland, 1981, p. 75); para ele, essa nossa compreensão ordinária de comportamentos como sendo originados em estados mentais e atitudes proposicionais "*não dá sinais de ser suavemente integrável à síntese emergente das diversas ciências físicas, químicas, biológicas, fisiológicas, e neurocomputacionais*" (Churchland, 1996, p. 8). Embora ele afirme que não seja correto descartar por completo a possibilidade de uma integração futura, ele definitivamente não tem esperanças de que isso vá ocorrer: "*A impotência explicativa e longa estagnação da PSC inspiram pouca fé de que suas categorias serão claramente refletidas no arcabouço da neurociência*" (Churchland, 1981, p. 75).

Churchland aponta que o pessimismo a respeito da redução da terminologia mentalista da Psicologia do Senso Comum às categorias da neurociência não é exclusividade de seu materialismo eliminativo: dualistas e funcionalistas também não acreditam em tal redução por razões distintas mas que mantêm a integridade da PSC; já Churchland afirma que a verdadeira razão para ser pessimista a respeito dessa redução é que "*a Psicologia do Senso Comum é uma descrição radicalmente inadequada de nossas atividades internas, confusa e defeituosa demais para sobreviver por meio de redução interteórica*" (Churchland, 1981, p. 72).

Churchland compara a PSC a uma "física do senso comum", que guiou os seres humanos durante longo período da história até que nossas teorias científicas fossem formuladas, e que até hoje é acionada por leigos em seus cálculos mais corriqueiros (como, por exemplo, na hora de atravessar a rua evitando ser atingido por um carro). Mas, da mesma forma que os termos e conceitos dessa "física do senso comum" não têm o mesmo rigor e precisão que os da física científica, Churchland defende que os termos intencionais e proposicionais empregados pela PSC são inadequados, e ele prevê que a PSC inevitavelmente será eliminada e substituída pela neurobiologia, e seus termos intencionais e proposicionais rejeitados em favor de explicações de padrões de ativamento neural e conexões sinápticas.

Em um artigo escrito por Paul e sua esposa Patricia, os Churchlands explicam que reduções interteóricas são relações entre "*dois arcabouços conceituais distintos para a descrição de fenômenos*", e não "*a relação entre domínios distintos de fenômenos*"; mais precisamente, o objetivo de uma redução entre teorias é "*mostrar que*

o que pensávamos serem dois domínios na verdade é um domínio só, embora ele tenha sido descrito em dois (ou mais) vocabulários diferentes" (Churchland & Churchland, 1990, p. 70). Os Churchlands citam exemplos de reduções interteóricas bem-sucedidas, como a dedução das três leis planetárias de Kepler pelas três leis da mecânica newtoniana, a reconcepção de *temperatura* como *energia cinética*, a unificação conceitual da química clássica (de valência) na física (quântica) atômica e subatômica, e a redução das três leis newtonianas à teoria especial da relatividade, sendo que sobre o último caso eles afirmam: "*como geralmente ocorre com reduções interteóricas, a nova teoria nos permite explicar muito do que a teoria anterior era incapaz de explicar*" (Churchland & Churchland, 1990, p. 69). Mais do que isso, porém, principalmente no caso da inovação teórica de Einstein, o resultado é que:

> a teoria anterior se revela não apenas restrita, mas falsa em aspectos importantes. Espaço e tempo não são distintos, como Newton supunha, e simplesmente não existem propriedades intrínsecas como massa e comprimento que são invariáveis sobre todas referências inerciais.
> (Churchland & Churchland, 1990, p. 71)

Em contraste com esses casos em que a ontologia da teoria anterior é suavemente reduzida a uma teoria nova e também outros casos em que a identidade interteórica é menos evidente e a teoria antiga "sobrevive" apenas em uma forma extremamente modificada, Churchland aponta que existem pares de teorias em que "*a teoria mais antiga, e sua ontologia antiquada, são completamente eliminadas em favor da ontologia mais útil e das leis mais bem-sucedidas da nova teoria*" (Churchland & Churchland, 1990, p. 72). Nesse sentido, Churchland vê a situação atual da Psicologia do Senso Comum como similar à da alquimia frente ao surgimento da química elementar, à da cosmologia aristotélica em comparação com a mecânica clássica, e ao vitalismo à época do desenvolvimento da química orgânica. Obviamente, sua intenção

com essas analogias é afirmar que a Psicologia do Senso Comum é a tal teoria antiquada e equivocada que tenta competir com uma teoria científica mais rigorosa e correta, a neurociência, e que se não há integração entre as teorias, a primeira deve ser abandonada em favor da segunda. A similaridade, por exemplo, com a teoria do flogisto proposta por Stahl no século XVIII e posteriormente refutada por Lavoisier, é que, para ele, os estados mentais e atitudes proposicionais postulados pela PSC são conceitos vazios de significado, que não apontam para nada real; termos como "crenças" e "desejos" deverão dar lugar a teorias melhores sobre a armazenagem de informações e ativação de comportamentos: caso sua redução a processos neurais se mostre impossível, diz ele, ficará provada a falsidade da PSC (Churchland & Churchland, 1990, p. 73).

A resposta de Fodor vem em pontos distintos. Primeiro, em termos gerais, ele afirma que a história da ciência não justifica a expectativa de uma unificação mais abrangente do conhecimento científico:

> o desenvolvimento da ciência tem testemunhado a proliferação de disciplinas especializadas ao menos tão frequentemente quanto tem testemunhado sua eliminação, então o entusiasmo pela ideia de que finalmente existirá apenas a física não pode ser uma mera indução sobre sucessos reducionistas passados.
> (Fodor, 1975, p. 9-10)

Mais do que isso, Fodor acredita que a adoção de uma postura fisicalista não exige a redução da taxonomia das ciências especiais à taxonomia da física:

> Se a ciência deve ser unificada, então todas as taxonomias devem se aplicar às mesmas coisas. Se a física deve ser a ciência básica, então cada uma dessas coisas deverá ser uma coisa física. Mas não é requerido que as taxonomias empregadas pelas ciências especiais sejam elas próprias reduzidas à taxonomia da física. Isso não é requerido, e provavelmente não é verdadeiro.
> (Fodor, 1975, p. 25)

Fodor explica que, para ele, não é a "*nossa relação epistêmica com o mundo*" que

explica a existência das ciências especiais; antes, é a própria natureza da realidade ou a constituição do mundo que justifica as ciências especiais: "*nem todos os tipos (nem todas as classes de coisas e eventos sobre as quais podem ser feitas importantes generalizações contrafáticas) são, ou correspondem a, tipos físicos*" (Fodor, 1975, p. 24).

Assim, mais especificamente no que diz respeito ao status da Psicologia do Senso Comum frente a outras teorias científicas, a postura funcionalista de Fodor motiva sua expectativa sobre a viabilidade e a autonomia de uma psicologia científica intencional (que, portanto, vindica a PSC); ao mesmo tempo, esta postura o leva a rejeitar o ultimato reducionista de Churchland, que para Fodor é absolutamente desnecessário e equivocado:

> Nós não possuímos o aparato teórico apropriado para a taxonomia psicológica de eventos neurológicos. Se no fim das contas a decomposição funcional do sistema nervoso corresponder precisamente à sua decomposição neurológica (anatômica, bioquímica, física), então teremos apenas razões epistemológicas para estudar a primeira em vez da última. Mas suponha que não existe tal correspondência. Suponha que a organização funcional do sistema nervoso traspassa sua organização neurológica. Então a existência da psicologia depende não apenas do fato de que neurônios são tão depressivamente pequenos, mas antes [depende] do fato de que a neurologia não postula os tipos requeridos pela psicologia.
> (Fodor, 1975, p. 24)

Fodor observa que nosso arcabouço conceitual mentalista intuitivo é "*consequência de literalmente milênios de tentativas de entendermos uns aos outros e a nós mesmos*", e concede que avanços científicos poderão eventualmente transformar essa tradição, mas que esse deverá ser um processo gradual, e que será "*surpreendente se conseguirmos dar conta do comportamento com categorias explicativas inventadas a partir do nada*" (Fodor, 1975, p. 198). Como fica evidente em outras de suas obras, Fodor não acredita

ser preciso inventar tais categorias explicativas quando já temos uma que funciona tão bem.

2.5: Alternativas Hipotéticas à Psicologia do Senso Comum

Churchland reconhece que é por meio da Psicologia do Senso Comum que nós ordinariamente compreendemos a nós mesmos e às pessoas ao nosso redor, mas afirma que isso não deve nos impedir de buscar alternativas:

> A PSC é nada mais e nada menos que uma teoria culturalmente entrincheirada a respeito de como nós e os animais superiores funcionam. Ela não possui características especiais que a tornam empiricamente invulnerável, nem funções únicas que a tornam insubstituível, nem absolutamente nenhum tipo de status especial.
> (Churchland, 1981, p. 84)

Assim, admitindo que pode parecer difícil até mesmo simplesmente imaginar o que poderia substituir a PSC, Churchland conclui seu artigo clássico de 1981 indicando três exemplos de mudanças que avanços na neurociência poderiam oferecer. Como ele explica, estes são três cenários hipotéticos em que, do primeiro ao terceiro, "*a concepção operante de atividade cognitiva é progressivamente divorciada das formas e categorias que caracterizam a linguagem natural*" (Churchland, 1981, p. 84), se distanciando, portanto, da Psicologia do Senso Comum.

No primeiro cenário, desenvolvimentos nas pesquisas a respeito da *estrutura e atividade do cérebro* resultam em "*uma nova cinemática e dinâmica correlata para o que agora é conhecido como atividade cognitiva*" (Churchland, 1981, p. 84). Churchland imagina uma teoria cujo domínio de aplicação se estenderia a qualquer tipo de cérebro (não apenas humanos), a qual se integraria a outras teorias estabelecidas,

como a biologia evolutiva e a termodinâmica de não-equilíbrio. Essa teoria apresentaria um novo vocabulário sobre estados complexos de configuração neural em um espaço de fase de quatro ou cinco dimensões, com leis governando "*a interação, movimento, e transformação de estados 'sólidos' naquele espaço, assim como suas relações com qualquer tipo de transdutor sensório ou motor que o sistema possui*" (Churchland, 1981, p. 85). Embora não sendo exaustiva na especificação desses estados de configuração (ou "sólidos"), Churchland imagina que essa teoria ofereceria, para fenômenos de curto prazo, "*excelentes explicações e previsões sobre mudanças internas e comportamento externo*", e para fenômenos de prazo mais longo, "*descrições poderosas e unificadas do processo de aprendizagem, da natureza de doenças mentais, e de variações em caráter e inteligência em todo o reino animal bem como entre diferentes indivíduos humanos*" (Churchland, 1981, p. 85). De acordo com esta teoria, ficaria evidente que "*simplesmente não existem estados governados por leis* [sobre atitudes proposicionais] *do tipo que a Psicologia do Senso Comum postula. As leis que realmente governam nossas atividades internas são definidas em estados e configurações cinéticas mais complexas e diferentes, assim como também são* [definidos] *os critérios normativos para integridade de desenvolvimento e virtude intelectual*" (Churchland, 1981, p. 85). Em um cenário como este, Churchland considera provável que em pouco tempo as pessoas aprenderiam esse novo vocabulário sobre estados cinéticos, adquirindo facilidade em aplicar *as leis que governam nossas interações e projeções comportamentais* também em atribuições de primeira pessoa, assim efetivamente eliminando do uso cotidiano a Psicologia do Senso Comum e seu arcabouço conceitual de estados mentais e atitudes proposicionais.

Se no primeiro cenário Churchland fala do surgimento de um novo vocabulário

para tratar do que hoje entendemos como atividade cognitiva e ordinariamente explicamos com base na Psicologia do Senso Comum, no segundo cenário ele prevê que avanços na neurociência poderiam resultar no nascimento não apenas de um novo vocabulário, mas até mesmo de uma *linguagem* completamente nova. Em contraste com a hipótese proposta por Noam Chomsky de que o cérebro é composto fundamentalmente por estruturas inatas dedicadas à aquisição de linguagem, Churchland sugere a hipótese de que as estruturas inatas do cérebro têm como "*função original e primária a organização da experiência perceptiva*", a capacidade linguística sendo apenas um subproduto evolutivo adicional e incidental (Churchland, 1981, p. 86). Churchland imagina que pesquisas que confirmassem essa segunda hipótese mostrariam que nossa linguagem natural explora apenas uma parte superficial da maquinaria complexa do cérebro, e assim, uma compreensão melhor da *organização e processamento da percepção* nos permitiria desenvolver uma "*linguagem muito mais sofisticada*" com *estruturas sintáticas e semânticas* alternativas, que refletiriam "*a estrutura subjacente a nossas atividades cognitivas em maior detalhe do que a linguagem natural*" (Churchland, 1981, p. 86). Como ele sugere, esse "*novo sistema de comunicação verbal*" baseado em uma gramática da percepção, uma gramática combinatória complexa, transcenderia os limites impostos pela linguagem ordinária, como os valores de verdade e falsidade ou relações de implicação entre sentenças, e resultaria efetivamente no desaparecimento das atitudes proposicionais e, assim, na eliminação da Psicologia do Senso Comum como um todo.

Os dois primeiros cenários falam de mudanças conceituais que resultariam em um novo vocabulário ou, mais profundamente, em uma nova linguagem. De maneira ainda mais radical, o terceiro cenário propõe que desenvolvimentos científicos poderiam

resultar em uma forma completamente nova de comunicação. Churchland propõe que, assim como os hemisférios de um mesmo cérebro se comunicam de maneira não-verbal, avanços suficientemente complexos na tecnologia neurocientífica nos permitiriam replicar esse tipo de comunicação entre cérebros diferentes. Ele explica que pessoas cujo corpo caloso (a estrutura nervosa que reúne os dois hemisférios do cérebro) é removido cirurgicamente experimentam uma série de "*déficits comportamentais que indicam uma perda de acesso de um hemisfério a informações que ele costumava receber do outro* [hemisfério]"; entretanto, em pessoas com *agenesia do corpo caloso* (ausência congênita daquela estrutura) esses efeitos são minimamente percebidos, indicando que "*os dois hemisférios aprenderam a explorar informações transportadas em vias menos diretas que os conectam por entre regiões subcorticais*" (Churchland, 1981, p. 87). Assim, Churchland sugere que, da mesma maneira que um cérebro em desenvolvimento aprende a explorar vias alternativas para transmissão de informações entre os dois hemisférios, seria possível que, munidos da estrutura física necessária, dois cérebros distintos também pudessem aprender a trocar informações. Ele propõe que um transdutor instalado em alguma região apropriada do cérebro poderia transformar a atividade neural relevante em micro-ondas a serem transmitidas por algum tipo de antena localizada na testa, a qual também receberia o sinal emitido por outras pessoas e o interpretaria de volta em ativação neural. Com uma tal estrutura, que funcionaria de maneira análoga ao corpo caloso porém permitindo a transmissão de informações entre cérebros distintos, Churchland prevê que nós aprenderíamos uma forma inteiramente nova de comunicação, permitindo uma troca de informações mais rápida e fiel, muito superior às limitações impostas pela comunicação verbal. Nesse cenário, Churchland acredita que as bibliotecas deixariam de conter livros "*com sentenças e argumentos*" e passariam a oferecer acervos de "*gravações de atividade neural*"; indo além, ele afirma

que as pessoas finalmente interpretariam outros indivíduos "*da mesma maneira que o hemisfério direito 'entende' e 'concebe' o hemisfério esquerdo – intimamente e eficientemente, mas não proposicionalmente!*" (Churchland, 1981, p. 88).

Estes três cenários apresentam situações gradativamente mais especulativas em que a Psicologia do Senso Comum seria eliminada, primeiro com o surgimento de um vocabulário alternativo, segundo com o desenvolvimento de um novo tipo de linguagem, e terceiro com a implementação de uma nova forma de comunicação não-verbal. Embora estes cenários pareçam fantasiosos, Churchland obviamente não pretende escrever uma história de ficção científica, mas sim propor possibilidades reais de desenvolvimentos futuros e explorar suas implicações para nossa compreensão ordinária de nós mesmos e dos outros ao nosso redor. Considerando seriamente estas sugestões como afirmações passíveis de comprovação empírica, o fato é que até hoje nenhum cenário similar a estes se tornou realidade, e nosso modo de comunicação continua sendo verbal e nosso vocabulário ordinário fortemente baseado no arcabouço conceitual mentalista da Psicologia do Senso Comum.

2.6: Cognição, Computação, e a Vindicação da PSC

Fodor concorda com Churchland que a Psicologia do Senso Comum é uma teoria; mas, enquanto Churchland argumenta que ela é uma teoria falsa e que deve ser eliminada em favor de uma teoria mais científica a respeito do funcionamento do cérebro e de padrões de ativamento neuronal, Fodor defende que a PSC é uma teoria praticamente correta e perfeitamente integrável às melhores teorias científicas. Como

esse tipo de afirmação requer suporte, Fodor dedicou grande parte de sua carreira a articular uma proposta de conciliação entre o arcabouço conceitual mentalista ordinário e descrições científicas de nosso funcionamento cognitivo. Nesse contexto Fodor se compromete com uma Teoria Representacional da Mente (TRM), articulando mais precisamente um subtipo da TRM, que é a Teoria Computacional da Mente (TCM). Ele propõe que uma "Linguagem do Pensamento" (ou "*mentalês*") organiza informações no cérebro de maneira análoga a uma linguagem de máquina em um computador:

> computadores reais tipicamente usam ao menos duas linguagens diferentes: uma linguagem de input/output na qual eles se comunicam com seu ambiente e uma linguagem de máquina na qual eles se comunicam consigo mesmos (i.e., na qual eles realizam suas computações). (…) Embora a máquina precise de um compilador para usar a linguagem de input/output, ela não precisa de outro compilador para a linguagem de máquina. (…) A linguagem de máquina difere da linguagem de input/output no sentido de que suas fórmulas correspondem diretamente a estados e operações físicas computacionalmente relevantes da máquina.
> (Fodor, 1975, p. 66)

Semelhantemente à linguagem de input/output e à linguagem de máquina no caso de computadores, Fodor sugere que "*o primeiro se aplica a predicados na linguagem natural que as pessoas aprendem e o segundo a predicados na linguagem interna na qual elas pensam*" (Fodor, 1975, p. 66). Essa linguagem do pensamento não seria "aprendida", mas faria parte da configuração mental prévia que nos permite processar informações (possibilitando, assim, qualquer tipo de aprendizado, até mesmo o de nossa língua materna, utilizada para input/output).

Com base nessa linguagem de pensamento e na analogia mente-máquina, Fodor propõe: "*processos intencionais são operações sintáticas definidas sobre representações mentais*" (Fodor, 2000, p. 1), e mais ainda, "*a ação que um agente realiza é a consequência de computações definidas sobre representações de ações*

possíveis" (Fodor, 1975, p. 31). Nessa perspectiva, o mesmo se aplica a atitudes proposicionais e estados mentais em geral, os quais "*são relações entre organismos e representações internas*" (Fodor, 1975, p. 198), de tal maneira que "*acreditar que P é estar em uma determinada relação à sentença em mentalês que significa que P*" (Crane & McLaughlin, 2009, p. 212). Mas, apesar de dotados de valor semântico (significado), "*estados mentais causalmente relacionados sucedem uns ao outros de acordo com princípios computacionais que se aplicam formalmente às representações*" (Fodor, 1975, p. 198); isso quer dizer que, embora "*transições computacionais de uma sentença em mentalês a outra sentença em mentalês façam* (…) *sentido dado o significado das sentenças*", tais sentenças na verdade "*participam de processos computacionais apenas em virtude de suas características sintáticas e formais*" (Crane & McLaughlin, 2009, p. 212), exatamente como ocorre nos processamentos de computadores reais.

