



**Universidade do Estado do Rio de Janeiro**

Centro Biomédico

Instituto de Medicina Social

Cláudia Maria Passos Ferreira

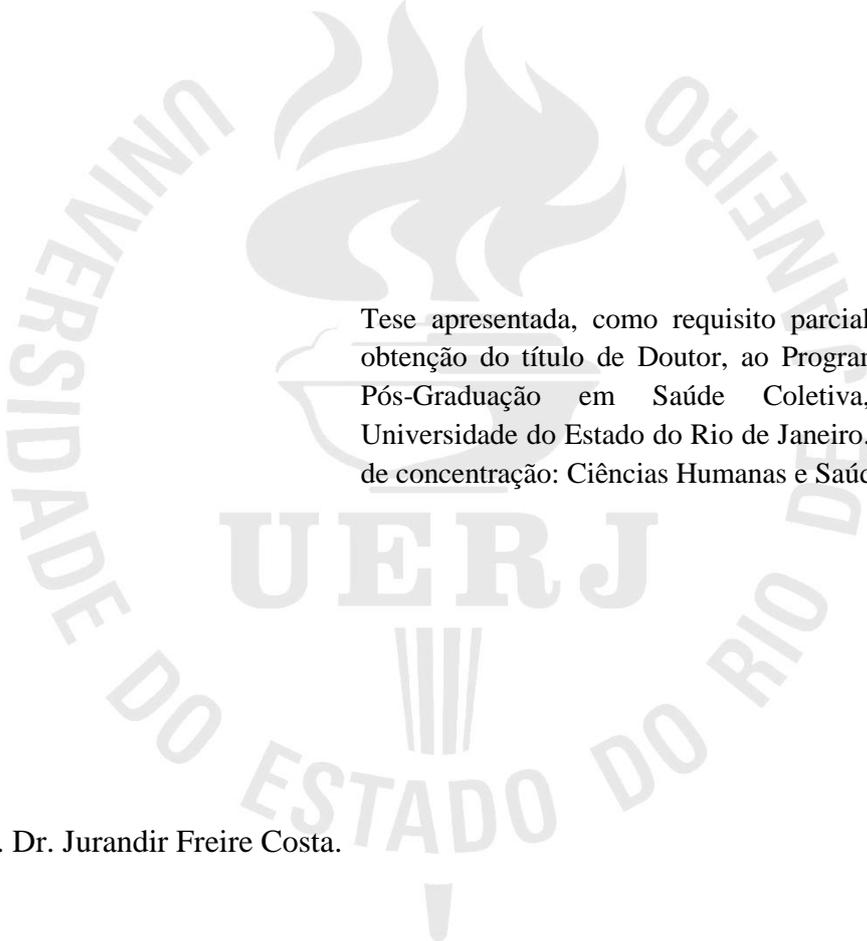
**Origens da interioridade: autoconhecimento e externalismo**

Rio de Janeiro

2006

Cláudia Maria Passos Ferreira

**Origens da interioridade: autoconhecimento e externalismo**



Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Saúde Coletiva, da Universidade do Estado do Rio de Janeiro. Área de concentração: Ciências Humanas e Saúde.

Orientador: Prof. Dr. Jurandir Freire Costa.

Rio de Janeiro

2006

CATALOGAÇÃO NA FONTE  
UERJ / REDE SIRIUS / BIBLIOTECA CB-C

F383 Ferreira, Cláudia Maria Passos.  
Origens da interioridade: autoconhecimento e externalismo /  
Cláudia Passos Ferreira. – 2006.  
152f.

Orientador: Jurandir Freire Costa.  
Tese (doutorado) – Universidade do Estado do Rio de Janeiro,  
Instituto de Medicina Social.

1. Self (Psicologia) – Teses. 2. Interiorização – Teses. 3. Teoria da autoconsciência – Teses. 4. Consciência – Teses. I. Costa, Jurandir Freire. II. Universidade do Estado do Rio de Janeiro . Instituto de Medicina Social. III.Título.

CDU 159.923.2

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese, desde que citada da fonte.

---

Assinatura

---

Data

Cláudia Maria Passos Ferreira

**Origens da interioridade: autoconhecimento e externalismo**

Tese apresentada, como requisito parcial para obtenção do título de Doutor, ao Programa de Pós-Graduação em Saúde Coletiva, da Universidade do Estado do Rio de Janeiro.  
Área de concentração: Epidemiologia.

Aprovada em 29 de maio de 2006.

Banca Examinadora:

---

Prof. Dr. Jurandir Freire Costa (Orientador)  
Instituto de Medicina Social – UERJ

---

Prof. Dr. Benilton Carlos Bezerra Junior  
Instituto de Medicina Social – UERJ

---

Prof. Dr. Francisco Javier Ortega  
Instituto de Medicina Social – UERJ

---

Prof. Dr. Octávio Domont De Serpa Junior  
Instituto de Psiquiatria – IPUB – UFRJ

---

Prof. Dr. Roberto Horácio de Sá Pereira  
Instituto de Filosofia e Ciências Sociais – IFCS – UFRJ

Rio de Janeiro

2006

## DEDICATÓRIA

Aos meus pais e meus irmãos.

A Marcelo, without whom not.

## AGRADECIMENTO

Esta tese é fruto de um trabalho de pesquisa desenvolvido junto ao PEPAS/ IMS desde 2001. Agradeço aos colaboradores do PEPAS, em particular a Jurandir Freire Costa, pelo desafio de pensar os problemas subjetivos pertinentes a nosso mundo contemporâneo. A pesquisa foi desenvolvida em duas etapas, uma no Brasil e outra na França. Durante a fase brasileira, contei com o apoio da FAPERJ, e durante a fase francesa, contei com o apoio da CAPES. Os apoios dessas agências, como de seus funcionários, e da equipe do IMS viabilizaram a pesquisa.

Além da interlocução com os membros do PEPAS, três outros lugares foram de importância fundamental para os rumos teóricos do trabalho. Os cursos de filosofia do IFCS com Roberto Horácio e Wilson Mendonça, cuja clareza argumentativa e o interesse vivo pela filosofia me permitiram o entendimento de questões fundamentais da tese. As diversas atividades do Institut Jean Nicod em Paris, que me permitiram expandir os horizontes de meu pensamento brasileiro. E o *Séminaire Sauvage* com Philippe Rochat e Pedro Salem no Café e l'École Militaire, que, pelo exercício permanente de *notre pensée triadique*, me permitiram criar uma bela memória para minha vida. Esses encontros teóricos trouxeram um enriquecimento inimaginável para as idéias aqui apresentadas.

Algumas pessoas contribuíram de forma valiosa com comentários, sugestões e críticas.