O projeto de Fodor, portanto, é o de explicar como os conceitos e a ontologia da Psicologia do Senso Comum que as pessoas utilizam em suas explicações cotidianas podem ser compreendidos e incorporados a uma perspectiva científica; assim, sua teoria sobre nossa arquitetura cognitiva visa explicar como é possível que tenhamos estados mentais intencionais e que interpretemos o comportamento de outras pessoas atribuindo a elas atitudes proposicionais:

> A Teoria Representacional da Mente (...) pretende explicar como é possível existirem estados que têm as propriedades semânticas e causais que ordinariamente consideramos que as atitudes proposicionais têm. De fato, a TRM propõe uma teoria de o que as atitudes proposicionais são.
> (Fodor, 1990, p. 5)

E a resposta, para ele, é que "*a maneira como informação é armazenada, computada, aceita, rejeitada ou processada de algum outro modo pelo organismo explica seus estados cognitivos e, particularmente, suas atitudes proposicionais*", de tal maneira que

alguns processos orgânicos "*satisfazem descrições como 'armazenar, aceitar, rejeitar, computar, etc., P' e que o organismo aprende, percebe, decide, lembra, acredita, etc., o que quer que seja porque ele armazena, aceita, rejeita, ou computa o que quer que seja*" (Fodor, 1975, p. 75).

É com essa perspectiva computacional da mente que Fodor pretende vindicar a Psicologia do Senso Comum. Ele reconhece que alguns cientistas cognitivos têm postulado "*processos mentais surpreendentes a torto e a direito*", mas que o cerne da abordagem que ele propõe é o reconhecimento da "*facticidade das atribuições de atitudes proposicionais a organismos e a consequente necessidade de explicar como organismos chegam a ter as atitudes a proposições que eles têm*" (Fodor, 1975, p. 198).

2.7: Conclusão

O debate entre Churchland e Fodor acerca da Psicologia do Senso Comum tem como centro a discordância dos dois filósofos no que diz respeito à adequação do nosso arcabouço conceitual intencional ordinário em relação ao conjunto de nossos conhecimentos científicos, com Churchland de um lado defendendo que a PSC e sua terminologia mentalista serão finalmente eliminadas pela neurociência, e Fodor de outro lado defendendo que elas deverão ser consideravelmente integradas às ciências cognitivas. No desenvolvimento das posições de ambos, fica evidente a influência do contexto intelectual apresentado no capítulo anterior, como no caso das discussões científicas e filosóficas sobre o recurso a estados mentais representacionais na psicologia e sua definição como ciência do comportamento (behaviorista) ou da mente;

sobre relações interteóricas e a unidade do conhecimento científico, no contexto do positivismo lógico; sobre abordagens reducionistas e a identidade de estados mentais com estados neurais em oposição a uma abordagem funcionalista da mente; e por fim, sobre avanços neurocientíficos e a analogia mente-máquina (em contraste com a abordagem computacional clássica adotada por Fodor, Churchland viria posteriormente a investir no modelo computacional *conexionista* ou de *processamento paralelo e distribuído* como instanciando sua proposta que privilegia padrões de ativamento na rede neuronal).

Vale ressaltar que o foco nestas diversas questões do contexto intelectual só foi possível graças à concordância dos dois filósofos no que diz respeito à natureza da Psicologia do Senso Comum como uma teoria que as pessoas empregam no dia a dia para entender, explicar e prever o comportamento delas mesmas e de outras pessoas. Parte central dessa perspectiva *teórica* da Psicologia do Senso Comum é sua constituição a partir de generalizações nômicas sobre relações entre comportamentos e estados mentais (e atitudes proposicionais). No entanto, como será discutido no próximo capítulo, a ideia de que a Psicologia do Senso Comum é uma teoria *porque é composta de relações nômicas entre comportamentos e estados mentais* só pode ser aceita se tivermos bons motivos para acreditar que ela de fato é composta por tais leis ou esboços de leis; isto é, a questão sobre se as pessoas de fato baseiam seus julgamentos ordinários em generalizações como as que Churchland e Fodor sugerem é uma questão empírica, e não pode ser assumida *a priori*.

Um último comentário sobre a proposta eliminativista de Churchland é que ele é ambíguo quanto ao caráter de sua tese, se de teor descritivo ou normativo. Em alguns

momentos ele claramente se porta como um "profeta" que antevê a (supostamente) inevitável eliminação da Psicologia do Senso Comum, primeiro afirmando que o arcabouço conceitual mentalista não se integrará à neurociência e, segundo, afirmando que avanços na neurociência resultarão (ao menos) em um novo vocabulário sobre estados e transformações neurais (vocabulário este que será adotado não apenas em círculos científicos, mas também por leigos). Interpretando seu eliminativismo materialista desta maneira como uma tese *descritiva*, uma previsão do que acontecerá, a única conclusão possível é que, ao menos até hoje, nenhuma de suas afirmações foi confirmada e a Psicologia do Senso Comum permanece estável e guiando nosso entendimento interpessoal. Em sua defesa, porém, Churchland argumenta que suas previsões são a respeito da *neurociência completa*, e que ainda estamos longe desse desenvolvimento científico avançado. Por outro lado, em outros momentos Churchland parece propor uma tese *prescritiva* ou *normativa*, que visa afirmar não um *estado de coisas futuro*, mas um estado de coisas *ideal*, defendendo a importância de que a Psicologia do Senso Comum seja eliminada e de que formulemos teorias científicas que transcendam seu arcabouço conceitual. Em ambos os casos, a posição de Fodor é que a eliminação da PSC não ocorrerá (contradizendo a leitura descritiva) porque tal eliminação é de fato desnecessária (contradizendo a leitura normativa).

CAPÍTULO 3:

Transição Conceitual

No capítulo 2 caracterizamos o debate filosófico acerca da Psicologia do Senso Comum conforme articulado por Paul Churchland e Jerry Fodor, o primeiro propondo o Materialismo Eliminativo pela eliminação da PSC em favor de teorias neurocientíficas do comportamento, e o segundo defendendo a integridade do arcabouço conceitual mentalista ordinário e sua integrabilidade com a ciência cognitiva por meio da formulação de uma teoria computacional da mente fundamentada na linguagem do pensamento. Neste capítulo examinaremos a transição conceitual que ocorre nos anos 80 e 90, transformando de tal maneira a perspectiva filosófica sobre a Psicologia do Senso Comum que justifica a classificação do debate contemporâneo (Capítulo 4) como distinto daquele inicial entre Churchland e Fodor (Capítulo 2).

3.1: Tarefas de Crença Falsa

Uma inovação experimental importante na década de 1980 que viria a influenciar a literatura filosófica sobre a Psicologia do Senso Comum foi a criação dos "testes de crença falsa" ou "tarefas de crença falsa" pelos psicólogos Heinz Wimmer e Josef Perner. No experimento original, os participantes são crianças entre três e nove anos de idade que observam a seguinte cena:

> um protagonista coloca um objeto em local *x* e então [os participantes] veem que na ausência do protagonista o objeto é transferido de *x* para *y*. Como a transferência é uma surpresa, eles [os participantes] tinham que assumir que o protagonista ainda acreditava que o objeto estava em *x*. Os participantes tinham que indicar onde o protagonista irá

procurar o objeto quando retornar.
(Wimmer & Perner, 1983, p. 103)

Como exemplo das diversas variações replicadas posteriormente, a criança observa um cenário como o este: uma pessoa (o protagonista) recebe um objeto pequeno (e.g., um doce qualquer, como um chocolate) e tem a opção de esconder esse objeto em uma de duas caixas que estão sobre uma mesa; digamos que essa pessoa escolhe esconder seu chocolate dentro da caixa número 1. A pessoa então sai de cena, e o experimentador retira o chocolate da caixa 1 e, sem que a pessoa em questão possa saber, esconde-o na caixa 2. A pessoa então volta à sala, e o experimentador pergunta para a criança onde ela acha que essa pessoa irá procurar seu chocolate, se na caixa 1 (onde ela havia escondido o chocolate) ou se na caixa 2 (onde o chocolate de fato está escondido).

Qualquer adulto em funcionamento cognitivo típico é capaz de responder que a pessoa irá procurar o chocolate na caixa 1, onde ela pensa que ele está, embora ele tenha sido transferido para a caixa 2; ou seja, adultos conseguem compreender que, porque a pessoa não viu a transferência do chocolate da caixa 1 para 2, aquela pessoa não sabe que o chocolate agora se encontra na caixa 2 e por isso ela deve ter uma crença equivocada a respeito da localização do chocolate. De fato, no experimento original de Wimmer e Perner este foi o resultado obtido com as crianças mais velhas. O que o experimento original indica e estudos posteriores suportam, entretanto, é que antes dos quatro anos de idade, crianças são incapazes de reconhecer que o protagonista deve ter uma crença falsa, e assim respondem ao experimentador que o protagonista irá procurar por seu chocolate na caixa 2, onde a criança sabe que o chocolate realmente se encontra. A interpretação mais comum na literatura é que esse resultado indica que só após os quatro anos de idade é que crianças têm uma compreensão representacional de crenças,

ou seja, vendo crenças como representações internas da pessoa observada, as quais podem ser verdadeiras ou falsas independentemente da crença que a própria criança tem. Nas palavras dos autores do estudo original, "*isso mostra que assim que a criança é capaz de representar a crença de outra pessoa, ela também consegue restringir sua interpretação da intenção que a pessoa manifesta às crenças daquela pessoa*" (Wimmer & Perner, 1983, p. 104) – ou seja, quando dotada da capacidade de formar crenças sobre a crença de outra pessoa, a criança passa a interpretar a intenção daquela pessoa (e.g., a intenção de pegar o chocolate) como sendo restringida pelo que a pessoa acredita (no caso, o que a pessoa acredita sobre onde o chocolate está). De maneira geral, o experimento pode ser visto como um "*diagnóstico sobre se uma criança atingiu um estado de maturidade no que diz respeito a componentes-chave do repertório conceitual que constitui nossa psicologia do senso comum*" (Davies & Stone, 1995, p. 3).

3.2: Mecanismos de Leitura Mental

Avanços experimentais como os das tarefas de crenças falsas e outros testes similares fazem parte do surgimento, no decorrer dos anos 80, de uma nova perspectiva no que diz respeito à Psicologia do Senso Comum, que visa elucidar quais são os mecanismos que fundamentam nossa capacidade de atribuir estados mentais às pessoas com quem interagimos no dia a dia. Para tornar explícito o foco no aspecto prático de nosso entendimento interpessoal ordinário, muitos passam a adotar o rótulo "*leitura mental*" (em inglês "*mind-reading*") para designar nossa habilidade de "ler" a mente de outras pessoas. Esse termo não deve ser confundido com atividades paranormais como a telepatia e o acesso sobrenatural ao que os outros pensam; antes, neste contexto o termo

se aplica ao que ocorre em nossa experiência cotidiana quando atribuímos estados mentais às pessoas com quem interagimos. Nesse sentido particular, de acordo com o filósofo Alvin Goldman, "*Ler uma mente é formar um julgamento, crença, ou representação de que uma determinada pessoa está (no passado, presente, ou futuro) num estado mental ou experiência* específica" (Goldman, 2008). Meramente reconhecer que a pessoa está sorrindo, chorando, dizendo algo, ou realizando uma atividade qualquer não é leitura mental: "*Para atribuir um estado mental, o julgamento deve empregar um conceito ou categoria mental*" (Goldman, 2008). A questão, portanto, passa a ser determinar *como* a leitura mental ocorre, isto é, como é que nós chegamos a formar julgamentos a respeito dos estados mentais de outras pessoas.

3.2.1: A "Teoria da Teoria"

Uma resposta possível à pergunta sobre como nós lemos a mente de outras pessoas já estava presente no debate entre Churchland e Fodor sobre a Psicologia do Senso Comum. Como discutido no capítulo anterior, os dois filósofos acreditavam que nós entendemos as pessoas ao nosso redor e somos capazes de explicar e prever o seu comportamento porque temos uma teoria composta de leis e esboços de leis que estabelecem relações entre comportamentos e estados mentais. Essa resposta, pressuposta por Churchland e por Fodor, passa ser identificada como "Teoria da Teoria" (do inglês "Theory Theory") porque é a *teoria* de que atribuímos estados mentais a outras pessoas a partir de uma *teoria* sobre comportamentos e estados mentais.[9]

[9] A origem do nome "Teoria da Teoria" é comumente atribuída a Adam Morton (1980): cf. Baker (1999), Marraffa (2011)

Como explica Goldman (que é contrário a essa posição), a Teoria da Teoria diz que "*nossa capacidade de ler mentes é implementada por uma teoria intuitiva da mente, um corpo de conhecimento ou crenças a respeito das relações causais entre estados mentais de um lado, e comportamento, ambiente e outros estados mentais de outro lado*" (Goldman & Mason, 2006). Tanto Churchland quanto Fodor propõem que nosso entendimento interpessoal e nossa habilidade de atribuir estados mentais a outras pessoas ("ler" a sua mente) fundamenta-se na aplicação de uma teoria que relaciona nomologicamente comportamentos e estados mentais. Esta teoria (que eles acreditam que as pessoas usam para ler a mente dos outros) é, portanto, o que eles entendem como a própria Psicologia do Senso Comum, isto é, o "*arcabouço conceitual pré-científico que todos os humanos normalmente socializados empregam para compreender, prever, explicar e manipular o comportamento de outros humanos*", o qual inclui conceitos como "*crença, desejo, dor, prazer, amor, ódio, alegria, medo, suspeita, memória, reconhecimento, raiva, simpatia, intenção etc.*" (Churchland, 1996, p. 3). O ponto é que tanto Churchland quanto Fodor podem ser vistos como adeptos da Teoria da Teoria precisamente porque adotam uma *teoria* segundo a qual a leitura mental ocorre por meio da aplicação de uma *teoria*.

3.2.2: A "Teoria da Simulação"

Uma outra resposta à pergunta sobre como nós formamos julgamentos a respeito dos estados mentais de outras pessoas surge na segunda metade da década de 1980. De maneira independente, dois pesquisadores publicam no ano de 1986 críticas à Teoria da Teoria, propondo que não é a aplicação de uma teoria o que fundamenta nosso

entendimento de outras pessoas, mas sim processos de "replicação" (Heal, 1986) ou "simulação" (Gordon, 1986); o termo preferido por Gordon acaba sendo aceito mais amplamente, e a teoria alternativa à Teoria da Teoria é conhecida como "Teoria da Simulação".

O que a Teoria da Simulação propõe é que, em vez de recorrermos a um corpo teórico a respeito da mente, as nossas atribuições de estados mentais a outras pessoas são resultado de processos em que simulamos estar na situação da pessoa observada e inferimos os estados mentais desta pessoa com base nos estados que teríamos se estivéssemos naquela situação. Processos de simulação podem ser identificados, no nível pessoal, como "*imaginativamente 'nos colocarmos no lugar da outra pessoa'*" (Gordon, 2009), mas também podem ocorrer em um nível sub-pessoal, quando simplesmente reagimos de maneira intuitiva e inconsciente como se de fato estivéssemos no lugar da outra pessoa. No primeiro caso, que Alvin Goldman chama de "simulação de nível alto" (*"high level" simulation*), nós deliberadamente tentamos entender como seria estar na situação da pessoa observada, talvez até mesmo teatralizando mentalmente uma cena, e como resultado "*podemos detectar suas atitudes proposicionais*"; já no segundo caso, que Goldman identifica como "simulação de nível baixo" (*"low level" simulation*), processos "automáticos" ou "involuntários" permitem que nos identifiquemos com a outra pessoa, detectando suas emoções e dor, e às vezes até mesmo experimentando reações semelhantes às que observamos (Goldman & Mason, 2006).

Uma forma de explicar esses processos de simulação em termos cognitivos é como a utilização de nosso "*sistema de controle de comportamento*" como um "*modelo*

manipulável de outros sistemas" (isto é, um modelo do sistema de controle de comportamento de outras pessoas); o que torna o processo de simulação diferente de outros usos do nosso s*istema de controle de comportamento* é que, primeiramente, o sistema é desligado (*off-line*) para que o resultado (*output*) não seja um comportamento "*mas apenas previsões ou expectativas de comportamento*" (Gordon, 2009); como o sistema de controle de comportamento opera "off-line", é possível que o processo de simulação gere uma "*tomada de decisão vicária*" na qual o observador efetua uma escolha própria como se estivesse na situação observada, mas, no entanto:

> mais tipicamente [a simulação] para no objetivo mais modesto de estabelecer quais opções seriam mais atrativas (de modo que não ficaríamos surpresos ao ver a outra pessoa preferindo uma delas) e quais não são atrativas (de maneira que ficaríamos surpresos ao ver a pessoa preferindo uma delas).
> (Gordon, 2009)

A Teoria da Simulação logo receberia suporte empírico por parte da neurociência com a descoberta dos "neurônios espelho":

> No começo dos anos 90 nosso grupo de pesquisa na Universidade de Parma na Itália (…) [descobriu algo] acidentalmente em uma classe surpreendente de neurônios no cérebro do macaco que disparam quando um indivíduo realiza uma simples ação motora direcionada a um objetivo, como pegar um pedaço de fruta. A parte surpreendente foi que estes mesmos neurônios também disparam quando o indivíduo observa outro realizar a mesma ação. Como esse subgrupo de células recém-descobertas parecia diretamente refletir ações realizadas por outrem no cérebro do observador, nós os chamamos de neurônios espelho.
> (Rizzolatti, Fogassi, & Gallese, 2006)

Ou seja, este mesmo grupo de neurônios é ativado não apenas quando o macaco realiza uma ação motora, mas também quando "*observa outros indivíduos (macacos ou humanos) executar ações similares*" (Gallese, 2007). A descoberta de sistemas neurais de espelhamento semelhantes em humanos logo motivou interpretações mais gerais, como esta de Goldman:

> Neurônios espelho são uma classe de neurônios que disparam quando um indivíduo (macaco, humano, etc.) experimenta um determinado evento mental ou cognitivo endógeno e também quando ele observa um sinal de que outro indivíduo experimenta ou está prestes a experimentar o mesmo tipo de evento mental ou cognitivo.
> (Goldman, 2008).

A proposta de Goldman e outros é que estas descobertas dão embasamento à Teoria da Simulação, sugerindo que a leitura mental é realizada por meio de processos simulatórios – ou ainda, processos de espelhamento (Gallese & Goldman, 1998). Nessa interpretação, neurônios-espelho proveem uma ponte entre indivíduos distintos, permitindo "*uma experiência interna direta, e portanto a compreensão da ação, intenção ou emoção de outra pessoa*" (Rizzolatti, Fogassi, & Gallese, 2006).

São diversos os exemplos de situações em que processos simulatórios podem contribuir com a nossa compreensão dos outros: ao ver alguém se machucar, nós às vezes sentimos algo parecido com a dor que observamos, ainda que em escala bem menor; ou ao ver em um filme uma aranha subindo pelo corpo de um personagem, nós às vezes temos aquela sensação estranha de que há algo subindo pelo nosso próprio corpo; "*os neurônios-espelho de um observador disparam quando ele vê uma faca afiada ser aplicada ao corpo de outra pessoa*" (Goldman, 2008); e assim por diante. O argumento dos proponentes da teoria da simulação (dentre os quais destacadamente Goldman) é que exemplos como esses podem envolver processos de espelhamento neuronal, enquanto outros (de nível alto) envolvem o uso da imaginação explícita de "como seria estar na situação que observamos". O que essa diversidade mostra, portanto, é que não há relação de identidade entre espelhamento e simulação: o espelhamento é visto como um mecanismo que permite a simulação, mas não é necessariamente o único; assim, enquanto a existência de sistemas de espelhamento dá

suporte à Teoria da Simulação, não é requerido que todo e qualquer processo de simulação envolva tais sistemas de espelhamento.

3.3: Stephen Stich e o "Despertar de um Sono Dogmático"

Ao lado de avanços experimentais como a criação das tarefas de crença falsa e do desenvolvimento da Teoria da Simulação em oposição à Teoria da Teoria, outro fator central na transição conceitual que marca a literatura filosófica sobre a Psicologia do Senso Comum é presente na figura do filósofo Stephen Stich.

Tanto Churchland quanto Fodor se mantiveram consideravelmente fiéis a suas posições no decorrer de suas carreiras, de forma que em seus vários artigos e livros encontramos poucas mudanças reais, mas principalmente a defesa de suas propostas contra os argumentos de críticos. Ao contrário de Churchland e Fodor, entretanto, Stich notoriamente iniciou sua carreira como defensor ferrenho do eliminativismo, mas, ao encontrar contra-argumentos suficientemente fortes, passou a defender a Psicologia do Senso Comum. Nesta reorientação, Stich introduziu à literatura uma distinção conceitual entre um sentido *internalista* e um sentido *externalista* de "teoria da Psicologia do Senso Comum". À luz desta distinção foi possível ver que o debate entre Churchland e Fodor partia do pressuposto de que a PSC é uma teoria protocientífica (e portanto sujeita a corroboração científica) internamente armazenada por meio da qual as pessoas supostamente fazem suas habituais atribuições de estados mentais; entretanto, tal pressuposto é questionável, e não é suficientemente defendido pelos dois autores, o que serviu de motivação para Stich articular a sua própria teoria da PSC, unificando em

uma “teoria híbrida” as correntes da “Teoria da Teoria” e da “Teoria da Simulação”. Nas próximas subseções veremos estes desenvolvimentos, tratando por fim, na seção final deste capítulo, de outros avanços nas ciências cognitivas que viriam a influenciar o debate contemporâneo.

3.3.1: Fase Eliminativista

A primeira fase da obra de Stich no que diz respeito à Psicologia do Senso Comum é articulada em seu livro *From Folk Psychology to Cognitive Science: The Case Against Belief* (1983). Neste livro Stich defende que a noção ordinária de crença e as ciências cognitivas são absolutamente incompatíveis entre si, e que, portanto, apenas uma delas deve sobreviver. Influenciado por sua formação na filosofia da linguagem, Stich analisa a noção de crença (e de atitudes proposicionais em geral) a partir de teorias de individuação de conteúdo sentencial, buscando elucidar “*o que queremos dizer de uma pessoa quando atribuímos a ela uma crença*” (Stich, 1983, p. 6); grande parte de sua preocupação, portanto, dizia respeito à *denotação* ou *extensão* dos predicados mentalistas que formam nossas explicações ordinárias. Mas, além de considerações sobre linguagem e significado, Stich pretende determinar se os conceitos intencionais da Psicologia do Senso Comum podem ser úteis para teorias das ciências cognitivas. Ele examina – e rejeita – a Teoria Representacional da Mente, formulando em contrapartida a “Teoria Sintática da Mente”, segundo a qual “*relações causais entre estados cognitivos espelham relações formais entre objetos sintáticos*” (como a relação gramatical entre palavras de uma frase, levando em conta apenas função sintática, e não a função semântica, isto é, de significado). Stich considera este um “*paradigma mais*

adequado para as ciências cognitivas" (Stich 1983, p. 149) do que o paradigma defendido por Fodor; o resultado de sua proposta, porém, é a rejeição do arcabouço conceitual da Psicologia do Senso Comum.

A conclusão eliminativista de Stich nesta primeira fase é motivada por alguns pontos principais: primeiro, concordando com Churchland, ele aponta que a Psicologia do Senso Comum é um programa de pesquisa degenerativo, sem grandes descobertas e progresso nos últimos séculos; segundo, para ele o caráter essencialmente semântico da PSC (ou a centralidade da intencionalidade) indica sua inaplicabilidade a pessoas que não sejam plenamente desenvolvidas em seu exercício linguístico (como é o caso de crianças ou de adultos "excepcionais"); terceiro, ele argumenta que embora uma multiplicidade de estados mentais sejam descritos em culturas diferentes, não é razoável esperar que todos eles sejam necessariamente compatíveis com os estados previstos pela PSC, e que variações culturais possam resultar em arcabouços conceituais diferentes daquele criticado por Churchland e defendido por Fodor.

3.3.2: A "Teoria" da PSC: Internalismo e Externalismo

Quase dez anos depois, em "*Deconstructing the Mind*" (1996), Stich compara o desenvolvimento de uma posição filosófica com a atividade de um tecelão: "*quando as coisas vão bem, cada fio de um argumento agrega força e suporte aos outros, e gradualmente padrões interessantes surgem. Mas quando as coisas vão mal – quando um dos fios se rompe – a trama inteira parece se desmanchar. Uma brecha pequena se torna grande, e no fim não há mais nada*" (Stich, 1996, p. 3). Com essa metáfora Stich

aponta para sua experiência pessoal de perceber um furo argumentativo na posição eliminativista: da primeira premissa de que crenças, desejos e outros estados mentais fazem parte da teoria da Psicologia do Senso Comum, e da segunda premissa de que a PSC é uma teoria seriamente equivocada e incoerente, não se segue a conclusão de que crenças, desejos e os estados mentais em questão de fato não existam. Stich descreve a percepção deste erro argumentativo como "o despertar de um sono dogmático", após o qual ele rejeita o eliminativismo como infundado e reorienta sua pesquisa em direção a uma nova avaliação da Psicologia do Senso Comum.

Considerações práticas sobre *como* a Psicologia do Senso Comum ocorre (isto é, sobre como nós chegamos a formar julgamentos a respeito dos comportamentos e dos estados mentais de outras pessoas e de nós mesmos) levam Stich a uma nova compreensão ontológica de *o que* a Psicologia do Senso Comum é, isto é, de como devemos interpretar sua natureza. Em um artigo intitulado "What Is Folk Psychology?", escrito com Ian Ravenscroft, originalmente publicado em 1994 e depois incluído em seu livro *Deconstructing the Mind*, de 1996, Stich e Ravenscroft propõem que há sentidos diferentes com que a Psicologia do Senso Comum pode ser entendida como uma *teoria*. Em uma leitura *internalista* de "teoria", a PSC é entendida como uma teoria no sentido de ser uma "*'estrutura de conhecimento' internamente representada que é utilizada pelo mecanismo cognitivo que fundamenta nossas capacidades interpretativas ordinárias*" (Stich & Ravenscroft, 1996, p. 132); nessa interpretação internalista, a teoria ou "estrutura de conhecimento" da PSC pode ser armazenada em "*um sistema que consiste de regras e princípios ou generalizações sentenciais*" ou mesmo em um "*modelo mental ou rede conexionista que não permite mapeamento a um grupo de proposições*" (Stich & Ravenscroft, 1996, p. 132) – a primeira possibilidade sendo mais próxima da

proposta de Fodor, e a segunda da perspectiva de Churchland; em ambos os casos, a despeito de *como* esta tal teoria é armazenada, a interpretação internalista pode ser identificada com a Teoria da Teoria uma vez que propõe que a leitura mental ocorre por meio da aplicação de uma teoria internamente armazenada – sendo assim a posição que a Teoria da Simulação pretende criticar.

A grande inovação conceitual de Stich é o reconhecimento de que a PSC pode ser interpretada como teoria também em um sentido *externalista* de "teoria", por um lado designando o "*conjunto de generalizações que as pessoas automaticamente reconhecem e aceitam*", e, por outro lado designando "*uma teoria que sistematiza essas generalizações ordinárias de maneira perspícua*" (Stich & Ravenscroft, 1996, p. 132). O primeiro caso é compatível com a Teoria da Simulação: mesmo se é verdade que lemos a mente de outras pessoas por meio de processes simulatórios e não pelo uso de uma teoria internamente armazenada, "*teóricos da simulação não negam o fato óbvio de que as pessoas têm intuições a respeito de platitudes psicológicas ordinárias*" (Stich & Ravenscroft, 1996, p. 134), e que portanto podem invocar tal corpo de informações culturais como input parcial de seu sistema de controle de comportamento. No segundo caso externalista, a "sistematização perspícua" diz respeito às várias teorias diferentes que filósofos, psicólogos e cientistas cognitivos, dentre outros, formulam para explicar de forma explícita como eles interpretam as generalizações que julgam fazer parte de nossa prática interpretativa ordinária; em outras palavras, nesta última leitura externalista, tanto a Teoria da Teoria quanto a Teoria da Simulação tratam-se de "teorias" da Psicologia do Senso Comum, isto é, sistematizações conceituais que pesquisadores propõem para explicar a Psicologia do Senso Comum (as quais podem ser verdadeiras ou não sobre o que realmente acontece em nosso entendimento

interpessoal e nossas interações cotidianas).

Em outras palavras, essa diferenciação entre interpretações internalistas e externalistas deixa claro que dizer "*a teoria da Psicologia do Senso Comum*" pode significar, no sentido internalista, "a teoria que as pessoas rotineiramente utilizam de maneira implícita para interpretar o comportamento de outras pessoas" ou então, no sentido externalista, "a teoria filosófica que sistematiza nossas intuições mentalistas ordinárias". Assim, a afirmação eliminativista de que "*a Psicologia do Senso Comum é uma teoria falsa*" é ambígua, podendo ser entendida como a afirmação de que "a teoria internamente representada que fundamenta nossas interpretações comportamentais ordinárias é falsa" ou ainda como a afirmação de que "a teoria filosófica que sistematiza nossas intuições mentalistas ordinárias é falsa"; no primeiro caso, se a Teoria da Simulação é provada verdadeira, o eliminativismo passa a ser vazio, já que afirma a falsidade de uma teoria internalizada que nem sequer existe; no segundo caso, se o eliminativismo afirma que a sistematização filosófica de nossas intuições ordinárias é equivocada, podemos superar o eliminativismo simplesmente formulando uma sistematização conceitual mais adequada sobre nosso entendimento interpessoal ordinário. Essa conclusão foi o que motivou Stich a abandonar de vez seu projeto eliminativista; mas mais do que mera mudança pessoal, essa conclusão também levou a uma transição conceitual que motivou as abordagens filosóficas contemporâneas da Psicologia do Senso Comum que veremos nos último capítulo desta dissertação.

3.3.3: A Teoria Híbrida

A partir de sua conclusão a respeito dos diferentes sentidos internalistas e externalistas de "teoria" que são utilizados nesta literatura, Stich vê a necessidade de formular uma sistematização filosófica da Psicologia do Senso Comum que seja empiricamente adequada. Rejeitando as descrições práticas tanto da Teoria da Teoria quanto da Teoria da Simulação, mas baseando-se em aspectos de ambas, ele propõe uma abordagem *híbrida* ao fenômeno de atribuição de estados mentais a outras pessoas (ou "leitura mental"): "*A leitura mental é um fenômeno complexo e multifacetado, muitas vezes melhor explicado por uma abordagem rica de informações, e muitas outras vezes melhor explicado pela simulação*" (Stich & Nichols, 2003, p. 251). Visando substanciar essa abordagem híbrida, Stich e seu ex-aluno e colaborador Shaun Nichols oferecem alguns exemplos de casos em que a leitura mental é mais claramente o resultado de processos simulatórios e outros casos em que simulação não é suficiente e um corpo de informações deve fundamentar a leitura mental.

A *previsão de inferências* é um fenômeno que parece mais plausivelmente envolver processos simulatórios, principalmente no caso de inferências não-demonstrativas (isto é, inferências indutivas, e não inferências dedutivas em que a conclusão é a consequência lógica das premissas). Stich e Nichols dão o seguinte exemplo: imagine que Fred ouve no rádio uma breve notícia dizendo que o presidente dos Estados Unidos renunciou; quem Fred acha que assumirá em seu lugar? A resposta dos autores é:

> Nós rapidamente geramos a previsão de que Fred acha que o vice-presidente se tornará presidente. Nós sabemos perfeitamente bem, e assim supomos que Fred também sabe, que há muitas maneiras em que essa inferência poderia estar errada. O vice-presidente poderia ser

> assassinado; o vice-presidente poderia renunciar antes de tomar posse; um escândalo poderia levar à remoção do vice-presidente; poderia haver um golpe de estado. É fácil imaginar histórias em que o vice-presidente não se tornaria o novo presidente. Ainda assim, nós prevemos as inferências não-demonstrativas de Fred sem hesitação. E na maioria dos casos como esse, nossas previsões estão corretas.
> (Stich & Nichols, 2003, p. 244)

Como Stich e Nichols explicam, esse tipo de caso é mais facilmente explicável com base em leitura mental por meio de simulação do que com base na teoria da teoria: a única explicação para o sucesso de nossas atribuições de inferências, na Teoria da Teoria, seria *mais informação* – isto é, que nós de alguma maneira adquirimos uma teoria notadamente completa sobre como as pessoas raciocinam –, mas exigir isso não parece razoável. Antes, eles argumentam que uma analogia entre a previsão de inferências e a previsão de intuições gramaticais deixa o ponto mais evidente. Em um experimento em que temos que prever as intuições gramaticais de uma outra pessoa que fala o mesmo idioma que nós, a Teoria da Teoria requereria que tenhamos uma teoria sobre "*os processos que fundamentam a produção de intuições gramaticais em outras pessoas*"; uma hipótese mais simples é que "*nós contamos com nossos próprios mecanismos para gerar intuições linguísticas, e uma vez que tenhamos determinado nossa própria intuição sobre um caso particular, nós a atribuímos [à outra pessoa]*" (Stich & Nichols, 2003, p. 244). O mesmo, segundo eles, se aplica ao caso da previsão de inferências não-demonstrativas, como no exemplo de Fred: em vez de contar com um complexo corpo de informações a respeito das intuições de outras pessoas e como elas raciocinam, é mais simples assumir, em casos gerais, que nós simulamos a inferência e só então atribuímos a conclusão à outra pessoa.

Em contraste com a previsão de inferências, a *atribuição de desejos* parece ser um fenômeno que requer mais do que processos de simulação. Stich e Nichols afirmam:

> Existem vários processos que podem gerar crenças a respeito dos desejos da pessoa observada. Em alguns casos usamos informações a respeito do comportamento verbal e não-verbal da pessoa (incluindo suas expressões faciais) para determinar o que ela quer. Em outros casos nós atribuímos desejos com base no que outras pessoas dizem sobre a pessoa. E muito provavelmente uma variedade de indicações [ou dicas] e fontes de dados são usados no processo de atribuição de desejos.
> (Stich & Nichols, 2003, p. 246)

De maneira geral, de acordo com Stich e Nichols, as atribuições de desejo feitas com base em simulação não são bem sucedidas: não basta imaginar o que desejaríamos no lugar da pessoa, nem mesmo "espelhar" seu estado mental com base nos inputs comportamentais – nós precisamos de *informações*. Um exemplo cotidiano é de quando vamos comprar um presente de aniversário para um amigo: não adianta pensar "se fosse meu aniversário eu gostaria de ganhar X"; antes é preciso consultar uma série de informações que temos a respeito da pessoa para concluir "ele provavelmente preferiria ganhar Y". Outro exemplo da vida real que aponta para a necessidade de informações na atribuição de desejos se dá com crianças: ao observar um adulto que aparenta estar triste ou doente, uma criança poderá oferecer ao adulto algum objeto de conforto que ela tem (como um cobertor ou um brinquedo); a suposição que a criança faz é que assim como ela deseja aquele objeto em um momento de sofrimento, o adulto também o fará. Essa conclusão é obviamente equivocada, e pode ser evitada com mais informações a respeito da mente de um adulto, o que pode lhe trazer conforto, etc. Este último exemplo poderia ser interpretado como sugerindo que crianças são mais propensas a depender de processos de simulação; mesmo se esse for o caso, o ponto para Stich e Nichols é que nossas atribuição de desejos são mais corretas quando temos à nossa disposição um corpo de informações do que quando contamos apenas com a possibilidade de simular e atribuir aos outros o que nós desejaríamos naquela situação.

Em suma, a Teoria Híbrida proposta por Stich e Nichols combina aspectos da Teoria da Teoria e da Teoria da Simulação, e sugere que tipos diferentes de leitura mental dependem de processos cognitivos diferentes: como visto nos exemplos acima, os "*mecanismos de detecção de desejos são ricos de informação*", enquanto "*as estratégias usadas para a previsão de inferências e planos são semelhantes às estratégias simulatórias paradigmáticas*" (Nichols & Stich, 2003, p. 149). Assim, eles concluem, nossas atribuições de estados mentais ocorrem graças a uma combinação de informação e de simulação: "*A leitura mental é um fenômeno complexo e multifacetado, e alguns aspectos da leitura mental devem ser baseados em processos informativamente pobres e de caráter simulatório, enquanto outros aspectos são baseados em processos plenos de informação*" (Stich & Nichols, 2003, p. 243).

Um ponto interessante nestas diferentes teorias de leitura mental é que, ao darem suas explicações de como nós lemos a mente de outras pessoas (ou seja, leitura mental de terceira pessoa), tanto a Teoria da Simulação quanto a Teoria Híbrida revigoram discussões a respeito do conhecimento que temos de nossos próprios estados mentais (ou seja, a "leitura mental de primeira pessoa"). A explicação da Teoria da Teoria, como ficou claro no debate entre Churchland e Fodor (capítulo 2), era que possuímos uma teoria internalizada (a Psicologia do Senso Comum) que é constituída de generalizações nomológicas sobre estados mentais e comportamentos; nessa perspectiva, nós interpretamos não apenas o comportamento de outras pessoas mas também os nossos próprios a partir dessa teoria. Mas ao rejeitarem essa descrição de nossos processos de leitura mental, alguns proponentes da Teoria da Simulação e da Teoria Híbrida se veem na necessidade de explicar o conhecimento que temos de nossa própria mente. Para Nichols e Stich a resposta é clara: é muito mais plausível que nosso conhecimento de

nossos próprios estados mentais provêm de "*um mecanismo de monitoramento que é independente dos mecanismos que nos permitem detectar os estados mentais de outras pessoas e raciocinar sobre estados mentais*" (Nichols & Stich, 2003, p. 199); assim, sua sugestão é que "*a mente vem pré-dotada de um conjunto de mecanismos dedicados à leitura de nossa própria mente*" (Nichols & Stich, 2003, p. 199). Mas, obviamente, uma discussão dos detalhes dessa proposta requereria pesquisas mais aprofundadas.

3.4: Avanços nas Ciências Cognitivas

Também importantes para a transição conceitual que ocorreu na literatura filosófica sobre a Psicologia do Senso Comum foram os desdobramentos nas ciências cognitivas que sucederam ao contexto intelectual discutido no primeiro capítulo desta dissertação. Nesta seção final, apontarei apenas alguns dos avanços que contribuem para o surgimento da discussão contemporânea a ser examinada no último capítulo.