Agradeço em especial a Jurandir Freire Costa, por sua orientação precisa, sua ousadia intelectual e seu compromisso ético, fonte permanente de inspiração. Agradeço também a Benilton Bezerra, Bernard Andrieu, Fernando Vidal, Francisco Ortega, Marcelo Land, Octavio Serpa, Pedro Salem, Philippe Rochat, Roberto Horácio, que, em momentos diferentes, discutiram pontos relevantes, que contribuíram de forma significativa para as questões apresentadas na tese. Marcelo Land leu os manuscritos, ajudando a rever erros e imprecisões dos argumentos.

Outras pessoas contribuíram com inúmeros gestos de carinho e atenção, que tornaram minha trajetória muito mais agradável. Aos meus amigos brasileiros, Janne Calhau, Denise, Gabi, Rossano, Nando, Carla, Célia, Bethânia, Nanda, Lulli,. Aos meus amigos 'franceses', Marcos, Alessandra, Bernard, Pascal, Pérola, Darry, Luciana e Hélène, pelos inúmeros esforços em tornar familiar a experiência do estrangeiro em mim.

Agradeço também à equipe do COIJ do Instituto Philippe Pinel, em especial Adriana Gonzaga e Fernando Ramos, cuja generosidade e confiança no meu trabalho me possibilitou uma dedicação exclusiva a esse projeto.

Agradeço ao meu pai pelas correções da redação final da tese e à minha família pelo suporte emocional constante.

## RESUMO

FERREIRA, Claudia Maria Passos. Origens da interioridade: autoconhecimento e externalismo. 2006. 152 f. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro. 2006.

O objetivo da tese é investigar a constituição da interioridade a partir de uma abordagem externalista. Os processos pelos quais o autoconhecimento é constituído são considerados como estando associados ao desenvolvimento da perspectiva da primeira pessoa. Adotar uma perspectiva de primeira pessoa é tornar-se capaz de fazer referência a si mesmo e conhecer seus próprios estados mentais e corporais. A autoconsciência e o autoconhecimento foram tradicionalmente subsumidos à idéia de Descartes da autoridade da primeira pessoa. Segundo a tese cartesiana, teríamos acesso privilegiado e não-empírico aos nossos estados mentais que se expressaria por meio de um conhecimento. A tese central do externalismo afirma, ao contrário, que o conteúdo dos estados mentais é constituído, em parte, pelas relações com o ambiente. A adoção da tese externalista coloca em dúvida a suposição cartesiana de que temos acesso privilegiado aos conteúdos de nossos pensamentos, restringindo, assim, a autoridade da primeira pessoa. O externalismo perceptivo de Davidson, por exemplo, oferece uma solução ternária eu-intérprete-mundo para as origens do autoconhecimento. A tese de Davidson é apresentada como reconciliando o autoconhecimento e as idéias centrais do externalismo. São apresentados dois modelos da gênese externa do eu: 1) os modelos ecológicos que propõem um desenvolvimento do sentido de eu como uma função das interações do eu com o ambiente; 2) o modelo psicológico de Winnicott que propõe a emergência do sentido de eu a partir de uma relação ternária entre o eu, os outros e os objetos transicionais. Defendemos a tese de que o modelo psicológico de Winnicott é o mais adequado para descrever a conceitualização epistemológica de Davidson das origens externalistas do autoconhecimento.

**Palavras-chave:** Self. Interiorização. Autoconsciência. Autoconhecimento. Externalismo.

## ABSTRACT

FERREIRA, Claudia Maria Passos. *Origins of interiority: self-knowledge and externalism*. 2006. 152 f. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro. 2006.

The aim of this thesis is to study the constitution of interiority from an externalist perspective. The process by which self-knowledge is formed is considered in relation to the development of the first person perspective. From this first person perspective, one is capable to self-refer, to know one's own mental and physical states. Self-consciousness and selfknowledge are discussed in relation to Descartes' idea of an authority of first person perspective. The Cartesian idea contends that this perspective gives privilege and nonempirical access to one's own mental states. On the contrary, externalism's central thesis contends that the content of the knowledge pertaining to one's own mental states originates from interactions with the environment. Thus, the externalist's thesis puts into question the Cartesian idea of a privileged epistemic access of the first person. The perceptual externalism of Davidson, for example, proposes a triadic solution to the origins of self-knowledge. Davidson's thesis is presented as reconciling self-knowledge and the basic externalist ideas. Two externalist models on the origins of self-knowledge are further discussed: 1) ecological models proposing a development of a sense of self as a function of interactions with the environment; 2) the psychological model of Winnicott proposing that the developing sense of self arises from the triadic relation between self, others, and transitional object(s). We defend the thesis that Winnicott's model is the most apt at describing the epistemological conceptualizing of Davidson regarding the externalist origins of self-knowledge.

**Keywords:** Internalization. Self-awareness. Self-knowledge. Externalism.

## RESUMÉ

FERREIRA, Claudia Maria Passos. Origines de l'intériorité connaissance du soi et externalisme. 2006. 152 f. Tese (Doutorado em Saúde Coletiva) – Instituto de Medicina Social, Universidade do Estado do Rio de Janeiro. Rio de Janeiro. 2006.

L'objectif de cette thèse est d'étudier la constitution de l'intériorité dans une démarche externaliste de la philosophie de l'esprit. Les processus de la constitution de la connaissance du soi sont envisagés comme étant liés au développement de la perspective de la première personne. Adopter la perspective de la première personne, c'est être capable soi-même de faire référence à soi, de connaître ses propres états mentaux et corporels. La conscience et la connaissance du soi étaient mises en relation avec à l'idée Cartésienne de l'autorité de la première personne. Selon la thèse de Descartes, nous devrions avoir un accès privilégié et nonempirique de nos propres états mentaux, et cet accès s'exprimait par l'intermédiaire d'une connaissance. La thèse centrale de l'externalisme, au contraire, soutient que le contenu des états mentaux est constitué par des relations avec l'environnement. L'adoption de la thèse externaliste met en doute l'idée Cartésienne d'un accès épistémique privilégié de la première personne. L'externalisme perceptive de Davidson, par exemple, propose une solution triadique du problème de l'origine de la connaissance du soi. Cette thèse de Davidson est présentée comme une alternative qui rend compatible la connaissance du soi et les idées externalistes. Deux modèles externalistes sur l'origine du soi sont aussi discutés : 1) les modèles écologistes qui proposent une genèse du sens du soi en fonction des interactions avec l'environnement ; 2) le modèle de Winnicott qui propose que la genèse du soi naisse de la relation triadique entre le soi, l'autrui, et l'objet transitionnel. Nous soutenons la thèse que le modèle psychologique de Winnicott est le plus approprié pour décrire la conceptualisation.

Mots-clés: Internalisation. Conscience de soi. Connaissance de soi. Externalism.

## SUMÁRIO

|       |  |     |
|-------|--|-----|
|       | <b>INTRODUÇÃO</b> .....  | 9   |
| 1     | <b>A PERSPECTIVA DA PRIMEIRA PESSOA</b> .....                      | 19  |
| 1.1   | <b>O que é se sentir como nós</b> .....                            | 19  |
| 1.2   | <b>Tornar-se uma perspectiva</b> .....                             | 23  |
| 1.2.1 | <u>O caráter subjetivo da consciência</u> .....                    | 23  |
| 1.2.2 | <u>A emergência da autoconsciência</u> .....                       | 28  |
| 1.3   | <b>Autoconsciência e autoconhecimento</b> .....                    | 39  |
| 1.3.1 | <u>Autoconhecimento e introspecção</u> .....                       | 42  |
| 1.3.2 | <u>Autoconsciência e “Eu-mentais”</u> .....                        | 49  |
| 1.3.3 | <u>Autoconhecimento e externalismo</u> .....                       | 54  |
| 2     | <b>EXTERNALISMO DO MENTAL</b> .....                                | 63  |
| 2.1   | <b>O debate individualismo versus externalismo</b> .....           | 63  |
| 2.2   | <b>Individualismo</b> .....  | 66  |
| 2.2.1 | <u>Descartes e o individualismo</u> .....                          | 66  |
| 2.2.2 | <u>Superveniência e relevância causal</u> .....                    | 71  |
| 2.3   | <b>Externalismo</b> .....  | 81  |
| 2.3.1 | <u>Os modelos de externalismo</u> .....                            | 82  |
| 2.3.2 | <u>O externalismo semântico de Putnam</u> .....                    | 83  |
| 2.3.3 | <u>O externalismo social de Burges</u> .....                       | 90  |
| 2.3.4 | <u>O externalismo das Terras Gêmeas e o autoconhecimento</u> ..... | 96  |
| 2.3.5 | <u>O externalismo perceptivo de Davidson</u> .....                 | 100 |
| 3     | <b>AS FONTES EXTERNAS DO EU</b> .....                              | 113 |
| 3.1   | <b>O lugar em que vivemos</b> .....                                | 113 |
| 3.2   | <b>A gênese dual do eu</b> .....                                   | 117 |

|       |   |     |
|-------|---|-----|
| 3.2.1 | <u>O <i>self</i> agente de Neisser</u> .....      | 119 |
| 3.2.2 | <u>O <i>self</i> narrativo de Dennett</u> .....   | 123 |
| 3.3   | <b>Triangulação Eu-Outro-Objeto</b> .....         | 128 |
| 3.3.1 | <u>O <i>self</i> expandido de Winnicott</u> ..... | 128 |
|       | <b>CONSIDERAÇÕES FINAIS</b> .....                 | 139 |
|       | <b>REFERÊNCIAS</b> .....                          | 142 |

## INTRODUÇÃO

*“In our languages of self-understanding, the opposition ‘inside-outside’ plays an important role. We think of our thoughts, ideas, or feelings as being “within” us, while the objects in the world which these mental states bear on are “without”. Or else we think of our capacities or potentialities as “inner”, awaiting the development which will manifest them or realize them in the public world.” (Taylor 1997: 111)*

Interno, interior e interioridade são palavras – adjetivos ou substantivos – usadas para descrever tudo que estaria dentro e se diferenciaria de tudo que estaria do lado de fora. Heurísticamente, usam-se os adjetivos *interno* para qualificar o mental e *externo* para qualificar tudo que se oporia ao mental. Nessa acepção, interno-externo confunde-se com dentro-fora da cabeça ou do corpo. Ao definir o interior em oposição ao exterior, criam-se pólos opostos que se co-determinam: mente-mundo, indivíduo-cultura, organismo-ambiente, eu-outro. Um dos problemas dessas oposições gira entorno das explicações sobre as determinações entre a mente e o mundo. Essa cisão entre o *interior mental* e o *exterior ambiental* culminou na busca teórica por intermediários entre a mente e o mundo.

Intuitivamente, diríamos que a mente existe *dentro* do indivíduo e os estados mentais ocorrem *dentro da cabeça*. Nesse sentido, as fronteiras da mente seriam determinadas pelas fronteiras do indivíduo. Nas ciências do Ocidente, a mente sempre foi concebida como delimitada por indivíduos e existindo dentro de indivíduos, constituindo assim aquilo que em grande parte os caracteriza e configura suas identidades. Os limites da mente foram identificados aos limites, senão da extensão do cérebro, ao menos da extensão da pele. Tanto a fronteira temporal da mente quanto a espacial seriam determinadas pelo corpo e pelo tempo de existência do indivíduo.

Essa concepção interiorizada da mente é a metáfora privilegiada do indivíduo moderno. O indivíduo moderno ocidental, como o retrata Taylor (1997), é aquele que rompeu com a transcendência e acredita poder encontrar suas coordenadas dentro de si mesmo, independente das redes sociais de interlocução que o constituíram. A identidade moderna caracteriza-se pela idéia de que somos um self introspectivo separado do mundo externo. Do ponto de vista epistêmico, o self possuiria uma percepção de si como um espaço interior que

guardaria pensamentos, desejos e intenções. A consciência seria algo independente, localizado no interior do corpo e, na sua vertente reducionista mais recente, no interior do cérebro<sup>9</sup>.

Essa forma de conceber a mente e o self adota o *indivíduo* como entidade paradigmática de análise do comportamento humano e dos estados psicológicos. O *indivíduo* como paradigma tem sido uma tendência constante nas teorias explicativas geradas pela psicologia, biologia e ciências sociais. Em *Boundaries of the mind*, Robert Wilson (2004) mostra como a psicologia, biologia e ciências sociais, coletivamente batizadas de *ciências frágeis*, estão organizadas, grosso modo, em torno de duas tendências teóricas opostas: o *internalismo*, que defende uma perspectiva individualista, e o *externalismo*, que enfatiza o papel determinante do ambiente externo na explicação do mental. Ambas as tendências expressam paradigmas predominantes na exploração da mente.

O *individualismo*<sup>10</sup> em psicologia afirma que os estados psicológicos deveriam ser construídos sem referência a qualquer coisa além da fronteira do indivíduo. As teses individualistas estão na origem do mentalismo cartesiano e ganharam novo fôlego no debate contemporâneo com o nativismo, fruto da virada lingüística e das ciências cognitivas (Wilson 2004). Os estudos sobre o funcionamento mental e as estruturas lingüísticas inatas originaram modelos internalistas da mente e uma semântica individualista que enfatizam o papel central que as propriedades internas do indivíduo ocupam na economia causal da mente.