O período tratado no primeiro capítulo abrangia apenas o início da "revolução cognitiva" e o surgimento das ciências cognitivas como área multidisciplinar de pesquisa. Nesta fase, também chamada de "primeira geração" das ciências cognitivas, era entendimento comum que "*a mente podia ser estudada em termos de suas funções cognitivas, ignorando quaisquer formas em que essas funções emergem do corpo e do cérebro*" (Lakoff & Johnson, 1999, p. 75). Como ficou evidente na discussão dos capítulos anteriores, entre filósofos e cientistas a mente "*era vista metaforicamente como um tipo de programa de computador abstrato que poderia ser rodado em qualquer hardware apropriado*", e tinha-se como princípio que "*o pensamento podia*

ser representado usando sistemas simbólicos formais" (Lakoff & Johnson, 1999, p. 76). Nessa perspectiva, a cognição nada mais era do que o "*processamento de informação dentro da cabeça de uma pessoa*":

> Para que isso acontecesse, a informação tinha de ser representada mentalmente; processos cognitivos podiam então operar sobre essas representações. Subsequentemente, a informação representada tinha de ser traduzida em comandos para o sistema motor, mas isto ocorria depois que o processamento cognitivo já havia encerrado.
> (Bechtel, Abrahamsen, & Graham, 1999, p. 86)

Característico dessa perspectiva sobre a cognição era o *solipsismo metodológico* (Fodor, 1980) como "*justificação teórica para ignorar tanto o mundo exterior quanto o corpo nas ciências cognitivas*" (Bechtel, Abrahamsen, & Graham, 1999, p. 86), considerando ambiente e corpo como, no máximo, provendo as entradas de informação (*inputs*) e recebendo as saídas (*outputs*), mas nunca participando do processamento cognitivo propriamente dito.

3.4.1: "Para baixo" e "Para fora"

Em contraste com essa primeira fase, a "segunda geração" das ciências cognitivas envolveu uma ampliação do conceito de cognição, abrangendo processos cerebrais, corpóreos e ambientais. Essa tendência é o que Bechtel, Abrahamsen, e Graham identificam como um movimento "para baixo" e "para fora": "*a ciência cognitiva está sendo puxada verticalmente para baixo em direção ao cérebro e horizontalmente para fora em direção ao ambiente*" (Bechtel, Abrahamsen, & Graham, 1999, p. 90). Esse movimento "para baixo" e "para fora" se faz evidente em diversos campos de pesquisa das ciências cognitivas, como é o caso da neurociência cognitiva. Desde o fim dos anos 80 e no decorrer dos anos 90, inovações tecnológicas permitiram

uma aceleração no desenvolvimento de *ferramentas experimentais e conceituais* na neurociência cognitiva, como "*a eletrofisiologia de neurônios, imageamento funcional cerebral, manipulações moleculares genéticas, e a computação neuronal*", resultando em novas compreensões de fenômenos tão diversos quanto "*sensação, percepção, tomada de decisão, controle motor, memória, linguagem, emoção, e consciência*" (Albright & Neville, 1999, p. lvii), e mais particularmente na produção de "*novos modelos de aprendizagem, categorização, processamento visual, memória sensorial, e teorias relacionando emoções à cognição*" (Clancey, 2009, p. 26). Um exemplo mais óbvio de como esses avanços "para baixo" nas ciências cognitivas contribuíram para discussões filosóficas sobre a Psicologia do Senso Comum é, conforme já mencionado, a descoberta dos neurônios espelho e especulações sobre o papel que eles exercem na leitura mental.

Na expansão horizontal, "para fora", abordagens ecológicas na psicologia antecedem alguns dos temas que caracterizam o debate contemporâneo a ser examinado no próximo capítulo. A teoria ecológica da visão articulada pelo psicólogo James Gibson em seu livro *The Ecological Approach to Visual Perception* (1979) é considerada pioneira na área, fornecendo uma alternativa direta a perspectivas computacionais da mente como processando representações internas. Mais precisamente, a teoria de Gibson "*desafiava a ideia de que* (…) *o sistema visual tem de* (…) *reconstruir um mundo tridimensional completo a partir de informações especificadas na imagem bidimensional na retina*" (Wilson & Foglia, 2011). Três pontos centrais da proposta de Gibson (seguindo a leitura de Chemero, 2009) são que: primeiro, "*a percepção é direta*" e "*não envolve computações ou representações mentais*", de forma que na percepção nós não precisamos "*acrescentar informações a*

sensações"; segundo, a "*percepção é primariamente voltada à orientação de ações, e não à* [mera] *coleta de informações*", tal que "*nós percebemos para fazer coisas*"; o terceiro ponto se segue desses dois:

> como a percepção não envolve a adição mental de informações a estímulos, mas é suficiente para guiar o comportamento de maneira adaptativa, [segue-se que] toda a informação necessária para guiar o comportamento adaptativo deve estar disponível no ambiente a ser percebido. Assim, (…) a percepção é de *affordances*, ou seja, oportunidades comportamentais ambientais diretamente perceptíveis. (Chemero, 2009, p. 23)

O conceito de *affordance*, introduzido por Gibson, diz respeito a diferentes possibilidades de ação que um mesmo objeto ou estrutura oferece a agentes diferentes, como por exemplo: "*uma árvore florida pode prover alimento a uma abelha mas não a um pássaro; e pode prover abrigo a um pássaro, mas não a uma abelha*" e "*o tronco que oferece um lugar de descanso para um pássaro não o faria para um porco*" (Shapiro, 2011, p. 100-101). Mas, ao contrário do que esses exemplos poderiam fazer parecer, para Gibson affordances não são propriedades do ambiente e nem do indivíduo: affordances "*implicam a complementariedade do animal e do ambiente*" (p. 127), e "*atravessam a dicotomia subjetivo-objetivo e nos ajudam a entender sua inadequação*" (Gibson, 1979, p. 129). Embora recebida de maneira crítica pelos seus contemporâneos no fim dos anos 70, a proposta de Gibson ganharia simpatizantes nas décadas seguintes.

Também contribuindo para o avanço horizontal das ciências cognitivas, em seu livro *The Embodied Mind* (1991), Varela, Thompson e Rosch propõem uma abordagem "*enativa*", segundo a qual "*(1) a percepção consiste em ação perceptualmente guiada*" e "*(2) estruturas cognitivas emergem de padrões sensório-motores recorrentes que tornam possível que ações sejam perceptualmente guiadas*" (Varela, Thompson & Rosch, 1991, p. 173); expressada de maneira um pouco menos críptica, sua proposta

enfatiza a ideia de que "*o mundo da experiência é* (…) *determinado por interações mútuas entre a fisiologia do organismo, seu circuito sensório-motor e o ambiente*", resultando em "*acomplamentos cérebro-corpo-mundo*" (Wilson & Foglia, 2011). Varela, Thompson e Rosch são influenciados pela tradição fenomenológica francesa e propõem uma aproximação com o budismo e a psicanálise, aspectos que fogem do escopo desta dissertação. Ainda assim, sua abordagem é digna de menção como representativa dessa "segunda geração" das ciências cognitivas, na qual "*processos cognitivos humanos são* [vistos como] *inerentemente sociais, interativos, pessoais, biológicos, e neurológicos, o que quer dizer que uma variedade de sistemas desenvolvem e dependem uns dos outros de maneiras complexas*" (Clancey, 2009, p. 11).

3.4.2: Cognição Corporificada, Situada e Estendida

Diretamente relevantes para o debate contemporâneo acerca da Psicologia do Senso Comum são as ideias, surgidas nesta segunda fase da ciência cognitiva, de que a cognição é *corpórea*, *situada* e *estendida*. Embora às vezes resumidas no título de "teoria da cognição corporificada" (em inglês, *theory of embodied cognition*), de acordo com o filósofo Lawrence Shapiro um título mais apropriado seria "o programa de pesquisa da cognição corporificada": não se trata de uma teoria unitária e coesa, já que seus "*compromissos e objetos de estudo* (…) *permanecem relativamente nebulosos*" (Shapiro, 2007, p. 338), mas sim de um "programa de pesquisa" que gira em torno de hipóteses e abordagens relacionadas, cuja diversidade é capturada em teses distintas que qualificam a cognição como sendo *corporificada*, *situada* e *estendida*.

Atese da cognição corporificada (em inglês, *embodied*) afirma, em linhas gerais, que "*a cognição depende não apenas do cérebro mas também do corpo*" (Robbins & Aydede, 2009, p. 3). Essa ênfase na importância do corpo em processos cognitivos prevê que "*sem o envolvimento do corpo tanto na percepção quanto na ação, pensamentos seriam vazios, e atividades mentais não teriam as características e propriedades que têm*" (Wilson & Foglia, 2011). Pesquisas a respeito desse aspecto da cognição têm como foco a contribuição de "*atributos físicos do corpo*" em processos cognitivos, buscando "*explicar o conteúdo da cognição* [a partir] *da natureza do corpo que contém o cérebro*" (Shapiro, 2007, p. 340). Nessa direção, o linguista George Lakoff e o filósofo Mark Johnson notam que, embora muitos de nossos conceitos intuitivos poderiam parecer arbitrários, nossa relação com o mundo é essencialmente metafórica, e que "*a metáfora permite que imagens mentais convencionais de domínios sensório-motores sejam utilizadas nos domínios da experiência subjetiva*" (Lakoff & Johnson, 1999, p. 45). Como exemplo, por que é que dizemos que "os preços *subiram*" ou que eles estão "*altos*" e inversamente que eles "*abaixaram*" ou estão "*baixos*" para indicar variações de quantidade? E mais precisamente, por que não usamos o contrário, isto é "baixo" para quantidades maiores e "alto" para as menores? A ideia de Lakoff e Johnson é que "*nós aprendemos algo novo apenas* (…) *em termos de algo mais familiar*" (Shapiro, 2007, p. 341), construindo cadeias de analogias a partir de nossa experiência corpórea: assim, a experiência (sensório-motora) de empilhar objetos ou de encher um recipiente com algum líquido e observar variações de quantidade e volume tornam intuitiva a relação metafórica de aumento com algo "acima" e diminuição com "abaixo" (Lakoff & Johnson, 1999, p. 51). O mesmo vale para a referência à experiência sensório-motora de tamanho para indicar magnitude, intensidade ou importância: "um problema grande" (objetos grandes costumam ser pesados e por isso

difíceis de carregar), "um grande dia" (objetos grandes e pesados podem exercer força e influência), e assim por diante. Mais um exemplo é o da relação entre tempo e orientação espacial, como quando dizemos que o futuro vem pela *frente*, e as coisas do passado ficam para *trás*, ou utilizamos gestos apontando respectivamente *à nossa frente* ou *às nossas costas*: na experiência sensório-motora de uma caminhada nós saímos do ponto de origem e, com o passar do tempo, ao nos movimentarmos, nos aproximamos do ponto de destino; no decorrer do trajeto, as experiências passadas (uma árvore que avistamos antes) ficam literalmente para trás, enquanto experiências futuras (a casa pela qual iremos passar mais adiante) estão literalmente à frente. Este mesmo exemplo da caminhada explica diversas analogias, como entre distância física e tempo (nas ideias de passado/futuro "próximo" ou "distante"), e entre destinações e propósitos/objetivos ("aonde você quer chegar com isso?"). Assim, na teoria de Lakoff e Johnson, os conceitos básicos sobre os quais construímos nossa compreensão essencialmente metafórica do mundo são integralmente dependentes do "*tipo de corpo que seres humanos possuem e a maneira com que esse corpo interage com o ambiente*" (Shapiro, 2007, p. 341) – sua teoria sendo, por isso, representativa da tese da cognição corporificada.

A tese da cognição <u>situada</u> (em inglês, *situated* ou *embedded*) propõe que nossa "*atividade cognitiva rotineiramente explora estruturas do ambiente natural e social*" (Robbins & Aydede, 2009, p. 3), tendo como foco "*estratégias que organismos empregam para ancorar processos cognitivos no ambiente* (…) *físico, social e cultural*" (Wilson & Foglia, 2011). Exemplos desse fenômeno incluem "*medidas que seres humanos tomam para organizar o ambiente a seu redor de uma maneira que 'facilita' o fardo cognitivo*":

> Seres humanos têm como hábito colocar suas chaves perto da porta para que não precisem se lembrar de onde elas estão; eles colocam pontos de referência para que não precisem memorizar rotas; eles catalogam arquivos alfabeticamente para minimizar a dificuldade da procura. De formas como estas seres humanos podem incorporar a estrutura do ambiente ao seu processamento cognitivo. O ambiente se torna cognitivamente 'amistoso' no sentido de que ele elimina etapas que [de outra maneira] poderiam ser requeridas de tarefas cognitivas. (Shapiro, 2007, p. 343)

Exemplos de pesquisa nesta área incluem o trabalho já mencionado anteriormente na percepção visual e o conceito de *affordance* em Gibson, bem como a pesquisa de Edwin Hutchins sobre as práticas representacionais empregadas por marinheiros para calcular a posição de seu navio a partir da codificação de informações de fontes distintas em mapas diferentes (Hutchins, 1995), e também as pesquisas sobre a cognição *distribuída* no contexto da prática científica, entre membros de um laboratório ou da comunidade científica (Giere, 2002) ou no raciocínio individual a partir de modelos (Nersessian, 2006). Outro caso de sua ocorrência na vida cotidiana inclui, por exemplo, quando tentamos montar algum equipamento sem ler o manual de instruções, apenas vendo quais peças encaixam umas nas outras por meio de tentativa e erro, como se fosse um quebra-cabeça (Kirsh, 2009, p. 299).

A tese da cognição estendida (em inglês, *extended*) sugere que os limites de alguns processos cognitivos podem "*se estender para além dos limites de organismos individuais*" (Robbins & Aydede, 2009, p. 3), de forma que "*aspectos do ambiente físico, social, e cultural de um agente podem* (…) *parcialmente constituir o sistema cognitivo do agente*" (Wilson & Foglia, 2011). Um exemplo clássico de cognição estendida é apresentado por Andy Clark e David Chalmers em seu artigo *The Extended Mind* (1998). Clark e Chalmers examinam um caso normal de crença armazenada na memória biológica: uma mulher chamada Inga fica sabendo de uma exposição

interessante no Museu de Arte Moderna e decide ir; "*ela pensa por um momento e se lembra que o museu fica na Rua 53, então ela anda até a Rua 53 e vai ao museu. Parece evidente que Inga acredita que o museu fica na Rua 53, e que ela acreditava nisso mesmo antes de consultar sua memória*" (Clark & Chalmers, 1998). Em contraste com esse caso normal, eles propõem um caso anômalo, e concluem que não há motivos suficientes para que ambos não sejam considerados equivalentes. O personagem do segundo caso é Otto, que sofre de Alzheimer, e, assim como alguns portadores dessa doença fazem, Otto tem sempre consigo um caderno: "*Quando ele aprende alguma informação nova, ele toma nota neste caderno. Quando ele precisa de alguma informação antiga, ele procura no caderno*"; assim, quando Otto fica sabendo da exibição no Museu de Arte Moderna e decide ir, "*ele consulta seu caderno, que diz que o museu fica na Rua 53, e então ele anda até a Rua 53 e vai ao museu*" (Clark & Chalmers, 1998). A conclusão de Clark e Chalmers é que "*Otto andou até a Rua 53 porque queria ir ao museu e que ele acreditava que o museu fica na Rua 53*", e que "*assim como Inga tinha sua crença mesmo antes de consultar sua memória, parece razoável dizer que Otto acreditava que o museu fica na Rua 53 mesmo antes de consultar seu caderno*" (Clark & Chalmers, 1998); para eles, o caderno de Otto exerce exatamente a mesma função que a memória biológica em outras pessoas (como Inga), e que o fato da informação estar armazenada além dos limites da pele é irrelevante, não devendo nos impedir de conceder a paridade funcional dos processos cognitivos de Inga e de Otto.

Esse exemplo desencadeou uma série de críticas, em grande parte porque Clark e Chalmers dão um passo além e propõem que não são apenas os processos cognitivos que se estendem além dos limites do corpo, mas que a própria *mente* de Otto é estendida

de maneira a compreender seu caderno.[10] Mas esse passo extra não é necessário, sendo possível reconhecer que em circunstâncias bastante comuns artefatos tecnológicos e ambientais podem não apenas contribuir, mas de fato fazer as vezes dos componentes biológicos da memória e do processamento de informações, como quando utilizamos o Google para lembrar da letra de uma música que "está na ponta da língua", ou papel e caneta para efetuar um cálculo complexo que envolve várias etapas.

Em suma, embora controversos e longe de serem estabelecidos empiricamente como *fatos* da cognição, a consideração destes três aspectos *potenciais* da cognição como corporificada, situada e estendida resultam em uma compreensão da "*atividade mental como* [sendo] *dependente da situação e do contexto em que ela ocorre*", quer se trate de "*uma situação ou contexto relativamente local (no caso da corporeidade) ou relativamente global (no caso da cognição situada e estendida)*" (Robbins & Aydede, 2009, p. 3).

[10] Cf., por exemplo, as críticas de Baker, 2009 e Farkas, 2012.

CAPÍTULO 4:

O Debate Contemporâneo

Os desenvolvimentos experimentais e conceituais descritos no capítulo 3 exerceram tamanha influência entre filósofos e cientistas ocupados com a Psicologia do Senso Comum que o debate contemporâneo pouco tem a ver com o debate inicial entre Paul Churchland e Jerry Fodor. A pressuposição (característica do debate Churchland-Fodor) de que nosso entendimento interpessoal decorre da aplicação de uma teoria internalizada é identificada como a *Teoria da Teoria* (TT), a qual é desafiada pela hipótese de que são processos simulatórios que nos levam a atribuir estados mentais a outras pessoas (*Teoria da Simulação*, ou TS). A introdução da distinção entre sentidos *internalistas* e *externalistas* de "teoria" (Stich & Ravenscroft, 1996) motivam a articulação de uma sistematização explícita (teoria no sentido externalista) conciliatória entre TT e TS (a *Teoria Híbrida*, ou TH), segundo a qual algumas instâncias de leitura mental requerem apenas processos simulatórios enquanto outras dependem de um corpo de informações. Muito do debate contemporâneo gira em torno exatamente do contraste entre versões da TT, TS e TH, tendo como foco os mecanismos de leitura mental que nos permitem atribuir estados mentais a outras pessoas, bem como a arquitetura cognitiva requerida para acomodar a implementação inclusiva de processos tão variados quanto a aplicação de corpos conceituais, de um lado, e simulações e espelhamento, de outro.

Apesar da importância dessa discussão sobre a natureza da leitura mental (se essencialmente teórica-informacional, simulatória, ou ambos), neste capítulo final terei como foco um tema ainda mais fundamental que surge no debate contemporâneo sobre

a Psicologia do Senso Comum a partir da influência das teorias da cognição corporificada, situada e estendida – a saber, as questões sobre se a leitura mental é um fenômeno real (isto é, se as pessoas de fato atribuem estados mentais umas às outras) e se ela é necessária para nosso entendimento interpessoal. Esse questionamento fundamental é motivado pela ideia, característica da segunda geração das ciências cognitivas, de que nosso funcionamento cognitivo não se define pelo processamento de representações internalizadas, mas sim pela contribuição mútua do corpo e do ambiente natural e sociocultural nas interações que caracterizam nosso entendimento interpessoal. Embora diversos autores partilhem dessa perspectiva, aqui me limitarei ao exame do eliminativismo de Matthew Ratcliffe como representativo desse movimento, oferecendo por fim minhas críticas à sua proposta e articulando as bases de uma abordagem pluralista em defesa da leitura mental e da Psicologia do Senso Comum (portanto em um sentido particular da PSC, diferente da "defesa" de Fodor). Como ficará claro, esta abordagem pluralista é, em princípio, compatível com diferentes teorias da arquitetura cognitiva humana, dizendo respeito apenas ao *explanandum* (o fenômeno a ser explicado, a saber, a leitura mental), e não ao *explanans* (a explicação do mesmo, sua implementação).