Em reação aos reducionismos e equívocos promovidos pelos modelos explicativos individualistas, vêm ganhando força as explicações externalistas. O *externalismo* da mente defende a idéia de que os conteúdos de nossos pensamentos são determinados por nossos

---

<sup>9</sup> Ao analisar a antropologia do indivíduo contemporâneo, autores como Bauman, (1998), Sennett (2002), Costa (2004), Ehrenberg (2000, 2004), Bezerra, (1996), Andrieu (2003, 2004), Ortega (2002) e Vidal (2002) mostram o surgimento de configurações identitárias que buscam a localização material da interioridade, seja no cérebro ou nas experiências sensoriais. Ao longo do século XX, observa-se o deslocamento crescente dos atributos sentimentais para o uso de predicados biológicos como definidores da personalidade. A despseudologização dos indivíduos provocou o enfraquecimento da matriz identitária apoiada na educação sentimental e a biologização da pessoa moderna (Russo & Possiano 2003) com valorização de processos identificatórios apoiados na materialidade corpórea. A figura antropológica do *sujeito sentimental* descrito por atributos psicológicos deu lugar ao *sujeito cerebral* descrito por atributos neurobiológicos (Vidal & Ortega 2004)

<sup>10</sup> A tese não privilegia a análise sociológica e histórica da interioridade nas sociedades individualistas e não aborda o surgimento da figura antropológica do *indivíduo moderno*. A análise sociológica das sociedades modernas (Dumont 1966, Ehrenberg 2004) mostra que o *indivíduo*, o agente empírico singular, é não apenas a matéria primeira do grupamento, mas principalmente a fonte de *valor supremo* dessas sociedades, em contraposição às sociedades hierarquizadas nas quais o valor supremo é atribuído a signos coletivos e ao pertencimento à coletividade. A noção de *individualismo* será utilizada no sentido específico de tomar a entidade “indivíduo” como paradigma explicativo do que determina os estados mentais da pessoa. A análise da relação entre o ‘indivíduo moderno’ e os paradigmas individualistas da pesquisa pode ser encontrado em Ehrenberg (2004).

contextos sociais, interpessoais e institucionais. Contudo, adotar uma tese externalista não é necessariamente sinônimo de abandono da dimensão da *interioridade* do indivíduo. Mesmo num relato externalista do mundo, a experiência psicológica introspectiva guarda relevância explicativa na compreensão do que é a pessoa. Nessa tese, adotaremos uma *perspectiva externalista* na investigação dos processos que constituem o que chamamos de *interioridade*.

Depois das contundentes críticas à psicologia introspectiva (Bezerra 1996, Bouveresse 1976, Costa 1994, Ryle 1984), indagar-se sobre as *origens da interioridade* parece reavivar metáforas mortas desprovidas de relevância descritiva no contexto atual e cuja ilogicidade já foi exaustivamente mostrada pela filosofia da linguagem na crítica ao *mito da interioridade*.

O *mito da interioridade* (Bouveresse 1976) afirma a existência de operações internas ao sujeito que seriam fundamento da linguagem e do conhecimento sobre a experiência individual. Isso levou à criação de conceitos psicológicos que descreveriam os estados internos do sujeito como processos inacessíveis de ‘fora’, dos quais o sujeito seria o único a ter conhecimento e que teriam somente uma relação causal contingente com o exterior. O que interessa na crítica filosófica do *mito da interioridade* explicitado por Wittgenstein (Descombes 2004) é menos mostrar o emprego linguageiro das noções de interior, mas, sobretudo, mostrar o emprego metafórico transferido da noção que consiste em identificar os atos mentais de alguém a partir de sua expressão *discursiva*. A transferência metafórica é praticada quando transferimos a descrição mental para a descrição lingüística identificando os pensamentos pela proposição lingüística que seria necessária formar para comunicá-los. A mitologia surge da criação de uma ‘linguagem do pensamento’, um ‘discurso interno’ que se localizaria ‘dentro’ do espírito ou ‘dentro’ do cérebro.

Wittgenstein foi o primeiro a argumentar contra a existência de uma ‘linguagem privada’. A idéia de que existe uma ‘linguagem privada’ está baseada na idéia de que as operações básicas da linguagem seriam operações mentais cujo resultado consistiria na expressão verbal (externa) da vontade e das afecções anímicas (interna), ou seja, a idéia de que as palavras estariam referidas a algo interno, como as sensações. Como mostra Wittgenstein, não existe uma linguagem que não possa ser intersubjetivamente observável ou partilhável, que só se refira às sensações exclusivas de seu único usuário, e cujas expressões refiram-se apenas aos processos internos desse usuário (Faustino 1995).

Essa dicotomia interno-externo dominante na filosofia de inspiração cartesiana deu origem à noção solipsista de mente como o espaço interior habitado por entidades mentais privadas que se opõem ao mundo exterior. O ‘espaço interior’ é um recipiente para os conteúdos que representam e espelham a realidade externa. A *mente interior* é desprovida de caracteres extensos e guarda a noção de imaterialidade da alma e dos processos mentais. A dicotomia interno-externo está apoiada na impressão natural de que, por trás do nosso comportamento, haveria uma distinção entre o mundo físico, contendo matéria e objetos tangíveis, inclusive corpos humanos, e o mundo privado dos fenômenos mentais. Haveria um contraste entre o mundo dos objetos materiais e o mundo das idéias etéreas que constituiriam propriedades privadas do indivíduo.

Como reação a esse dualismo e ao introspectivismo, surgiu o behaviorismo, cuja expressão mais importante é a crítica de Gilbert Ryle (1984) à noção de *mente interior*. A crítica de Ryle se centra na noção de interioridade como equivalente a eventos privados, conhecidos de modo imediato pela introspecção. A solução behaviorista foi recusar a existência dos fenômenos mentais reduzindo-os às disposições comportamentais. Nessa acepção, as proposições psicológicas de primeira pessoa são *exteriorizações*, expressões da esfera mental. As exteriorizações não são *acerca do* comportamento, são *iguais* ao comportamento. A exteriorização de uma sensação, por exemplo, é um comportamento lingüístico entendido como continuação de um comportamento pré-lingüístico que se tornará um jogo de linguagem quando compreendida por outra pessoa. O aprendizado da gramática da exteriorização das sensações é o aprendizado de um modo de manifestação de uma prática social.

Um dos problemas dessa posição foi criar um problema oposto, como um *mito da exterioridade*, no qual o mental é reduzido a tudo que é externo e exteriorizado.

Não obstante, expandir a mente para além da fronteira do que caracteriza o interior psíquico do indivíduo não significa necessariamente o abandono das metáforas espaciais na caracterização do mental. O dualismo interno-externo, embora duramente criticado por Wittgenstein e seus seguidores, tem sido revisto e reatualizado e continua presente na descrição da mente e do comportamento do indivíduo.

Um dos lugares em que o vocabulário interno-externo persiste é na explicação do comportamento intencional. A discriminação entre interno e externo ao organismo parece

necessária para explicar a ação intencional humana, ou mesmo o comportamento de qualquer sistema. O comportamento é um movimento endógeno produzido internamente, um movimento que tem sua origem causal dentro do sistema cujas partes estão em movimento. Nesse sentido, interno não significa simplesmente dentro do corpo, sob a pele, interno inclui a idéia de uma parte própria ou integral do organismo exibindo o comportamento. No entanto, o fato de afirmar que um comportamento é interno ao organismo não significa dizer que é produzido no interior do mesmo, ou independe de suas interações exteriores, mas tão somente que está associado aos movimentos endógenos do organismo concebido como um sistema em interação na direção do mundo. Considerando os conteúdos mentais do agente como aquilo que explica a sua ação intencional, pode-se afirmar que o mental está sempre delimitado ao agente situado. Podemos identificar o mental apenas em pessoas localizadas espacialmente. Não existe o mental fora da existência espacial corporal. O agente situado existe do lado de fora, nas exteriorizações da mente, na história e na linguagem, em ações e falas. Contudo, o comportamento como exteriorização do agente para ser compreendido como um ato intencional do agente precisa ser considerado como sendo gerado pelo próprio agente, por sua rede de crenças e desejos.

Outro lugar em que a utilização dos termos interno e externo parece persistir é na idéia de uma experiência introspectiva interna. É inegável a experiência existencial de habitarmos um corpo com vida interior. É inegável o fato de que podemos ter pensamentos introspectivos que não comunicamos aos outros. A palavra introspecção vem do latim “introspecto”, que significa ‘ação de olhar para dentro’, ‘penetrar’. Essa ação captura a forma como tradicionalmente concebemos nossa apreensão dos estados mentais (ver Bezerra 1996). Esses termos expressam, numa linguagem espacial, uma divisão entre o mundo interno e o mundo externo. Contudo, a conotação espacial desta linguagem é puramente metafórica. Dizer que um estado ou uma entidade são internos à mente não é dizer que isso implica uma dada fronteira espacial. Se a distinção entre ‘o que está dentro da mente’ e ‘o que repousa fora da mente’ fosse apenas espacial, a diferença entre a introspecção e a exterocepção seria trivial, uma mera questão de olhar em diferentes direções.

Parece haver uma aparente contradição entre a *exteriorização do agente* e a *experiência introspectiva*. Contudo, a inclusão da exteriorização do agente como critério do mental não torna obsoleta a noção de *interioridade*. Ao contrário, para entendermos a

exterioridade humana, ou seja, a ação intencional do agente, precisamos redefinir o que é a *interioridade*. Se a mente não é mais caracterizada por aquilo que permanece no seu interior, se não é mais concebida como um recipiente de representações, fantasias ou idéias, o que caracterizaria a experiência da vivência interior descrita pelo relato da primeira pessoa ou pela experiência introspectiva?

Em “*O Self e A Interioridade*” (inédito), Costa mostra que esta idéia corrente espacial do ‘self interior’ ou de uma experiência interior pode ser analisada de duas formas. Uma se refere ao *sentido* do termo e outra à sua *função* na vida mental. Em relação ao *significado* do termo, afirma Costa, “utilizamos os vocábulos de forma *teoricamente infundada, mas psicologicamente justificada*.” Isso quer dizer que nos equivocamos no sentido dos termos, mas temos razões psicológicas que justificam seu emprego. Costa sugere que “consideremos os *predicados mentais* denominados de ‘interiores’ como a descrição, feita de outra perspectiva, de atividades “exteriores” dos organismos humanos na interação com o meio. Interior e exterior são duas formas fenomênicas da mesma relação concreta e material do self com seu mundo” (Costa inédito). Ele sugere que empreguemos o termo *interioridade* como “um fenômeno público relacional, fundado na corporeidade física e sobre o qual podemos fazer afirmações empíricas testáveis”. Nessa acepção, o *interior* é aquilo que é “formado pela auto-percepção corporal, pelas representações dos objetos e situações ambientais e pelas narrativas em primeira pessoa sobre nós mesmos”; e o *exterior* é formado “pelos objetos e eventos cujas existências ou realidades independem de percepções particulares de sujeitos particulares” (Costa inédito).

E qual é a *função* preenchida pela experiência interior na vida mental? Como afirma Costa (inédito), a *interiorização* dos predicados mentais é a condição necessária para a *estabilização do sentido do self*, pois o sentimento de continuidade da existência depende da internalização das expressões externas do self sobre o meio e do fluxo de trocas entre o mundo interno e o externo. O self, continua Costa:

“é o efeito do esforço constante dos organismos humanos para unificar o dessemelhante, predizer o imprevisível, reproduzir o que nos satisfaz e explorar o que não conhecemos. Esse conjunto de atividades físico-mentais é o que chamamos de criatividade e que depende da vida interior para se exprimir” (Costa inédito).

Portanto, para expressão de nossa ação criativa no mundo é necessária a experiência de ser uma existência contínua. *A função do self e da experiência da interioridade está na criação de um sentimento de continuidade do ser de onde possa emergir a ação criativa.*

Portanto, uma das funções da interioridade é estabilizar um sentido de continuidade da identidade que favoreça a integração entre a exteriorização do agente pela ação criativa e a experiência introspectiva. Uma das características que identificam uma pessoa essencialmente como a pessoa que ela é depende da internalização das ações externas do agente que estão na base do sentimento de ser um self que persiste no tempo. A *continuidade da consciência reflexiva* estabelecida pela memória do passado de nossas ações e a projeção de expectativas futuras, ou seja, aquilo que caracteriza nossa vida mental, é um dos candidatos psicológicos para esse estatuto de identificação.

Nessa acepção, a *interioridade* é caracterizada pelo sentimento de continuidade da experiência subjetiva da pessoa. Há uma experiência que é ter consciência de nossos vários estados mentais, ou seja, uma fenomenologia de nossa vida mental que é tida e experimentada. Um dos aspectos dessa vida interior é o perspectivismo. Ou seja, o fato de experimentarmos o mundo de uma dada perspectiva. Um ser autoconsciente é aquele que não somente é sujeito de suas experiências fenomenais, mas possui a capacidade de conceber a si mesmo como uma perspectiva de primeira pessoa. A autoconsciência e o autoconhecimento são essenciais para a constituição da perspectiva da primeira pessoa. Como diz Lynne Baker (1998), “um ser consciente torna-se autoconsciente ao adquirir uma perspectiva da primeira pessoa – uma perspectiva a partir da qual pensa-se sobre si mesmo como um indivíduo frente ao mundo, como um sujeito distinto de tudo o mais”. A atividade introspectiva, o autoconhecimento e a autoconsciência são aspectos centrais de nossa vida mental interior e fazem parte daquilo que nos define como pessoas. A forma como nos constituímos como pessoa é diretamente determinada pelo modo como nos auto-percebemos, auto-representamos e pelas narrativas que construímos.