4.1: O Eliminativismo de Matthew Ratcliffe

Adeptos da Teoria da Teoria, da Teoria da Simulação e da Teoria Híbrida podem discordar sobre quais são os mecanismos que nos permitem ler a mente de outras pessoas, mas eles concordam em um aspecto mais básico de suas teorias, uma vez que sustentam que nós entendemos outras pessoas e somos capazes de explicar e prever seus

comportamentos porque, de uma maneira ou de outra, temos a habilidade de ler suas mentes e atribuir atitudes proposicionais a elas. O filósofo britânico Matthew Ratcliffe rejeita essa ideia de que nosso entendimento interpessoal é fruto da leitura mental (ideia esta que ele chama simplesmente de "Psicologia do Senso Comum"); antes, ele propõe um modelo de cognição social inspirado nos aspectos fenomenológicos da experiência intersubjetiva e no programa de pesquisa da cognição corporificada, assim favorecendo uma perspectiva segundo a qual a cognição social é *corporificada*, *situada*, e *distribuída* ou *estendida*.

De maneira geral, Matthew Ratcliffe é influenciado pela proposta de que a cognição não é realizada apenas por processos internos mentais ou cerebrais, mas é uma atividade que envolve todo o corpo e se estende até o ambiente em que um indivíduo se encontra, de maneira que nós somos profundamente marcados por nosso contexto social, cultural, linguístico, político etc., e de tal forma que não existe atividade epistêmica subjetiva absoluta, descontextualizada, não-relacional. Assim, Ratcliffe defende as ideias da cognição *corporificada*, *estendida* e *situada,* que aqui representam enfoques distintos porém complementares. A afirmação de que a cognição é incorporada ou corporificada pode ser identificada com a noção de "propriocepção" (ou seja, da percepção do próprio corpo, sua posição, movimento e orientação) e de que a "*atividade corpórea e a manipulação de certos tipos de ambientes estruturados são integrais à cognição*" (Ratcliffe, 2007, p. 108); já a afirmação da cognição como estendida complementa a citação anterior, sugerindo que "*as estruturas do ambiente com que interagimos são, frequentemente, não apenas auxiliares cognitivos ou acompanhamentos à cognição, mas completos participantes nos processos cognitivos*" (Ratcliffe, 2007, p. 108); por fim, ao afirmar que a cognição é situada, Ratcliffe defende

que nossa atividade cognitiva não é apenas constituída por nosso corpo e artefatos ao nosso redor, mas é inescapavelmente marcada pelo contexto social: "*O entendimento social é inextricável do mundo social. Interpretações dos outros geralmente pressupõem um contexto compartilhado de atividade social*" (Ratcliffe, 2007, p. 86). A partir dessas considerações, Ratcliffe conclui que "*a habilidade de entender e antecipar outras* [pessoas] (…) *é parcialmente situado no ambiente social*" (Ratcliffe, 2007, p. 107) e é "*inseparável da interação com o ambiente social compartilhado*" (Ratcliffe, 2007, p. 118). Ao contrário do que é assumido na literatura da Psicologia do Senso Comum, Ratcliffe argumenta que não precisamos "ler a mente" de outras pessoas para entendê-las, mas dependemos apenas de uma compreensão do contexto social e ambiental em que um determinado encontro interpessoal ocorre:

> Muito de nossa habilidade para entender e interagir uns com os outros depende não da Psicologia do Senso Comum, mas de uma compreensão compartilhada do entrelaçamento de papéis sociais, normas sociais e funções de artefatos.
> (Ratcliffe, 2007, p. 21)

E mais ainda, "s*e combinamos a percepção de ação, gesto e expressão com* (…) *entendimento situacional* (…)*, nós temos uma compreensão sofisticada de ações sem* [que] *qualquer contribuição da Psicologia do Senso Comum* [seja] *necessária*" (Ratcliffe, 2007, p. 149), isto é, sem que seja preciso atribuir estados mentais.

Partindo da afirmação de que nossa cognição é, nos sentidos apontados, corporificada, estendida e socialmente situada, Ratcliffe propõe que as teorias filosóficas da Psicologia do Senso Comum (no sentido externalista) que explicam a cognição como resultando da leitura mental são incompatíveis com nossa prática interpretativa interpessoal. Ele questiona se nós realmente empregamos o arcabouço conceitual crença-desejo (o arcabouço conceitual da PSC) em atribuições de estados

mentais nas nossas interações cotidianas, e conclui com uma resposta amplamente negativa: nós não empregamos o arcabouço conceitual da PSC porque não atribuímos estados mentais.

Ao menos três passos fundamentam tal conclusão. O primeiro passo dessa negativa é a constatação de que nós raramente nos encontramos na posição de observadores isolados, mas estamos sempre inseridos em um contexto de interação social:

> Relações eu-você são geralmente caracterizadas por interação expressiva, gestual e dialógica entre os participantes. Esta interação (…) já é em si constitutiva de nossa habilidade de entender uns aos outros. O entendimento interpessoal geralmente não é uma questão de empregar habilidades internas em contextos de observação. Ao contrário, as habilidades em parte residem na própria interação.
> (Ratcliffe, 2007, p. 22)

A questão é que, por ser tão profundamente relacional, nosso entendimento das outras pessoas é fundamentalmente diferente da atitude desvinculada de um cientista para com seu objeto de estudo. Ele reitera:

> o entendimento interpessoal não é sempre fenomenologicamente descritível como A atribuindo estados internos a B ou mesmo A interpretando B. Tal entendimento é constituído pela interação entre eles. B não é apenas interpretado por A mas é também participante do processo pelo qual A interpreta A, B e o relacionamento entre eles.
> (Ratcliffe, 2006, p. 40)

Um exemplo bem claro que Ratcliffe dá para apoiar essa posição é a de um casal visitando uma casa com a intenção de possivelmente comprá-la:

> o julgamento de uma pessoa não é formulado independentemente e depois socialmente mediado, mas estruturado desde o começo pela interação com o companheiro. (…) O que acontece geralmente é que avaliações emergem em meio a olhares mútuos, expressões, “hums” e “ahs” e, é claro, comentários. A análise final geralmente não é a de A ou a de B ou uma mistura do que A pensou e B pensou. Ela emerge na interação entre eles.
> (Ratcliffe, 2006, p. 40)

O segundo passo de Ratcliffe na negação da ideia de que nós empregamos a estrutura conceitual crença-desejo em atribuições de estados mentais se dá na sugestão de que estados mentais são diretamente observados, e não apenas inferidos: "*o que uma pessoa vê não é um indicador externo de uma coisa interna escondida, mas a presença de uma experiência, emoção, intenção ou atitude*" (Ratcliffe, 2007, p. 145). Ele afirma que não devemos distinguir estados mentais internos de suas manifestações externas, e completa: "*O que quero sugerir é que o que percebemos é mesmo literalmente uma emoção, uma intenção ou mesmo um pensamento"* (Ratcliffe, 2007, p. 145). A consequência disso é que, uma vez que não somos nunca observadores isolados, mas estamos engajados em uma troca que efetivamente forma as crenças em questão (primeiro passo), a observação de um estado mental em outra pessoa é parte integral do processo formativo de tal estado em nós mesmos, e vice-versa:

> já que disposições para certos gestos e expressões, e às vezes os próprios gestos e expressões, podem ser diretamente obtidos pela percepção do comportamento de outra pessoa, segue-se que as pessoas influenciam diretamente a vida mental umas das outras por meio da interação.
> (Ratcliffe, 2007, p. 149)

O terceiro e conclusivo passo na negação da ideia de que nós empregamos a estrutura conceitual crença-desejo em atribuições de estados mentais é que o próprio arcabouço conceitual em questão é desnecessário: nos raros casos em que explicamos e prevemos comportamentos, a compreensão das regras sociais é suficiente para fundamentar o trabalho interpretativo que permite nossas interações. Ratcliffe reconhece que em certo sentido a capacidade de prever comportamentos é imprescindível para atividades corriqueiras como andar em uma rua movimentada; entretanto nesse tipo de caso nós não dirigimos nossos esforços interpretativos para pessoas específicas, mas apenas "navegamos" a rua evitando possíveis colisões. Ele diz:

> trabalhos recentes em Psicologia do Senso Comum assumem que o entendimento interpessoal é alcançado pelo emprego de habilidades cognitivas internas de modo a solucionar um problema posto pelo ambiente externo, que é [o problema de] prever o que os outros provavelmente vão fazer. Mas eu sugiro que a habilidade de entender e antecipar os outros (…) é parcialmente constituída pelo ambiente social e é geralmente uma questão de interação com tal ambiente mais do que de contemplação desvinculada. Em vez de tentar atribuir causas internas de comportamento de modo a prever comportamentos (…), nós usamos o ambiente compartilhado para regular nosso comportamento, assim nos tornando mais previsíveis uns aos outros. (Ratcliffe, 2007, p. 107)

Em outras palavras, ele rejeita a ideia de que nós vemos o comportamento dos outros como um quebra-cabeça de um ambiente externo, e propõe que, inseridos em um processo conjunto, nós nos fazemos mutuamente interpretáveis ao nos adequarmos ao contexto compartilhado, e assim não precisamos de fato "prever" comportamentos. Dessa maneira, por exemplo, o mero reconhecimento de que em determinado contexto uma pessoa exerce a função de atendente bancário, garçom, motorista de ônibus, vendedor de carros, etc., já nos fornece indicações de como podemos esperar que ela irá se comportar naquela situação. Em contextos interativos, uma vez que as normas do ambiente compartilhado e aspectos da situação já permitem a interpretação, os recursos de explicação e predição se tornam redundantes.

Em suma, para Ratcliffe, além de nunca observarmos os outros de maneira desvinculada (primeiro passo) e de geralmente percebermos muitos de seus estados "mentais" diretamente (segundo passo), a própria psicologia crença-desejo (o arcabouço conceitual mentalista da PSC) é desnecessária, sendo mais importante para nosso entendimento interpessoal uma compreensão da complexa rede de regras sociais compartilhadas e de aspectos contextuais. Mas o ponto central para Ratcliffe não é apenas que a leitura mental é tornada desnecessária e redundante pelas indicações contextuais providas pela interação; de maneira ainda mais radical, ele propõe que a

atribuição de estados mentais e de atitudes proposicionais não é parte de nosso "*discurso e práticas interpretativas cotidianas*", sendo apenas "*algo inventado pela filosofia e ciência cognitiva*" (Ratcliffe, 2007, p. 224). Ele explica que a expressão "Psicologia do Senso Comum" costuma designar uma destas quatro ideias: "*(i) O que as pessoas fazem. (ii) O que as pessoas pensam que fazem. (iii) O que filósofos e cientistas pensam que as pessoas fazem. (iv) O que filósofos e cientistas pensam que as pessoas pensam que elas fazem*" (Ratcliffe, 2007, pp. 55-56)[11]. O eliminativismo clássico de Churchland reconhecia a PSC como uma descrição verdadeira de uma prática social e de uma teoria falsa que as pessoas utilizam para fundamentar essa prática, ou seja, o ponto (i) acima; mas, seguindo a distinção internalista/externalista proposta por Stich, Ratcliffe defende que a Psicologia do Senso Comum como atribuição de estados mentais é apenas um construto de sistematizações filosóficas: "*A PSC é, eu sugiro, uma mistura ambígua de (iii) e (iv)*" (Ratcliffe, 2007, p. 56). Ele afirma que:

> O eliminativismo de Churchland recomenda a eliminação de algo que nem sequer faz parte de nossa vida social. É a eliminação de o que certos filósofos e cientistas cognitivos pensam que as pessoas pensam, e não o que as pessoas realmente pensam.
> (Ratcliffe, 2007, p. 224)

Em contraste com o eliminativismo de Churchland, Ratcliffe propõe que teorias que explicam a cognição social como baseada na leitura mental (as quais ele chama de "Psicologia do Senso Comum") são incompatíveis com nosso entendimento interpessoal ordinário porque nós não atribuímos estados mentais uns aos outros. Sua posição portanto resulta em um tipo de eliminativismo: da ideia de que a leitura mental é apenas um construto filosófico e não algo que as pessoas de fato realizam em suas interações diárias, ele conclui que ela deveria ser eliminada de nossas teorias da cognição social.

[11] Na distinção proposta por Stich, a "teoria da PSC" teria sentido internalista apenas no ponto (i).

4.2: Críticas

Uma leitura crítica da proposta eliminativista de Ratcliffe deve reconhecer seu mérito. A literatura filosófica a respeito da Psicologia do Senso Comum de fato tem priorizado uma imagem racionalista que pode não ser um retrato tão fiel da realidade de nossa prática interpretativa intersubjetiva. Parece irresistível a ideia de que nossa atividade cognitiva é profundamente marcada tanto por nossa estrutura biológica quanto pelo contexto social, cultural, linguístico, etc. em que estamos inseridos. É também bastante plausível a sugestão de que nossa interpretação do comportamento alheio não ocorre de maneira desvinculada, como um observador "interior" isolado que tenta decifrar seu ambiente "exterior". Mas, embora a razoabilidade dessas ideias pareça dar suporte às ideias de Ratcliffe, minha impressão é que a ausência de leitura mental torna a proposta de Ratcliffe no mínimo inadequada e potencialmente inconsistente e auto-refutante. Nesta seção apresentarei dois argumentos independentes para substanciar essa afirmação.

Primeiro (4.2.1), eu apresento dois "cenários de colapso" (*breakdown scenarios*) que desafiam o eliminativismo de Ratcliffe ao proporem que ao menos algumas instâncias de entendimento interpessoal dependem da atribuição de estados mentais e atitudes proposicionais. Segundo (4.2.2), eu examino o papel que a racionalidade e as normas sociais exercem no modelo de Ratcliffe e argumento que ao menos um tipo mínimo de leitura mental é requerido. Estes dois argumentos independentes revelam a insuficiência da abordagem anti-leitura mental de Ratcliffe, e sugerem que atribuições de estados mentais e atitudes proposicionais persistem como fenômenos a serem

explicados por qualquer teoria adequada sobre como as pessoas interagem e entendem umas às outras.

Esta conclusão não deve ser interpretada como nos comprometendo a abordagens tradicionais onde o entendimento interpessoal é obtido *apenas* por meio da leitura mental; antes, ela torna possíveis abordagens pluralistas nas quais o entendimento interpessoal é obtido por meio de uma combinação de leitura mental e interpretação contextual situada, às vezes envolvendo primariamente um, outras vezes o outro, ou ainda possivelmente ambos. Na terceira e última subseção (4.2.3) articulo uma crítica final a Ratcliffe, propondo que, dados seus compromissos teóricos, seria conceitualmente viável para ele adotar uma abordagem pluralista. Embora os detalhes desta proposta estejam além do escopo desta dissertação, na última seção deste capítulo (4.3) eu discuto alguns dos princípios que devem guiar uma tal abordagem pluralista.

4.2.1: Contraexemplos

Ratcliffe propõe um modelo de cognição social no qual interações são situadas em um contexto compartilhado constituído de uma rede complexa de normas e papéis sociais bem como as possibilidades de ação (*affordances*) de objetos e da estrutura física do ambiente. Em sua visão, o meio social e físico não é apenas pano de fundo para a ação, mas parte integrante de ações, de tal maneira que o entendimento interpessoal é obtido por meio da percepção dos componentes contextuais da ação:

> A percepção das atividades das pessoas, embutida em situações compartilhadas, é frequentemente tudo o que precisamos para facilitar o nível de entendimento requerido para interação e coordenação bem-sucedida. Mesmo quando explicamos ações em termos de razões, nós

> frequentemente o fazemos nos referindo a aspectos de situações e não a estados psicológicos.
> (Ratcliffe, 2007, p. 186)

E mais ainda, "*a percepção de ação, gesto e expressão em situações compartilhadas faz todo o trabalho* [interpretativo]" (Ratcliffe, 2007, p. 151). Sua posição pode portanto ser resumida nas afirmações de que, porque nós dependemos apenas de uma compreensão da situação compartilhada, (1) a atribuição de estados mentais e atitudes proposicionais é desnecessária, e, ainda mais, (2) a atribuição de estados mentais e atitudes proposicionais não é algo que nós fazemos.

"Cenários de colapso" ("*breakdown scenarios*") podem ser uteis ao tornarem explícito o papel funcional de um componente no todo, nos permitindo concluir, por exemplo, que se um carro para de funcionar quando uma determinada peça quebra, então aquela peça deve exercer um papel importante no funcionamento normal do carro. Consideremos então dois cenários de colapso da cognição social, ou seja, casos em que *não há entendimento interpessoal*, para ver o que eles revelam sobre o que é requerido para o entendimento interpessoal *bem-sucedido*.

Imagine que Carlos, um jovem paulistano, está visitando o Rio de Janeiro pela primeira vez. Ele passeia distraído, maravilhado com a cidade, e entra na estação do metrô com o objetivo de dirigir-se a um determinado ponto turístico. Ao avistar o trem chegar e as portas se abrirem, ele entra no vagão mais próximo, e é recebido com hostilidade pelos passageiros, que expressam reprovação com olhares e palavras grosseiras. Em uma questão de segundos Carlos percebe que todos os passageiros são mulheres, e vê uma placa dizendo que em horários específicos aquele vagão é destinado

apenas para passageiros do sexo feminino. Constrangido, ele sai do vagão o mais rápido possível. Naqueles poucos segundos, do ponto de vista de uma das passageiras que reagiu com hostilidade, aquele homem (Carlos) está infringindo uma regra que é de conhecimento público, então ele provavelmente o faz porque deseja usar um espaço que não lhe é de direito, ou pior, porque tem intenções espúrias para com as passageiras; do ponto de vista de Carlos, no entanto, que apenas estava distraído e nunca ouvira falar da existência de vagões exclusivos no metrô carioca, a reação negativa dos outros passageiros é a princípio difícil de entender (sendo possível que ele tenha se perguntado "Será que estou sujo? Será que pisei em alguma coisa que cheira mal?"), mas, depois, ao notar a placa informativa e a cor rosa do lado de fora do vagão tudo se esclarece. A interpretação mais óbvia desse exemplo (que pesaria em favor da primeira afirmação de Ratcliffe) é que ele mostra que uma compreensão das normas sociais que se aplicam a uma determinada situação compartilhada é mais informativa para o entendimento interpessoal do que a capacidade de "ler" a mente de outras pessoas. Mais sutilmente, no entanto, este caso exemplifica como nós *tentamos* atribuir estados mentais a outras pessoas (como a passageira do exemplo) mesmo quando fatos sobre a situação bastariam para explicar o comportamento de alguém (ou, no caso de Carlos, a falta de familiaridade com tais fatos sobre a situação); isso contradiz a segunda afirmação de Ratcliffe, de que nós não lemos a mente de outras pessoas, mesmo se a primeira afirmação segue sem ser questionada.

Considere outro cenário. Imagine a situação em que uma pessoa é subordinada a outra, como um supervisor no local de trabalho. Existem certos direitos e deveres explícitos envolvidos em cada um dos papeis neste relacionamento, assim como muitas expectativas sociais implícitas relativas a cada cargo. Agora, suponha que Sílvia (a

supervisora) encarrega Pedro (o empregado) de um determinado projeto, definindo o resultado esperado, mas não dando instruções específicas sobre como o projeto deverá ser desenvolvido. Imagine, ainda, que uma vez que Pedro tenha finalizado este projeto com sucesso, Sílvia critica duramente o trabalho de Pedro *por causa da maneira como ele o desenvolveu*, apesar de o resultado esperado ter sido atingido. Neste tipo de cenário, se mais nenhuma explicação for dada, Pedro provavelmente ficará confuso e pensará algo como "Eu não entendo a Sílvia!" ou "Eu simplesmente não sei o que ela quer que eu faça!". Vários fatos poderiam explicar porque Sílvia criticou pedro apesar de ele ter realizado o projeto que lhe foi incumbido da melhor maneira que ele sabia, e de tê-lo feito de maneira bem-sucedida, obtendo o resultado esperado. Poderia ser, por exemplo, que Sílvia reagiu de maneira exagerada e impensada apenas porque ela está tendo um dia difícil; mas também poderia ser que Sílvia tem uma preferência pessoal pelo procedimento A em vez de B (ou, por exemplo, pela fonte Arial em vez de Times New Roman), a respeito da qual Pedro não tem a menor ideia; ou ainda, é possível que, embora ninguém mais na empresa pense assim, Sílvia pense que a realização daquele tipo de projeto exige que um determinado protocolo seja seguido, ou ao menos que alguém no cargo de Pedro deveria segui-lo. O que este caso ilustra é que as normas e papeis sociais em uma situação compartilhada são apenas parte do que torna a interação efetiva: neste exemplo, o que atrapalha o entendimento interpessoal não é uma falta de compreensão das normas e valores no contexto compartilhado nem das regras que determinam os papeis sociais e suas responsabilidades, mas sim a falta de conhecimento que Pedro tem do estado emocional de Sílvia hoje, ou de sua preferência pessoal dentre vários procedimentos aceitáveis, ou mesmo de suas *crenças* sobre quais são as regras procedimentais nesse caso ou sobre o que o cargo de Pedro implica. Mas, se a falta de entendimento interpessoal ao menos às vezes resulta da falha em leitura mental, então

parece que a atribuição bem-sucedida de estados mentais e atitudes proposicionais deve exercer um papel positivo no entendimento interpessoal bem-sucedido; esse caso desafia diretamente a primeira afirmação de Ratcliffe (de que a leitura mental é desnecessária), mas também permite a crítica da segunda afirmação (de que nós não lemos a mente de outras pessoas), uma vez que qualquer pessoa na triste situação de Pedro muito provavelmente buscaria explicações para a crítica que recebeu, ao menos algumas das quais plausivelmente envolveriam os estados mentais de Sílvia.

Embora estes dois cenários de colapso possam parecer triviais, eles invocam intuições que coincidem com a evidência empírica coletada nos últimos cinquenta anos sobre como as pessoas tendem a atribuir agência e estados mentais intencionais até mesmo a seres que não poderiam tê-los, como objetos e formas geométricas animadas (cf. Nichols no prelo, e sua discussão dos estudos originais de Heider & Simmel nos anos 40 bem como da pesquisa subsequente). De maneira similar com o que ocorre nestes estudos, no exemplo do metrô carioca Carlos não possui nenhuma crença ocorrente[12] sobre a situação (como a crença de que *no Rio de Janeiro existem vagões exclusivos* e que *este é um vagão exclusivo*), de maneira que, portanto, não existem crenças ocorrentes relevantes para a passageira hostil "ler" ou atribuir a Carlos; ainda assim, seria pouco surpreendente se essa passageira hipotética de fato tentasse resolver sua confusão atribuindo estados mentais a Carlos, como proposto no exemplo e como a evidência empírica sugere que as pessoas frequentemente fazem. Isto contradiz o que listei como a segunda afirmação de Ratcliffe, que nós não atribuímos estados mentais e atitudes proposicionais aos outros quando tentamos entendê-los.

[12] Uma crença ocorrente é uma crença que um indivíduo tem e que ele contempla conscientemente em um dado momento; em outros momentos, quando ele pensando sobre alguma outra coisa, a mesma crença é apenas disposicional: "*uma crença ocorrente vem e vai, dependendo de se as circunstâncias a elicitam ou não; a crença disposicional permanece*" (Schwitzgebel, 2010).

A motivação de Ratcliffe para rejeitar a leitura mental é ao menos parcialmente formulada como uma rejeição de abordagens espectatoriais ao entendimento interpessoal, as quais, ele argumenta, "*artificialmente removem* [o entendimento] *dos contextos práticos em que nós encontramos outras pessoas*" (Ratcliffe, 2007, p. 70). Muitos exemplos de entendimento interpessoal encontrados na literatura da Psicologia do Senso Comum tendem a ter a forma de pessoa A observando o comportamento da pessoa B e explicando que *B fez X porque ela acredita que Y*, ou prevendo que *B fará P porque ela deseja que Q e acredita que P implica Q*. Ratcliffe compara essa imagem clássica com a de um cientista observando um objeto, como uma pedra, e afirma que "*entender pessoas nunca é distanciado* [deste jeito]", mas "*envolve adotar uma perspectiva pessoal*" (Ratcliffe, 2007, p. 23). Ele propõe que essa perspectiva pessoal requer a apreciação dos aspectos contextuais nos quais a ação de uma determinada pessoa se situa, incluindo as normas sociais em jogo, bem como a estrutura do ambiente físico; entretanto, ele acrescentaria, essa apreciação não é realizada pelo ponto de vista de um espectador, e sim pela responsividade mútua e pela proximidade que só são possíveis na interação como coparticipantes situados em um mesmo contexto. Assim, o pensamento de Ratcliffe é que a atribuição de estados mentais e de atitudes proposicionais caracterizaria uma perspectiva observacional distanciada (ou "fria"), contrária à perspectiva pessoal que ele defende.

Mas não parece óbvio que a leitura mental é necessariamente distanciada e espectatorial, como quer Ratcliffe, e de fato não há nenhuma contradição lógica na ideia de uma leitura mental "envolvida" ou "situada". Além disso, o caso do local de trabalho acima sugere que uma boa compreensão das normas sociais que se aplicam a todos os

supervisores e empregados em uma situação similar pode não ser suficiente para evitar a falta de entendimento interpessoal; mesmo se o contexto compartilhado provê aproximações heurísticas para a coordenação simples, tomar a perspectiva pessoal com frequência requer que se vá além de generalizações sobre como uma pessoa pensa baseado em sua posição social, etnia, religião, etc., e, ao menos em alguns casos, envolverá a busca de uma compreensão melhor do que a pessoa em questão espera, deseja, prefere, ou mesmo o que ela acredita ser o papel de outra pessoa, portanto exigindo a atribuição de estados mentais em geral ou até mesmo de atitudes proposicionais propriamente ditas. Dessa maneira, os cenários de colapso acima apresentam desafios às duas afirmações de Ratcliffe listadas no início desta seção, já que eles indicam que a atribuição de estados mentais é ao menos às vezes necessária para que entendamos outros *como pessoas*, e que ao menos às vezes as pessoas de fato atribuem estados mentais umas às outras mesmo quando isto não é estritamente necessário. A leitura mental pode ou não ser tão universal e onipresente quanto os defensores das abordagens vistas em capítulos anteriores sugerem, mas ela permanece como um fenômeno que nossas teorias da cognição social precisam explicar, e cuja ausência torna inadequado o modelo que Ratcliffe oferece do entendimento interpessoal.

4.2.2: Racionalidade

A abordagem eliminativista de Ratcliffe à leitura mental baseia-se em uma perspectiva externalista da cognição social. Sua ideia é que, em vez de depender de nossas capacidades cognitivas internas para atribuir estados mentais internos para outras pessoas, o entendimento interpessoal é intrinsecamente situado; desta forma, nós somos

capazes de interagir e coordenar atividades com sucesso precisamente porque nós encontramos as outras pessoas em contextos onde redes complexas de normas e papeis sociais implícitos e explícitos, e *affordances* físicas regulam as ações de todos. Como interações ocorrem nestes contextos compartilhados, Ratcliffe explica, nós apenas assumimos que os outros "*farão 'o que as pessoas devem fazer' naquele tipo de ambiente*", sem qualquer necessidade de individualmente atribuir estados mentais a eles: "*Simultaneamente atribuir crenças e desejos a várias centenas de pessoas requereria um esforço cognitivo considerável, sem muito ou qualquer benefício*" (Ratcliffe, 2007, p. 91). Ele reconhece que as pessoas podem até quebrar tais regras explícitas e tácitas, algumas das quais dizem respeito a questões morais enquanto outras são apenas procedimentais; contudo, ele propõe que tais regras são componentes objetivos da situação, e quer as pessoas obedeçam ou quebrem as regras, nós interpretamos suas ações em relação a essas rotinas exteriores e normas sociais e ambientais, e não em relação aos estados mentais que pensamos que elas tem. Nesta seção eu examino o papel que Ratcliffe designa para normas sociais e a racionalidade em seu modelo de cognição social, e concluo que a ausência de leitura mental torna seu modelo inconsistente.

Como Donald Davidson notoriamente defendia, "*A interpretação bem-sucedida necessariamente investe de racionalidade básica a pessoa interpretada*" (Davidson, 2001, p. 211). Ratcliffe concorda com Davidson (cf. Ratcliffe, 2007, p. 96), mas, dada sua perspectiva externalista da cognição social, Ratcliffe articula uma visão externalista também da racionalidade:

> As normas que figuram em explicações do comportamento envolvem, é claro, normas de raciocínio, além de normas situacionais. Entretanto, (…) normas de raciocínio não são geralmente entendidas como características de processos mentais internos, assim como normas situacionais não o são.

(Ratcliffe, 2007, p. 195)

E ele acrescenta: "*As normas de raciocínio estão aí para todos seguirem, assim como as normas situacionais*" e "*Os padrões em questão são padrões compartilhados, em vez de serem processos psicológicos que são atribuídos a indivíduos*" (Ratcliffe, 2007, p. 217). Em outras palavras, assim como nós encontramos outras pessoas como se conformando ou não às normas sociais externas, nós também entendemos seu comportamento e raciocínio à luz de regras igualmente *externas* da racionalidade; então mesmo se Davidson está certo ao dizer que a interpretação naturalmente envolve ver outras pessoas como racionais, na visão de Ratcliffe isso significa apenas que nós vemos os outros como situados em uma situação que é constituída por normas que ditam não apenas o que é moral, socialmente aceito, e procedimentalmente adequado, mas também o que é racional.

Embora a princípio esta possa parecer uma boa maneira de resguardar a racionalidade sem sucumbir a um modelo de cognição social baseado em leitura mental, uma análise mais cuidadosa mostra que esta solução resulta de uma confusão que, quando desfeita, torna o modelo de Ratcliffe inconsistente exatamente por *requerer* a leitura mental. O filósofo Robert Audi explica que a racionalidade global de pessoas depende de sua capacidade geral de responder à razão, o que implica "*sua habilidade compreender certas verdades óbvias e de fazer certos tipos de inferências justificadas a partir delas, assim como a habilidade de raciocinar de maneira prática*" (Audi, 2001, p. 50). Nós podemos, portanto, discriminar, por um lado, o quê em uma inferência a torna 'justificada', por exemplo, e, por outro lado, saber fazer uma tal inferência justificada; enquanto ambos tratam da racionalidade, o primeiro diz respeito ao conteúdo objetivo da regra enquanto o segundo se refere à capacidade que uma pessoa

tem de seguir tal regra ou não. Ratcliffe parece ignorar esta distinção básica entre o conteúdo das regras de raciocínio e a capacidade de raciocinar. Ele afirma:

> Entender porque Henrietta raciocinaria de uma determinada maneira geralmente não requer mais do que entender porque nós mesmos raciocinaríamos de tal maneira. Nos dois casos, é apenas uma questão de apreciar o que se deve pensar ou fazer quando diante de uma determinada situação.
> (Ratcliffe, 2007, p. 217)

Mas enquanto o conteúdo das normas de raciocínio podem ser objetivos, a capacidade de responder a tais normas certamente não é universal. Esse ponto fica claro quando adaptamos o exemplo e imaginamos Henrietta como portadora de alguma deficiência intelectual grave: embora as normas da racionalidade estejam *aí para todos seguirem*, em tais circunstâncias, e dependendo da situação, não seria justificável esperar que Henrietta obedeça tais normas. O mesmo vale se, em vez de portar uma deficiência intelectual, Henrietta é uma pessoa insensata que rotineiramente age por impulsos e toma decisões sem pensar, contra seu melhor julgamento e sem considerar seus próprios interesses ou o de outras pessoas; notar que ela tem um nível limitado de responsividade à razão nos impediria de esperar que ela aja racionalmente, ou seja, que ela obedeça as regras que *estão aí para todos seguirem*.

Ainda tratando das regras da racionalidade, Ratcliffe afirma que "*tais normas não são usadas apenas para prever o comportamento de outras pessoas mas também para criticar e corrigir seu raciocínio ou para ensiná-las a raciocinar de uma maneira particular*" (Ratcliffe, 2007, p. 217). Entretanto, dependendo de quem for a pessoa em questão, poderia ser simplesmente irracional criticar, corrigir ou tentar ensiná-la a raciocinar de qualquer maneira, precisamente porque ela pode ser completamente incapaz ou ser muito limitada em sua capacidade de raciocinar. Aqui, portanto, se torna instrumental saber distinguir entre pessoas que não poderiam ter agido de maneira

diferente e aqueles cujo comportamento "não tem desculpas". Se nós interpretamos uma pessoa como inocente, nós podemos tentar instruí-la, o que pode ser bem-sucedido ou não dependendo tanto da complexidade das regras em questão quanto da capacidade que aquela pessoa tem para compreendê-las; nos piores casos podemos concluir que é perda de tempo e que a pessoa simplesmente não irá entender. Por outro lado, a suspeita de que uma pessoa agiu de má-fé pode afetar drasticamente nossa atitude em relação àquela pessoa, ativando o julgamento de que ela é, e.g., preguiçosa (no caso das normas de raciocínio) ou que ela é "esperta" ou maliciosa (no caso de ela ter quebrado alguma norma social visando ganho pessoal), e assim por diante. Aqui um mecanismo cognitivo mais complexo pode envolver o reconhecimento do outro como ciente da regra em questão e talvez simultaneamente ciente de que pode enganar os outros levando-os a pensar que é inocente – desta maneira envolvendo não apenas a atribuição de certas crenças e intenções, como também a atribuição da capacidade de leitura mental. Em casos mais simples, porém, a despeito de se julgamos que uma pessoa é inocente ou maliciosa, um tipo mínimo de leitura mental parece estar presente na percepção de alguém como racional, isto é, como capaz de processar informações de uma maneira tal que envolve uma compreensão satisfatória das regras (objetivas) da racionalidade.

Confiante de ter preservado a racionalidade em um modelo de entendimento interpessoal destituído de leitura mental, Ratcliffe conclui:

> As pessoas apreciam "o que se deve fazer" ou "o que se deve pensar" em várias situações e pressupõem, na maior parte, que os outros participam das mesmas normas. (…) Uma apreciação do bom raciocínio tem a mesma estrutura que o entendimento situacional. (Ratcliffe, 2007, p. 218)

Ratcliffe parece estar certo nesta conclusão, mas suspeito que seja de uma maneira que refuta seu próprio modelo: a similaridade entre as normas sociais e as normas do bom raciocínio parece ser que, embora seu conteúdo seja externo ou objetivo (e *esteja aí*

para todos seguirem), a capacidade de responder adequadamente a tais normas é algo que varia entre indivíduos diferentes; aquelas pessoas que nós percebemos como menos capazes de responder às normas (seja normas sociais ou normas da racionalidade) nós também tratamos de maneira diferente, sem esperar que elas ajam de maneira adequada em algumas situações; ao mesmo tempo, uma pessoa que nós julgamos que "deveria saber o que faz", nós consideramos responsável de maneiras das quais a pessoa inocente pode ser isenta. Além disso, seguindo a distinção de Audi, nós vemos que além de variar entre indivíduos diferentes, a capacidade de responder à razão até certo ponto é independente de sua realização ou aplicação – ou seja, uma pessoa pode ocasionalmente não agir ou pensar racionalmente embora seja geralmente capaz de compreender as regras da racionalidade; isso se torna evidente se alguém que conhecemos passa por algum evento traumático, e nós então ajustamos nossas expectativas sobre seu comportamento, assumindo que "ele não está em suas plenas faculdades mentais". Variando entre indivíduos distintos assim como no decorrer do tempo para um mesmo indivíduo, a responsividade à razão – parece razoável concluir – permanece uma capacidade interna de processar informações de uma maneira ou de outra, de acordo com o conteúdo das normas em cada caso; assim, a inferência da racionalidade por sua vez consiste na atribuição de uma capacidade interna de compreender e responder a razões (ainda que estas sejam externas). Nesta perspectiva, embora a inferência da racionalidade possa ou não envolver a atribuição de conjuntos particulares de atitudes proposicionais, é razoável interpretá-la como um tipo mínimo de leitura mental. Se esta interpretação está correta, então o modelo de Ratcliffe é inconsistente, e ele deve optar entre adotar a leitura mental para manter a racionalidade em seu modelo de cognição social, ou então seu modelo pode manter-se destituído de leitura mental mas com o alto preço de desistir da racionalidade.

4.2.3: A Linguagem como Extensão Cognitiva

As críticas apresentadas nas duas subseções anteriores (4.2.1 e 4.2.2) tinham uma perspectiva negativa, visando criticar a proposta de Ratcliffe ao argumentar que sua rejeição da leitura mental torna seu modelo de cognição social inadequado e inconsistente. Nesta última subseção crítica meu argumento assumirá uma perspectiva positiva, sugerindo que, dados os compromissos teóricos de Ratcliffe, seria conceitualmente viável para seu modelo acomodar a leitura mental conjuntamente à interpretação contextual. O ponto central para este argumento é a adoção da tese da cognição estendida por parte de Ratcliffe.

Vale relembrar que para Ratcliffe, a situação social e o ambiente físico são mais do que mero "background" para a ação ou simplesmente auxílios cognitivos, mas antes: "[nossas] *habilidades cognitivas de maneira geral são indissociáveis de nossas interações com ambientes socialmente e tecnologicamente estruturados*" (Ratcliffe, 2007, p. 109). Tomando como ponto de partida de sua proposta externalista o modelo de cognição estendida articulado por Andy Clark, Ratcliffe concorda com Clark que, por exemplo, os recursos que estão disponíveis a um autor enquanto ele escreve um texto acadêmico pode efetivamente transformar o processo e seu resultado, fazendo emergir ideias que o escritor não poderia ter tido em outro contexto:

> Durante o processo de escrever e rescrever, ideias surgiram a mim que eu não havia considerado antes, como se o próprio ato de escrever fosse parcialmente responsável pela sua gênese. Além disso, durante a leitura do rascunho impresso de um capítulo, eu me vi capaz de enxergar ambiguidades e defeitos de argumentação dos quais eu não estava ciente antes de começar a escrever ou durante o processo de redação.
> (Ratcliffe, 2007, p. 110)

Em parte, isso decorre das possibilidades proporcionadas pela linguagem escrita: ao

escrever uma ideia, nós não temos mais apenas uma compreensão abstrata daquele pensamento, mas sua representação física nos permite examinar tal pensamento e literalmente *ver* seus componentes, sutilezas de argumentação e raciocínio, etc; anotações contendo rascunhos de ideias, citações e referências a outros trabalhos, e mesmo representações modais da conexão entre ideias diferentes (e.g., um diagrama de flechas), todos estes são recursos que podem contribuir e efetivamente compor o processo de redigir e revisar um texto, e podem ser ampliados em escopo com o uso de ferramentas virtuais (em computadores individuais, os recursos de processamentos de texto, e na internet, o acesso a diferentes fontes de pesquisa) levando assim às inovações conceituais que Ratcliffe descreveu.