Contudo, não é a atividade introspectiva que está no ponto de origem da interioridade. Como veremos, a experiência de ser um self que persiste no tempo está associada à experiência prévia de habitar um corpo em interação com o ambiente. Nossa interioridade emerge de um complexo espaço de interação com pessoas e um mundo de objetos físicos e

culturais compartilhado. *As origens de nossa interioridade são, portanto, externas e relacionais.*

O objetivo dessa tese é investigar uma forma de conceber o universo psicológico da pessoa, seu *mundo interior*, que preserve o relato introspectivo da experiência humana em primeira pessoa e que leve em consideração a gênese externalizada dos processos mentais. O problema que se coloca é como compatibilizar a noção introspectiva do eu com as teses externalistas do mental. Dito de outro modo, como explicar a constituição do eu introspectivo capaz de ter estados mentais privados e acesso privilegiado a seus conteúdos mentais, cujos estados mentais são, ao mesmo tempo, determinados por suas relações sociais ambientais exteriores?

A tese está dividida em 3 capítulos.

No primeiro capítulo, analiso a experiência subjetiva de ser um ser autoconsciente. Começo por fazer a distinção entre *ser* um ponto de vista subjetivo e *se conceber e se conceitualizar* como um ponto de vista subjetivo. Analiso a perspectiva da primeira pessoa e os processos que a caracterizam, destacando dois eixos de investigação. Em primeiro lugar, investigo a emergência da perspectiva da primeira pessoa a partir da história do desenvolvimento ontogenético. Ter uma perspectiva de primeira pessoa é ser capaz de ter consciência de si e de seus estados mentais e corporais. A capacidade de compreender a experiência e de se conceber e se representar como uma perspectiva deriva da aquisição do desenvolvimento da capacidade de conceitualizar. Num segundo momento, analiso os processos mentais que caracterizam um ser autoconsciente: a autoconsciência e o autoconhecimento. Pretendo mostrar que a emergência da capacidade de autoconsciência e de autoconhecimento estão associadas à aquisição de conceitos, em particular, dos conceitos que permitem fazer referência ao eu e a seus estados mentais e corporais. Por fim, analiso quais são os obstáculos que uma abordagem externalista do mental pode trazer ao problema do autoconhecimento. Se o externalismo defende a tese de que os conteúdos mentais são em parte determinados pelos contextos interativos do indivíduo, ele precisa explicar como o conteúdo da experiência fenomenal é determinado pelo exterior.

O segundo capítulo é dedicado à investigação das teses externalistas e a sua compatibilização com o autoconhecimento. O que interessa, para os propósitos de investigação do *autoconhecimento e da gênese do eu*, é que a defesa da tese taxonômica

individualista ou externalista dos conteúdos mentais acarreta modificações consideráveis na forma como concebemos o mental e os processos introspectivos. Apresento a concepção individualista da mente e discuto os problemas oriundos das teses individualistas cartesianas que deram origem às concepções solipsistas do mental. O individualismo estuda os fenômenos mentais a partir do que acontece com a pessoa e como ela responde ao seu ambiente físico sem referência ao que acontece no contexto social no qual está situada. O individualismo postula a autoridade do conhecimento da primeira pessoa, de seus próprios estados mentais, como um conhecimento *a priori*, que para obtê-lo não precisaríamos investigar o ambiente ou conhecer o mundo externo.

Em seguida, traço a história conceitual endógena das teses externalistas e de seus modelos, restringindo-me às implicações dessas teses no que concerne ao campo da filosofia da mente, em particular à questão do autoconhecimento. Na perspectiva externalista, acredita-se que o pensamento é identificado pela relação entre o indivíduo e alguma coisa fora da sua cabeça. Se o pensamento não está completamente na cabeça, não pode ser apreendido pela mente no modo que é requerido pela autoridade da primeira pessoa. Isso parece significar, em alguma medida, uma renúncia ao autoconhecimento como um conhecimento *a priori* de nossos estados mentais. Apresento os diferentes modelos de externalismo e adoto o externalismo perceptivo Davidson baseado na origem intersubjetiva da mente. Ao definir conhecimento como, de início, intersubjetivo, Davidson inverte a lógica solipsista cartesiana e empirista que pretende construir uma visão do mundo a partir do indivíduo e seu ponto de vista subjetivo perceptivo. É a partir da interação triangular eu, outro e mundo e das trocas comunicativas que desenvolvemos um conhecimento compartilhado do mundo externo, o conhecimento do que o outro pensa e o autoconhecimento, ou seja, o conhecimento do que ocorre dentro de nós mesmos.

No terceiro capítulo, apresento os modelos externalistas da gênese do eu que enfatizam a interação com o ambiente. Esses modelos podem ser divididos em dois blocos: os modelos ecológicos e os modelos triádicos. A abordagem ecológica opera com uma lógica dual de constituição do eu, na qual o sentido do eu emerge da interação eu-mundo ou organismo-ambiente. Ou seja, o fator invariável da identidade é a relação organismo-ambiente. A abordagem triádica opera com uma lógica ternária, na qual o sentido do eu emerge da relação entre o eu, o intérprete e o mundo. O processo de desenvolvimento do eu em Winnicott é

apresentado como o modelo psicológico de emergência do sentido de continuidade do eu que exemplifica a conceitualização epistemológica triádica de Davidson. Para Winnicott, o fator invariável da identidade emerge da relação do eu com o outro que interpreta o mundo e apresenta os objetos. A concepção de *espaço transicional*, conceito pilar da metapsicologia de Winnicott, espaço para o qual convergem o eu, o outro e o objeto transicional, *é o espaço do qual se origina nossa interioridade*.

# CAPÍTULO I

## A PERSPECTIVA DA PRIMEIRA PESSOA

### 1. O Que É Sentir-Se Como Nós

Num dos episódios da série ‘Star Trek, The Next Generation’, Data, um dos personagens desse universo ficcional, é julgado por seu estatuto de ser ou não considerado uma pessoa. Data é uma forma de vida artificial, projetada para assemelhar-se a um humano. Ele é possuidor de um cérebro composto de positrônios que funciona como um computador central e exibe uma forma de consciência reconhecida por humanos. Esta é uma tecnologia ficcional dos cérebros dos robôs idealizada por Isaac Asimov, um dos pais da *science fiction*.

Em ‘The Measure of The Man’, vemos o Capitão Picard defender Data contra a ameaça de ser desmontado. A questão central é o direito que teriam os humanos que projetaram Data, seus criadores, de intervir sobre sua ‘vida artificial’ e, conseqüentemente, de destruí-lo. A defesa em favor de desmontar Data se baseia no argumento de que ele não é uma pessoa com direitos, mas sim uma máquina de propriedade da União da Federação dos Planetas. Em defesa de Data, Capitão Picard alega que ele preenche todos os critérios dos seres conscientes, como a capacidade de sentir, perceber e agir intencionalmente, e que, como tal, teria o direito à liberdade e à autodeterminação concedidos às pessoas. Picard alega que todos os seres são ‘criados’, mas que isso não os torna possessão de seus criadores. Data, portanto, teria o direito de tomar suas próprias decisões sobre sua vida.

Esse estimulante exercício ficcional traz à baila a questão sobre os critérios para definição do que é a personalidade. Daniel Dennett inicia o artigo “Conditions of personhood” (1981a) com uma definição que interroga nosso senso comum sobre o que é ser uma pessoa. Ele afirma que eu e você, leitor, certamente somos pessoas, mas não há garantias de que somos feitos da mesma composição biológica. Portanto, se o fato de pertencermos à espécie humana, se nossa condição biológica, não é uma condição necessária nem suficiente para caracterizar a pessoa, quais seriam, então, os critérios necessários? Dennett apresenta seis condições necessárias para a personalidade: ser racional; ser um ser a quem se atribui estados psicológicos e predicados intencionais; ser objeto das atitudes de pessoa (*personal stance*), ou

seja, ser tratado como pessoa; ser capaz de tratar reciprocamente os outros como pessoas; ser capaz de comunicação; exibir um tipo especial de consciência, a autoconsciência.

As três primeiras condições – racionalidade, intencionalidade, adoção de atitudes de pessoa – são condições necessárias, mas não suficientes para ser uma pessoa. Os seres que exibem essas três primeiras condições são *sistemas intencionais*<sup>11</sup> que se caracterizam por apresentar comportamentos que podem ser explicados e previstos por meio de atribuições de crenças e desejos. O quarto critério, *reciprocidade*, caracteriza os sistemas intencionais capazes de adotar uma atitude intencional frente ao mundo e às outras pessoas. São *sistemas intencionais de segunda-ordem* capazes de ter crenças e desejos acerca de suas próprias crenças e desejos, ou seja, são capazes de representar crenças, desejos e intenções de segunda ordem. Essas características são exibidas por alguns animais<sup>12</sup>. A quinta característica, ser capaz de comunicação verbal, ou seja, ser dotado de linguagem, exclui os animais não humanos da noção de pessoalidade. A última condição, a autoconsciência, é a idéia de que as pessoas são distinguíveis de outras entidades, pois exibem um tipo especial de consciência. A autoconsciência é o que nos permite ter consciência das nossas ações, ou melhor, ter consciência de que somos nós os responsáveis por nossas ações e que podemos ser indagados acerca das razões pelas quais as executamos e ser responsabilizados moralmente por nossos atos.

Dennett conclui o artigo ressaltando a normatividade inerente à noção de pessoa. Mesmo na presença das seis condições, persistirá sempre o caráter ideal e arbitrário do conceito de pessoa que depende da interação entre pessoas (*personnal interaction*). Qualquer definição metafísica que estabeleça condições ontológicas acerca da pessoalidade não escapará dos aspectos normativos que compõem a noção moral de pessoa. Em suas palavras:

---

<sup>11</sup> O *sistema intencional* é toda entidade cujo comportamento é previsível e explicável pela *atitude intencional*. Como afirma Dennett, "A atitude intencional é a estratégia de interpretar o comportamento de uma entidade (pessoa, animal, artefato, seja lá o que for) tratando-a como se fosse um agente racional que governa sua "escolha" da "ação" levando em "consideração" suas "crenças" e "desejos".(...) A estratégia básica da atitude intencional é tratar a entidade em questão como um agente, com a finalidade de prever - e, portanto explicar, num certo sentido - suas ações ou movimentos" (1996: 27). O sistema intencional é aquele cujo comportamento pode ser previsível pela imputação ao sistema da posse de certas informações e pela suposição de que é dirigido por certos objetivos. Para maiores detalhes ver Dennett (1981b).

<sup>12</sup> Os estudos em Psicologia Evolucionária têm mostrado que certos animais (em particular os primatas) possuem representações mentais de seu ambiente e dos comportamentos dos outros que os permitem resolver problemas e simular novas situações. Alguns chimpanzés chegam mesmo a exibir comportamentos que se assemelham à capacidade de ler o comportamento dos outros, um esboço do que chamamos 'teoria da mente'. Sobre a Psicologia Evolucionária ver Doug Jones (1999).

“se essas seis condições (estritamente interpretadas) fossem consideradas suficientes, elas não assegurariam que qualquer entidade real seria uma pessoa, pois nada satisfaria essas condições. A noção moral de pessoa e a noção metafísica de pessoa não são conceitos separáveis ou distintos, mas apenas dois pontos de repouso diferentes e instáveis do mesmo contínuo.” (1981a: 285)

Portanto, definimos o sentido de nossos critérios metafísicos a partir de nossas injunções éticas, como propõe Stephen White citado por Dennett (2004).

Contudo, é inegável que a presença dessas seis condições oferece critérios, mesmo que provisoriamente suficientes, para a compreensão da constituição do agente moral. Com certeza, a capacidade de **autoconsciência** – ou, ao menos, a possibilidade de adquiri-la, no caso dos bebês, ou o fato de já tê-la tido em algum momento, no caso dos que a perderam – tem sido um dos critérios definidores do que é uma pessoa.

Data é um ser que satisfaz essas condições. E é exatamente por isso que somos inclinados a conceder-lhe o estatuto de pessoa. Embora possamos restringir seus direitos, alegando que ele pertence a uma classe diferenciada de pessoa – como nas práticas de exclusão social em que limitamos os direitos àqueles que consideramos como não pertencendo ao nosso grupo, como não sendo “um de nós” como diria Rorty<sup>13</sup> – Data tem características que têm sido reconhecidas como definindo a pessoa. Ainda que esses critérios ontológicos estejam, em última instância, subordinados aos aspectos morais definidos pela comunidade a qual pertencemos, isso não torna obsoleto o seu uso como balizadores de nossos discursos e atitudes frente ao que consideramos uma pessoa. Dentre todos os critérios, **a capacidade de autoconsciência** tem sido tratada como um critério definidor decisivo.