Mas esta é apenas parte do quadro. Enquanto alguns desses aspectos da redação de um texto acadêmico se devem à ampliação da capacidade de processamento de ideias por meio de sua representação física (em papel ou na tela do computador), de maneira mais fundamental essas vantagens são extrapolações da extensão cognitiva proporcionada pela linguagem em si (e não apenas a linguagem *escrita*). Clark propõe que quer seja falada ou escrita, a linguagem "*torna disponíveis conceitos, estratégias e trajetórias de aprendizagem que simplesmente não estão disponíveis a cérebros individuais não-ampliados* (…) [e] *reconfigura os espaços computacionais que confrontam agentes inteligentes*" (Clark, 1998). Ele propõe que a linguagem aumenta nossa capacidade computacional ao proporcionar *expansão de memória*, *simplificação ambiental*, *coordenação e redução de deliberação "on-line"*, *alocação de atenção e recursos [cognitivos]*, e *manipulação e representação de dados*. De todas estas maneiras, Clark sugere que a linguagem nos fornece uma *dinâmica de segunda ordem*: "*assim que formulamos um pensamento em palavras (ou no papel), ele se torna um*

objeto para nós mesmos e para os outros. Como objeto, ele se torna o tipo de coisa sobre a qual podemos pensar" (Clark, 1998). Assim, a linguagem nos permite "*selecionar diferentes elementos de pensamentos complexos e examiná-los um de cada vez*", bem como "*estabilizar ideias muito abstratas na memória de trabalho* [ou de curto prazo]" e "*inspecionar e criticar nosso próprio raciocínio de maneiras que nenhuma outra modalidade representacional permite*" (Clark, 1998). Nesta perspectiva, portanto, o desenvolvimento linguístico abrange não apenas nossa percepção do mundo, mas inclui a possibilidade de meta-cognitivamente avaliarmos nosso próprio pensamento, percepção e capacidades de processamento cognitivo, assim como, um passo mais adiante, ponderar nossas próprias capacidades meta-cognitivas.

O ponto desta discussão sobre as extensões cognitivas proporcionadas pela linguagem é simples. Se a linguagem nos permite cristalizar pensamentos abstratos em um formato que pode ser avaliado por nós mesmos e por outros, tal ampliação cognitiva naturalmente resulta na possibilidade de assumirmos perspectivas e formarmos julgamentos a respeito de descrições de situações e, além disso, examinar (e formar julgamentos) sobre nossa própria perspectiva a respeito de tais descrições tanto quanto sobre as perspectivas de outras pessoas; com isso obtemos uma extensão de nossa compreensão não apenas do mundo e de nós mesmos, mas também de outras pessoas e de como elas pensam e agem – ou, em outras palavras, além de um aumento de nossas capacidades cognitivas e meta-cognitivas, um ganho também na cognição social. Ao menos em contextos dialógicos, a mediação linguística permite que interlocutores explicitamente compartilhem perspectivas e também que avaliem a perspectiva um do outro a respeito da situação e dos papeis e normas em jogo no ambiente compartilhado; plausivelmente, portanto, a extensão cognitiva proporcionada pela linguagem em

conjunção com a estrutura cultural fundamentada nessas capacidades estendidas e em repetidas interações linguisticamente mediadas, contribuiriam até mesmo a interações não verbais, levando-nos a formular pensamentos sobre o que os outros pensam ou querem (i.e., suas crenças e desejos), e levando-nos até mesmo a esperar que eles sejam capazes de ter pensamentos semelhantes sobre outras pessoas. Embora Ratcliffe concorde com a ideia geral da cognição estendida (até mesmo por meio da linguagem), ele limita os efeitos da facilitação de processamento à dependência das convenções sociais em jogo em situações compartilhadas. Mas essa parece ser uma limitação arbitrária. Se, como Ratcliffe defende, a cognição humana efetivamente explora recursos externos de tal forma que podemos razoavelmente descrever nossos processos cognitivos como estendendo além da mente/cérebro, então não há razão porque esta ampliação não incluiria, como sugerido acima, a extensão de nossas capacidades cognitivas para incluir a aptidão para atribuir estados mentais e atitudes proposicionais a outras pessoas.

Com esta discussão sobre como a linguagem proporciona uma extensão cognitiva, não pretendo comprometer-me com a tese da cognição estendida e nem com o programa de pesquisa da cognição corporificada; da mesma maneira, esta discussão não pressupõe que atitudes proposicionais de fato requeiram linguagem e um aparato cognitivo computacional. Antes, meu objetivo nesta subseção era apenas sugerir que o compromisso de Ratcliffe com teorias cognitivas externalistas não deveria impedi-lo de adotar uma abordagem pluralista que inclui tanto atribuições de estados mentais quanto práticas e estratégias interpretativas não-mentalistas – mesmo se para ele isso signifique assumir que a leitura mental é uma extensão cognitiva linguística-cultural.

4.3: Princípios de uma Abordagem Pluralista da Cognição Social

Os argumentos independentes apresentados na seção crítica (4.2) suportam minha conclusão de que Ratcliffe não está justificado em rejeitar a leitura mental: a ausência da leitura mental torna seu modelo de entendimento interpessoal inadequado (4.2.1) e inconsistente (4.2.2), ao passo que seria conceitualmente viável para ele aceitar a leitura mental como um componente de nosso entendimento interpessoal intuitivo (4.2.3). De maneira geral, isso fundamenta a conclusão mais ampla de que a leitura mental continua sendo um fenômeno que precisa ser explicado por nossas teorias da cognição social. Ao afirmar isto, não pretendo sugerir que a proposta de Ratcliffe deve ser rejeitada por completo em favor de um retorno a teorias que explicam a cognição social como resultando *apenas* da atribuição de estados mentais e atitudes proposicionais. Antes, os dois extremos tem suas desvantagens, de maneira que parece mais promissor aliar seus aspectos positivos em uma abordagem *pluralista*.

Esta perspectiva pluralista reconhece que a cognição social é uma atividade complexa e multifacetada, que pode empregar uma série de habilidades, desde aquelas que nos permitem entender situações e perceber ações em seu contexto social e físico, àquelas que nos permitem entender os outros como agentes dotados de intencionalidade (o que, por sua vez, pode envolver a atribuição de estados mentais por meio de processos simulatórios e/ou ricos de informação). O conflito entre abordagens tradicionais da Psicologia do Senso Comum e um modelo destituído de leitura mental como o de Ratcliffe é em grande parte devido a uma discordância mais fundamental sobre a natureza da cognição e seu funcionamento. Ainda assim, o ponto de partida da

abordagem pluralista que aqui se propõe é o simples reconhecimento da diversidade de fenômenos e práticas que subjazem nossa cognição social (o *explanandum*) – o que é neutro no que diz respeito a *como* estes mecanismos cognitivos serão explicados teoricamente (o *explanans*).

Embora não seja possível aqui articular em profundidade esta abordagem pluralista, encerro este capítulo e a dissertação apontando três dos princípios básicos relacionados, inspirados na obra da filósofa Kristin Andrews, que fundamentam uma possível abordagem pluralista. Estes princípios não pretendem encerrar o debate filosófico acerca da Psicologia do Senso Comum, mas antes indicar direções que parecem promissoras para pesquisas futuras.

O primeiro é o princípio da **Subjetividade da Cognição Social**, de acordo com o qual o entendimento interpessoal é uma atividade subjetiva, individual. Isso significa que, apesar de nosso entendimento interpessoal frequentemente emergir de contextos interativos (e por isso é interpessoal no sentido de ser *entre pessoas*), a cognição social é fundamentada em um esforço *subjetivo* de entender o outro. Segundo Andrews, a iniciativa pessoal de tentar compreender por que uma pessoa agiu de determinada maneira provém de um processo multifásico: em um primeiro estágio, "*um estado afetivo de curiosidade, perplexidade ou confusão*" a respeito de uma pessoa e seu comportamento nos leva a querer buscar uma explicação (Andrews, 2012, p. 121); na segunda etapa, engajamo-nos em comportamentos que visam gerar uma explicação, "*podendo incluir a manipulação do mundo físico (como comportamento exploratório), comportamento verbal (como pensar em voz alta), ou contemplação*", e podendo até mesmo incluir fazer perguntas relevantes à pessoa observada ou a outros observadores

(Andrews, 2012, p. 123); a terceira fase ocorre "*quando uma explicação é gerada tal que resolve a tensão que iniciou o processo em primeiro lugar*" (Andrews, 2012, p. 121); essa terceira fase Andrews caracteriza como "*ter a experiência de 'aha!'*", isto é, ter a experiência de ter encontrado uma resposta que se estava buscando, reduzindo o desconforto ou tensão cognitiva inicial (Andrews, 2012, p. 145).

Em contraste com essa estrutura multifásica de explicações do comportamento, Andrews afirma que, no dia a dia, a maioria de nossas previsões comportamentais ocorrem "*automaticamente, no nível subpessoal*", tomando a forma de "*expectativas de background*"[13] que nós só notamos conscientemente quando elas falham (Andrews, 2012, p. 49). Essa diferença ela afirma decorrer da *assimetria* entre explicações e previsões. No debate Churchland-Fodor (capítulo 2) assumia-se que explicações e previsões do comportamento são simétricas, dependendo do sentido da aplicação das generalizações nômicas da PSC – isto é, se a partir de estados mentais como condições iniciais em direção à conclusão de um comportamento como efeito (previsão) ou se a partir do comportamento observado como efeito em direção à conclusão dos estados mentais como condições iniciais (explicação). Ao contrário desta pressuposição, Andrews argumenta que previsões e explicações são assimétricas porque, por um lado, "*podemos prever comportamentos para os quais não temos uma explicação*" enquanto, por outro lado, "*podemos explicar eventos que não podemos prever*" (Andrews, 2012, p. 41-42). Ou seja, no primeiro caso, às vezes podemos prever corretamente como certa pessoa irá agir embora não entendamos a razão de tal comportamento e não saibamos explicá-lo; já no segundo caso, mesmo se um certo comportamento nos pega de surpresa, isso não nos impede de conseguir explicá-lo – como ela afirma: "*em nossas*

[13] Aqui "expectativas de background" poderia ser traduzido como "expectativas de segundo plano", significando expectativas que são mantidas apenas subconscientemente.

relações sociais cotidianas, nós frequentemente nos ocupamos em explicar comportamentos inesperados, comportamentos que nós não poderíamos ter previsto" (Andrews, 2012, p. 42).

Os outros dois princípios decorrem deste primeiro. O segundo princípio é o da **Diversidade de Conteúdo**. Conforme sugerido nos exemplos examinados anteriormente, nosso entendimento interpessoal intuitivo depende tanto do conhecimento de perspectivas e preferências de outras pessoas quanto de uma familiaridade com o ambiente e as regras que definem atividades naquela situação. Assim, em contextos diferentes é possível que nossas explicações e previsões de comportamentos alheios dependam mais diretamente de fatos sobre os estados mentais dos outros ou então de fatos sobre o contexto em que a interação está situada. Mais do que isso, dada a subjetividade da cognição social, indivíduos diferentes poderão ter opiniões diferentes sobre o que é "suficientemente explicativo" para um dado comportamento, assim como poderão ter expectativas diferentes assim construindo previsões divergentes do que os outros farão. Como o processo de explicação é subjetivo, não é possível qualificar uma determinada explicação como "certa" ou "errada": em nossa prática ordinária, de acordo com essa descrição multifásica, existem apenas explicações convincentes ou não. E em casos diferentes pessoas diferentes considerarão tipos diferentes de informação como satisfatoriamente explicativas, assim saciando a curiosidade ou confusão que motivou a busca de explicações em primeiro lugar. Desse modo, em alguns casos a atribuição de atitudes proposicionais é suficientemente explicativa: se em uma festa notamos que um amigo repetidamente olha para seu celular, podemos satisfazer nossa curiosidade ou confusão explicando (para nós mesmos) que ele está esperando uma ligação (ou seja, ele *acredita* que determinada

pessoa irá telefoná-lo). Em outros casos, porém, dizer que a pessoa tem uma determinada crença ou desejo não resolve o estado de tensão, enquanto a atribuição de outros tipos de estados psicológicos podem ser mais satisfatórios: imagine que vemos uma amiga chorando ao ler a notícia de um acidente automobilístico que resultou na morte dos envolvidos; seria fácil explicar que ela está chorando porque *acredita* que um conhecido morreu. Entretanto, imagine que neste caso descobrimos que ela não conhecia nenhuma das vítimas; é provável que nós permaneceríamos confusos, e se alguém dissesse que "ela está chorando porque *acredita* que uma pessoa morreu", nós poderíamos indagar "sim… mas ela não conhecia essa pessoa!"; se, porém, alguém dissesse "ela é sensível mesmo" ou "ela anda emotiva ultimamente", tais informações sobre sua disposição, estado de ânimo, ou traço de personalidade seriam mais explicativas do que meras informações sobre suas atitudes proposicionais.

A diversidade de conteúdo reconhece que explicações diferentes podem ser interpretadas de maneiras distintas também. No caso de comportamentos alheios, fatos sobre o contexto ou a situação às vezes podem ser mais informativos do que fatos sobre as atitudes proposicionais ou estados psicológicos de uma pessoa: ao perceber uma aglomeração na rua, podemos explicar tal comportamento coletivo "porque aconteceu um acidente" ou "porque o metrô está em greve" etc.; ou ainda, em se tratando de um comportamento individual, podemos explicar que um amigo chegou atrasado à festa porque ele trabalha do outro lado da cidade e hoje choveu bastante, causando um engarrafamento recorde. Já no caso de explicarmos um comportamento próprio, Andrews sugere: "*As pessoas geralmente explicam seus próprios sucessos em termos de características pessoais como traços de caráter, enquanto costumam explicar suas falhas em termos de características situacionais*" (Andrews, 2012, p. 156); assim, por

exemplo, se uma pessoa esqueceu de fazer um telefonema, é mais provável que ela explique seu comportamento dizendo que sua agenda está cheia e que ela está muito ocupada, e não reconhecendo que ela é desatenta e irresponsável.

O terceiro princípio, da **Diversidade de Estratégias e Mecanismos**, sugere que, em situações distintas, pessoas diferentes podem empregar práticas variadas para interagir com outras pessoas, coordenar atividades, e explicar e prever o comportamento dos outros. A partir da ideia de que explicações e previsões são assimétricas, este terceiro princípio sugere que uma diversidade de mecanismos cognitivos pode ser responsável por nossa formulação de explicações e previsões de comportamentos; mais ainda, é possível que mecanismos cognitivos diferentes sejam ativados na explicação e previsão do comportamento de estranhos ou de conhecidos. Ao mesmo tempo, este terceiro princípio sugere que, em casos em que não há leitura mental, a interação com ambientes fisicamente estruturados e a aquisição de conhecimentos sobre a organização social ou cultural do contexto pode ser suficiente para facilitar o entendimento interpessoal e a coordenação de atividades; já nos casos em que há leitura mental, tanto processos simulatórios quanto processos ricos de informação sobre outras pessoas podem contribuir para uma compreensão dos estados mentais dos outros. Dado o caráter subjetivo da cognição social, assim como fatos diferentes podem ser considerados como "suficientemente explicativos" por indivíduos diferentes, da mesma maneira é possível que uma mesma pessoa tenha critérios divergentes no decorrer do tempo como resultado da aprendizagem a partir de experiências intersubjetivas ou mesmo da aquisição de estruturas representacionais no desenvolvimento ontogenético.

As tarefas de crença falsa são particularmente relevantes para esse ponto. Como

discutido anteriormente (capítulo 3, seção 3.1), a inaptidão de crianças menores de quatro anos nas tarefas de crenças falsa foi inicialmente interpretada como indicando que antes dos quatro anos de idade crianças não possuem uma compreensão representacional de crenças e por isso são incapazes de entender que uma outra pessoa possui uma crença equivocada; no contexto da Teoria da Teoria, os resultados pareciam sugerir que só por volta dos quatro anos de idade é que crianças adquirem a "teoria" da Psicologia do Senso Comum, ou que só naquele ponto é que elas dominam a teoria suficientemente bem para de fato utilizá-la para atribuir estados mentais tais como crenças falsas. Estudos posteriores com versões diferentes da tarefa de crença falsa têm questionado essa conclusão. Uma variação interessante é a que utiliza o paradigma da violação de expectativa, segundo o qual, quando uma pessoa espera que algo aconteça mas sua expectativa é quebrada, ela tende a realizar comportamentos exploratórios com o objetivo de compreender a situação inusitada; comportamentos exploratórios podem envolver fazer perguntas, por exemplo, ou no caso de crianças em fase pré-linguística, a fixação da atenção naquela situação e a observação continuada por mais tempo. Um estudo com bebês de 15 meses de idade seguiu esse paradigma (Onishi e Baillargeon, 2005), e como resultado os bebês "*olharam durante mais tempo para uma cena em que o protagonista procurava um objeto em um lugar que ele não poderia saber que o objeto se encontrava (embora* [o objeto] *realmente estivesse lá)*", enquanto os bebês observavam durante um período menor de tempo "*uma cena em que o protagonista procurava um objeto onde ele tinha o escondido (embora* [o objeto] *não estivesse mais lá)*" (Buttelmann, Carpenter & Tomasello, 2009, p. 337). A diferença dos resultados parece significativa: bebês observavam por cerca de 10 segundos (em média) quando o protagonista agia de acordo com uma crença verdadeira (isto é, quando o protagonista tinha observado a mudança na localização do objeto), mas tinham a atenção fixa durante

até 27 segundos (em média 18 segundos) quando o protagonista deveria ter uma crença falsa (ou seja, quando o protagonista não tinha observado a mudança na localização do objeto, mas agiu como se soubesse onde ele estava, embora não fosse possível que ele de fato o soubesse); neste último caso, os autores do estudo concluem, "*os bebês esperavam que o protagonista procurasse pelo objeto onde ele acreditava que o objeto estava e observavam atentamente durante mais tempo quando ele não o fazia*" (Onishi e Baillargeon, 2005, p. 257), o que eles consideram indicar que bebês de apenas 15 meses de idade já possuem uma compreensão representacional da mente "*ao menos em forma rudimentar e implícita*" (Onishi e Baillargeon, 2005, p. 257). Estudos como este podem ser interpretados como corroborando o princípio da diversidade de estratégias e mecanismos ao sugerir que a cognição social recorre a capacidades distintas em diferentes fases da vida, e que mesmo antes do desenvolvimento linguístico pleno, seres humanos possuem formas simplificadas de leitura mental, as quais são potencialmente ampliadas e estendidas pela aquisição de conceitos próprios para tipos distintos de estados mentais e atitudes proposicionais (resultando até mesmo na capacidade de passar nas tarefas de crença falsa tradicionais, com atribuição explícita de atitudes proposicionais).

CONCLUSÃO

A literatura filosófica sobre a Psicologia do Senso Comum se estende desde a década de 1970, e abrange diversas questões sobre nosso entendimento interpessoal cotidiano, nossa capacidade de interação e coordenação de atividades, o arcabouço conceitual intuitivo que relaciona estados mentais e atitudes proposicionais a comportamentos, e os mecanismos cognitivos de "leitura mental" que nos permitem atribuir estados mentais a outras pessoas. Nesta dissertação eu examinei o desenvolvimento histórico desta literatura, identificando dois debates distintos, o primeiro (principalmente entre Paul Churchland e Jerry Fodor dos anos 70 aos anos 90) tendo como foco a relação entre a "teoria da Psicologia do Senso Comum" e teorias científicas (da neurociência e das ciências cognitivas), e o segundo (o debate contemporâneo) tendo como foco os mecanismos cognitivos de leitura mental e o papel das atribuições de estados mentais e atitudes proposicionais nas teorias da cognição corporificada, situada e estendida. Além do exame histórico do que argumentei serem dois debates distintos e da transição conceitual entre ambos, também apresentei aqui minha crítica à abordagem eliminativista contemporânea de Matthew Ratcliffe e, como alternativa, articulei os princípios de uma abordagem pluralista que combina leitura mental e interpretação contextual situada como fundamentais para a cognição social.

Por assumirem que nossa capacidade de compreender outras pessoas e explicar e prever seu comportamento resulta de nossa aplicação de uma *teoria empírica implícita* composta de generalizações nômicas sobre a relação entre estados mentais e comportamentos, as questões centrais para Churchland e Fodor dizem respeito à validade dessa teoria, a compatibilidade de sua ontologia mentalista com a de teorias

científicas como a neurociência e as ciências cognitivas (bem como a autonomia ou interdependência desses diversos campos da ciência entre si), e, portanto, o futuro do nosso entendimento ordinário de pessoas e seus comportamentos e estados mentais (se ele será vindicado pela ciência, no caso de Fodor, ou se será eliminado por ela, como defende Churchland).

A distinção entre sentidos internalistas e externalistas de "teoria", introduzida por Stich e Ravenscroft (1996), mostra que a discussão entre Churchland e Fodor sobre a Psicologia do Senso Comum é ambígua e confusa. Tanto Churchland quanto Fodor pressupõem que as pessoas possuem uma teoria sobre comportamentos e estados mentais da qual elas fazem uso para interpretar a si mesmas e às outras pessoas com quem interagem; essa teoria (no sentido internalista) bem como sua ontologia mentalista eles chamam de Psicologia do Senso Comum. Deste ponto de partida comum, Fodor articula uma teoria que busca vindicar nossa ontologia mentalista intuitiva e aproximá-la das ciências cognitivas; esta teoria (no sentido externalista, como uma sistematização conceitual explícita) ele defende como a representação mais fiel da teoria internalizada, e por isso frequentemente refere-se a ela como a Psicologia do Senso Comum; por outro lado, o eliminativismo de Churchland defende que a "Psicologia do Senso Comum é uma teoria falsa", às vezes com isso querendo dizer que a teoria (no sentido internalista) que as pessoas utilizam para guiar suas interações e interpretações mentalistas comportamentais é equivocada, e outras vezes com isso defendendo que a teoria (no sentido externalista) articulada por Fodor é inadequada e que não há espaço na ciência para nossas categorias mentalistas ordinárias.

Matthew Ratcliffe chama sua abordagem de um tipo de eliminativismo, segundo

o qual "leitura mental" e "atribuição de estados mentais" são meros construtos de sistematizações filosóficas e não possuem denotação ou extensão; para ele nosso entendimento interpessoal ordinário é fundamentalmente interativo e nossas interpretações comportamentais são baseadas nas normas sociais do ambiente compartilhado, e não em atribuições de estados mentais intencionais; assim, enquanto o eliminativismo original de Churchland previa que a Psicologia do Senso Comum é uma teoria falsa[14] que as pessoas utilizam no dia a dia e que deverá ser *eliminada* em favor de uma teoria científica mais adequada, o eliminativismo de Ratcliffe propõe que a Psicologia do Senso Comum é uma teoria falsa[15] que os filósofos criaram a respeito do que eles pensam que as pessoas fazem no dia a dia e que ela deve ser *eliminada* porque não corresponde a nada no mundo real – isto é, para ele a Psicologia do Senso Comum de que falam os filósofos como baseada na leitura mental simplesmente não existe.

Minha crítica à abordagem eliminativista de Ratcliffe é apresentada em argumentos independentes. A seção 4.2.1 tratou de dois cenários de colapso como contraexemplos à proposta de Ratcliffe, argumentando que seu modelo de cognição social é equivocado, primeiro porque *ao menos às vezes* a situação compartilhada não é suficiente para facilitar o entendimento interpessoal, e assim a atribuição de estados mentais é requerida para que interpretemos os outros como pessoas, e segundo porque a atribuição de estados mentais às outras pessoas é algo que tentamos fazer mesmo quando não é necessário ou particularmente informativo fazê-lo. A seção 4.2.2 apresentou o argumento de que o papel que Ratcliffe dedica às normas sociais e normas

[14] Como Churchland escreveu a maior parte de seus artigos sobre o eliminativismo antes que Stich tivesse proposto a distinção internalista/externalista, ele não posiciona sua abordagem à luz de tal distinção; em alguns momentos ele critica a PSC como uma teoria que as pessoas utilizam no dia a dia (portanto, no sentido internalista), e em outros momentos ele questiona a PSC conforme ela foi sistematizada por Fodor (portanto, no sentido externalista).

[15] No sentido externalista.

da racionalidade, mesmo em uma perspectiva externalista, é incompatível com sua rejeição da leitura mental: mesmo se concedemos que as normas são constituintes objetivos da situação compartilhada (que elas "*estão aí para todos seguirem*"), nossos encontros subjetivos ainda envolvem o reconhecimento nos outros de uma capacidade (ou sua ausência) de entender e seguir as normas objetivas, tanto sociais e quando da racionalidade; isso pode requerer a atribuição completa de atitudes proposicionais (e.g. quando julgamos que alguém é mal-intencionado), mas também pode constituir apenas um tipo mínimo de leitura mental (e.g. na percepção de que uma pessoa é inocente ou incapaz de compreender as regras básicas da racionalidade).

Nas três subseções da minha crítica à proposta de Ratcliffe, concluí que não é apenas conceitualmente viável que Ratcliffe adote uma perspectiva pluralista (4.2.3), mas que é de fato recomendável que ele o faça já que a ausência da leitura mental torna seu modelo de cognição social inadequado (4.2.1) e inconsistente (4.2.2). A implicação disso para o debate geral sobre a Psicologia do Senso Comum é que a leitura mental continua sendo um fenômeno que nossas sistematizações filosóficas e científicas (teorias no sentido externalista) precisam explicar, e isso porque ela é um fato da experiência cotidiana de como nós interagimos com outras pessoas e explicamos e prevemos seu comportamento. Na seção final da dissertação (4.3), articulei alguns princípios, inspirados no trabalho da filósofa Kristin Andrews, que definem uma abordagem pluralista à cognição social, combinando leitura mental e interpretação contextual situada como componentes indispensáveis de nosso entendimento interpessoal, e portanto indispensáveis também para uma teoria adequada da Psicologia do Senso Comum.

REFERÊNCIAS

ANDREWS, Kristin (2012) *Do Apes Read Minds?: Toward a New Folk Psychology*. MIT Press.

ANSCOMBE, G. Elizabeth M. (1957) *Intention.* Harvard University Press, 2000. Originalmente, Basil Blackwell.

AUDI, R. (2001) *The architecture of reason: the structure and substance of rationality.* Oxford University Press.

ALBRIGHT, T. D., & NEVILLE, H. J. (1999) Neurosciences. In: T*he MIT Encyclopedia of the Cognitive Sciences*. Robert A. Wilson & Frank C. Keil (Eds.). MIT Press. Cambridge, Massachusetts.

ARBIB, M. (2003) The Elements of Brain Theory and Neural Networks. In: The Handbook of Brain Theory and Neural Networks. Michael A. Arbib (Ed.). MIT Press. Cambridge, Massachusetts.

BAKER, Lynne R. (1999) Folk Psychology. In: *The MIT Encyclopedia of the Cognitive Sciences*. Robert A. Wilson & Frank C. Keil (Eds.). MIT Press. Cambridge, Massachusetts.

BAKER, Lynne R. (2009) Persons and the Extended-Mind Thesis. In: *Zygon: Journal of Religion and Science*, 44(3):642–658

BECHTEL, W., ABRAHAMSEN, A., & GRAHAM, G. (1999), The Life of Cognitive Science. In: *A Companion to Cognitive Science*. William Bechtel and George Graham (Eds.). Blackwell Publishing.

BECHTEL, W., & ZAWIDZKY, T., (1999), Selective biographies of major contributors to cognitive science. In: *A Companion to Cognitive Science*. William Bechtel and George Graham (Eds.). Blackwell Publishing.

BUTTELMANN, D., CARPENTER, M., & TOMASELLO, M. (2009) Eighteen-month-old infants show false belief understanding in an active helping paradigm. In: *Cognition,* 112:337–342. Elsevier.

CAT, J. (2013) The Unity of Science. In: *Stanford Encyclopedia of Philosophy.*

CHEMERO, A. (2002) Editor's Introduction: Reconsidering Ryle. In: *Electronic Journal of Analytic Philosophy*, Vol. 7. ISSN 1071-5800.

CHEMERO, T. (2009) *Radical Embodied Cognitive Science*. MIT Press.

CHOMSKY, N. (1959) Verbal Behavior by B. F. Skinner. In: *Language*, Vol. 35, No. 1 (Jan. - Mar., 1959), pp. 26-58; Linguistic Society of America.

CHURCHLAND, Paul (1970) The Logical Character of Action-Explanations. *The Philosophical Review.* Vol. 79, No. 2, pp. 214-236.

CHURCHLAND, Paul (1981) Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy.* Vol. 78, No. 2, pp. 67-90 (1981).

CHURCHLAND, Paul (1988) *Matter and Consciousness*. MIT Press.

CHURCHLAND, Paul (1996) Folk Psychology. In *On the contrary: critical essays, 1987-1997.* MIT Press, 1999. Originalmente publicado em *Companion to the Mind*, ed. S. Guttenplan, Oxford: Blackwells.

CHURCHLAND, Paul, & CHURCHLAND, Patricia (1990) Intertheoretic Reduction: A neuroscientist's field guide. In *On the contrary: critical essays, 1987-1997.* MIT Press, 1999. Originalmente publicado em *Seminars in the Neurosciences*, 249-56. (1990).

CLANCEY, W. (2009) Scientific Antecedents of Situated Cognition. In: *The Cambridge Handbook of Situated Cognition*. Philip Robbins & Murat Aydede (Eds.). Cambridge University Press.

CLARK, A. (1998) Magic Words: How Language Augments Human Computation. In: *Language And Thought: Interdisciplinary Themes.* P. Carruthers and J. Boucher (Eds). pp. 162-183. Cambridge University Press.

CLARK, A., & CHALMERS, D. (1998) The Extended Mind. In: *Analysis* 58:10-23.

CRANE, T., & MCLAUGHLIN, B. P. (2009) Introduction. In: *Synthese* 170:211-15. DOI 10.1007/s11229-009-9579-8

CREATH, R. (2011) Logical Empiricism. In: *Stanford Encyclopedia of Philosophy.*

DAVIDSON, Donald (2001) *Subjective, Intersubjective, Objective*. Oxford University Press.

DAVIES, M., & STONE, T. (1995) *Folk Psychology.* Blackwell.

DENNETT, D. (2002) Re-Introducing The Concept of Mind. In: Electronic Journal of Analytic Philosophy, Vol. 7. ISSN 1071-5800.

FARKAS, K. (2012) Two versions of the extended mind thesis. In: Philosophia, Volume 40, # 3, pp 435-447. DOI: http://dx.doi.org/10.1007/s11406-011-9355-0

FODOR, Jerry (1975) *The Language of Thought.* Harvard University Press.

FODOR, Jerry (1980) Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology. In: *Behavioral and Brain Sciences* 3, pp. 68-73. Cambridge University Press.

FODOR, Jerry (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind.* MIT Press.

FODOR, Jerry (1990) *A Theory of Content and Other Essays.* MIT Press.

FODOR, Jerry (2000) *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology.* MIT Press.

FODOR, Jerry (2008) *LOT 2: The Language of Thought Revisited.* Oxford University Press.

GALLESE, V. (2007) Embodied simulation: from mirror neuron systems to interpersonal relations. In: *Empathy and Fairness*. Wiley, Chichester. Novartis Foundation Symposium, 278: 3-12

GALLESE, V., & GOLDMAN, A. (1998) Mirror Neurons and the Simulation Theory of Mindreading. In: *Trends in Cognitive Sciences*, Vol. 2, N. 12, Dec.

GIBSON, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.

GIERE, R. (2002) Scientific Cognition as Distributed Cognition. In: *Cognitive Bases of Science*. Peter Carruthers, Stephen Stitch, and Michael Siegal (Eds.). Cambridge University Press.

GOLDMAN, A. (2008) Mirroring, Mindreading, and Simulation. In: *Mirror Neuron Systems: The Role of Mirroring Processes in Social Cognition*, J. A. Pineda, (ed.), pp. 311-330. New York: Humana Press.

GOLDMAN, A., & MASON, K. (2006), Simulation. In: *Handbook of the Philosophy of Scienc*e, vol. 12, Philosophy of Psychology, Paul Thagard (ed.). Elsevier.

GORDON, R. (1986) Folk Psychology as Simulation. In: *Mind and Language*, 1: 158–171; reprinted in M. Davies and T. Stone (eds.), Folk Psychology: The Theory of Mind Debate, Oxford: Blackwell Publishers (1995).

GORDON, R. (2009) Folk Psychology as Mental Simulation. In: *Stanford Encyclopedia of Philosophy.*

GOODMAN, Russell (2013) William James. In: *Stanford Encyclopedia of Philosophy.*

GRAHAM, G. (2010). Behaviorism. In: *Stanford Encyclopedia of Philosophy.*

HAUSER, Larry (2005) Behaviorism. In: *Internet Encyclopedia of Philosophy.*

HEAL, J. (1986) Replication and Functionalism. In: *Language, Mind, and Logic*, J. Butterfield (ed.), Cambridge: Cambridge University Press; reprinted in M. Davies and T. Stone (eds.), Folk Psychology: The Theory of Mind Debate, Oxford: Blackwell Publishers (1995).

HEMPEL, Carl (1942) The Function of General Laws in History. *The Journal of Philosophy*, Vol. 39, No. 2, pp. 35-48.

HORST, S. (2009) The Computational Theory of Mind. In: *Stanford Encyclopedia of Philosophy.*

HUTCHINS, E. (1995) *Cognition In The Wild*. MIT Press.

JORDAN, M. I., & RUSSELL, S. (1999) Computational Intelligence. In: *MIT Encyclopedia of the Cognitive Sciences*. Robert A. Wilson and Frank C. Keil (Eds.) The MIT Press. Cambridge, Massachusetts.

KIM, Alan (2006) Wilhelm Maximilian Wundt. In: *Stanford Encyclopedia of Philosophy.*

KIRSH, D. (2009) Problem Solving and Situated Cognition. In: *The Cambridge Handbook of Situated Cognition*. Philip Robbins & Murat Aydede (Eds.). Cambridge University Press.

KLEMENT, K. (2013) Russell's Logical Atomism. In: *Stanford Encyclopedia of Philosophy.*

LAKOFF, G., & JOHNSON, M. (1999) *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Basic Books.

LEVIN, J. (2013) Functionalism. In: *Stanford Encyclopedia of Philosophy.*

MARRAFFA, M. (2011) Theory of Mind. In: *Internet Encyclopedia of Philosophy.*

MCCANN, H. J. (1999) Practical Reasoning. In: *The Cambridge Dictionary of Philosophy*, Second Edition. Robert Audi (Ed.). Cambridge University Press.

MOREIRA, M. B., & MEDEIROS, C. A. (2007) *Princípios Básicos de Análise do Comportamento*. Artmed.

MORTON, A. (1980) *Frames of Mind*. Oxford University Press.

NERSESSIAN, N. (2006) Model-Based Reasoning in Distributed Cognitive Systems. *Philosophy of Science*, 73(5): 699–709.

NICHOLS, S. (no prelo) Mindreading and the Philosophy of Mind. In J. Prinz (ed.) *The Oxford Handbook on Philosophy of Psychology*. New York: Oxford University Press.

NICHOLS, Shaun, & STICH, Stephen (2003) *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press. (2003)

ONISHI, K. H., & BAILLARGEON, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258.

PAKALUK, M. (2005) *Aristotle's Nicomachean Ethics: An Introduction*. Cambridge University Press.

PARKER-RYAN, S. (2012) Ordinary Language Philosophy. In: *Internet Encyclopedia of Philosophy.*

PRESTON, A. (2006) Analytic Philosophy. In: *Internet Encyclopedia of Philosophy.*

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

RATCLIFFE, M. (2006) 'Folk Psychology' Is Not Folk Psychology. In: *Phenomenology and the Cognitive Sciences* (2006) 5: 31–52. DOI: 10.1007/s11097-005-9010-y. Springer.

RIVES, B. (2010) Fodor, Jerry. In: *Internet Encyclopedia of Philosophy.*

RIZZOLATTI G., FOGASSI L. & GALLESE V. (2006) Mirrors in the mind. In: *Scientific American*. Nov, 2006, 295(5): 54-61.

ROBBINS, P., & AYDEDE, M. (2009) A Short Primer on Situated Cognition. In: *The Cambridge Handbook of Situated Cognition*. Philip Robbins & Murat Aydede (Eds.). Cambridge University Press.

RUSSELL, Bertrand (1971) The Philosophy of Logical Atomism. In: *Logic and Knowledge; Essays 1901-1950*. Robert Charles Marsh (Ed.). New York: Capricorn Books, 1971. Originalmente publicado em *The Monist*, 28, 495-527; 29, 32-63, 190-222, 345-380 (1918, 1919).

RUSSELL, S. J., & NORVIG, P. (1995) *Artificial Intelligence: A Modern Approach*. Prentice-Hall. Englewood Cliffs, New Jersey.

RYLE, G. (2009) *The Concept of Mind. Routledge*. Originalmente 1949.

SEARLE, John (2004) *Mind: A Brief Introduction*. Oxford University Press.

SEARLE, John (1983) *Intentionality: An Essay in the Philosophy of Mind.* Cambridge University Press.

SCHWITZGEBEL, Eric (2010) Belief. In: *Stanford Encyclopedia of Philosophy.*

SHAPIRO, L. (2007) The Embodied Cognition Research Programme. In: *Philosophy Compass*, 2/2: 338-346, 10.1111/j.1747-9991.2007.00064.x

SHAPIRO, L. (2011) *Embodied Cognition*. Routledge.

SKINNER, B. F. (1957) *Verbal Behavior.* New York: Appleton-Century-Crofts, Inc.

STICH, Stephen (1983) *From Folk Pschology to Cognitive Science: The Case Against Belief.* MIT Press.

STICH, Stephen (1996) *Deconstructing the Mind.* Oxford University Press.

STICH, Stephen, & NICHOLS, Shaun (2003) Folk Psychology. In: *The Blackwell Guide to Philosophy of Mind*, Stephen Stich & Ted A. Warfield, (eds.). Oxford: Basil Blackwell, pp. 235-255.

STICH, S., & RAVENSCROFT, I. (1996) What is Folk Psychology?. In: *Deconstructing the Mind* (S. Stich). Oxford University Press.

UEBEL, T. (2011) Vienna Circle. In: *Stanford Encyclopedia of Philosophy.*

VARELA, F., THOMPSON, E. and ROSCH, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, MA: MIT Press.

WATSON, John B. (2008) *Clássico traduzido: a psicologia como o behaviorista a vê*. Temas em Psicologia, v. 16, n.2, pp. 289-301.

WATSON, John B. (1930) *Behaviorism*. Norton: New York.

WILSON, R. A., & FOGLIA, L. (2011) Embodied Cognition. In: *Stanford Encyclopedia of Philosophy.*

WIMMER, H., & PERNER, J. (1983) Beliefs About Beliefs: Representation and Constraining Function of Wrong Beliefs in Young Children's Understanding of Deception. In: *Cognition*, 13:103-128.

WITTGENSTEIN, L. (1921) *Tractatus Logico-Philosophicus*. C.K. Ogden (trad.), London: Routledge & Kegan Paul.

WITTGENSTEIN, L. (1953) *Philosophical Investigations.* G.E.M. Anscombe and R. Rhees (eds.), G.E.M. Anscombe (trad.), Oxford: Blackwell.