A autoconsciência tem sido considerada um traço distintivo da espécie humana. Mesmo que as pesquisas com chimpanzés (Gallup 1997, *apud* Baker 1998) demonstrem que os mesmos exibem uma capacidade cognitiva rudimentar de auto-reconhecimento (como aprender a se reconhecer no espelho após um tempo de exposição ao objeto), a capacidade de autoconsciência, de fazer referência a si mesmo e identificar a si mesmo, parece específica da nossa espécie. Somos as únicas criaturas capazes de pensar sobre nós como uma existência permanente no tempo (*persisting self*), isto é, a ter uma perspectiva da continuidade do eu. A

---

<sup>13</sup> Em “Droits de l’homme, rationalité et sentimentalité” (1994), Richard Rorty mostra que as atrocidades impostas a outros grupos, como os negros, judeus, loucos e mulheres, ocorrem quando não consideramos estes outros como “um de nós”. Rorty propõe que tentemos alargar tanto quanto possamos o sentido de “nós” a pessoas que pensávamos como sendo “eles”, de forma a garantir os direitos humanos e estendê-los a todos os *bípedes sem pluma*, ampliando *nossa* comunidade moral. (Claudia Passos Ferreira 1998)

autoconsciência é um aspecto essencial na construção de nossa identidade pessoal ao longo de nossa história de vida, pois nos permite escapar do aqui e agora de nossa ‘presença ao mundo’ e nos projetarmos numa dimensão futura como agentes morais responsáveis por nossas ações e decisões. Pode-se afirmar, portanto, que a autoconsciência é essencial para ser uma pessoa.

Mas, como definimos um ser autoconsciente?

Todas as formas de autoconsciência pressupõem um ponto de vista subjetivo. Em “What is it like to be a bat” (1992), Thomas Nagel mostra como um ser consciente se torna autoconsciente ao adquirir uma perspectiva de primeira pessoa. A perspectiva da primeira pessoa é a experiência subjetiva por meio da qual podemos nos conceber como um ser individual, distinto de todo o resto do mundo com o qual estamos lidando. Como mostra Nagel, todo ser consciente, mesmo seres sencientes não humanos, é sujeito de suas experiências no sentido de que as experimenta. Todo fenômeno subjetivo está conectado, portanto, como um ponto de vista singular.

No entanto, nem todo ser consciente é autoconsciente. A autoconsciência é um estado que nos torna consciente *de* nós mesmos. É a consciência de ter uma experiência subjetiva e possuir conceitos de si mesmos na perspectiva da primeira pessoa. Somente os animais humanos são capazes de exibir essa forma conceptual de perspectiva da primeira pessoa e podem ser ditos completamente autoconscientes, ou, como denomina Lynne Baker (1998), capazes de exibir uma versão *forte* da perspectiva da primeira pessoa.

Baker propõe uma distinção entre a perspectiva forte e a perspectiva fraca da primeira pessoa que parece útil para esclarecer a relação entre a autoconsciência e a identidade da pessoa. A perspectiva *fraca* é aquela exibida por criaturas capazes de resolver problemas e cujos comportamentos podem ser explicados em termos intencionais. São seres como os animais<sup>14</sup> e os recém-nascidos aos quais podemos atribuir crenças e desejos, pois parecem se comportar de uma certa perspectiva. Não é necessário que um organismo possua o conceito de primeira pessoa para que possam lhe atribuir um comportamento motivado por crenças e desejos. Segundo Baker, os fenômenos *fracos* da perspectiva da primeira pessoa são exibidos por seres sencientes, capazes de solucionar problemas por meio de atitudes perspectivais que explicam seu comportamento intencional. Como todos os estados psicológicos são

---

<sup>14</sup> Há um grande número de pesquisas em Psicologia Evolutiva que tenta mostrar a relação entre nossa forma de se desenvolver psicologicamente e a evolução de outras espécies em particular os macacos, chimpanzés, cães e corvos.

perspectivais na origem, assim também a crença, o desejo e o comportamento pertencem a um único organismo.

O que é característico da perspectiva *forte* da primeira pessoa é a capacidade de **reconhecer a si mesmo como um ponto de vista subjetivo** e de pensar sobre si mesmo como uma pessoa. Essa capacidade única dos humanos – de se conceber como ‘eu’ autoconsciente e capaz de introspecção – é fundamental para o desenvolvimento do que consideramos uma pessoa, *what-is-it-like-to-be-us*.

Nesse capítulo, analisarei a perspectiva da primeira pessoa e os processos que a caracterizam, destacando dois eixos de investigação: (1) como emerge a perspectiva da primeira pessoa na história do desenvolvimento ontogenético a partir da aquisição da capacidade de meta-representar e conceitualizar; (2) análise dos processos mentais que caracterizam um ser autoconsciente: autoconsciência e autoconhecimento. Pretendo mostrar que a emergência da capacidade de autoconsciência e de autoconhecimento estão associadas à aquisição de conceitos, em particular os conceitos que permitem fazer referência ao eu e a seus estados mentais e corporais. Por fim, analisarei quais são os obstáculos que uma abordagem externalista do mental pode trazer ao problema do autoconhecimento, ou seja, como compatibilizar a idéia de que os conteúdos mentais são parcialmente determinados pelo ambiente externo com a noção de que temos conhecimento privilegiado de nossos estados mentais e não nos equivocamos na atribuição de estados ao eu.

## **2. Tornar-se Uma Perspectiva**

### **2.1 O Caráter Subjetivo Consciência**

O debate acerca da experiência subjetiva da consciência tem sido dividido entre: os eliminativistas que renunciam ao caráter qualitativo da experiência, os *new mysterians* que tentam preservar o caráter de singularidade da experiência subjetiva e negam qualquer redução da consciência a processos neurofisiológicos, e aqueles que tentam oferecer um relato naturalista da consciência. O problema da consciência, *the hard philosophical problem* como chama Chalmers (1995), tem sido abordado a partir do problema da identidade entre o físico e o mental. Como mostra Nagel (1992), haveria um *gap* entre a consciência fenomenal e as explicações fisicalistas, na descrição fisicalista da experiência fenomenológica. Ele argumenta que uma teoria que identifica o mental ao físico reduziria o fenômeno da consciência a

processos neuroquímicos que eliminariam a qualidade da experiência subjetiva, o *qualia*. Nagel (1992) junto com Searle (2004), Chalmers (1995) e Varela (1991, 1996a, 1999, 2004) fazem parte do grupo dos *new mysterians*, que, apesar de acreditarem que a consciência não pode existir sem o suporte físico do cérebro, acreditam que a consciência e o caráter subjetivo da mente possuem uma dimensão privada da perspectiva da primeira pessoa que não podem ser inteiramente compreendidos ou naturalizados<sup>15</sup>. Isto é, a descrição fisicalista da consciência pode chegar a prever, do ponto de vista da terceira pessoa, as correlações entre as experiências subjetivas e as condições objetivas do cérebro, mas não pode explicá-las. Dennett, Rorty (1988), Andy Clark (1996), os Churchlands (Andrieu 1998) e Dretske (1997), fazem parte do grupo que defende a tese materialista de que nenhum fato consciente pode ser acessível somente a uma pessoa, a um único ponto de vista. Dennett (1991, 2001, 2003) propõe que a consciência seja investigada a partir de métodos científicos que esclareçam a

---

<sup>15</sup> Em comunicação oral, Francisco Ortega e Octávio Serpa chamaram à atenção para o equívoco e a conseqüente simplificação de identificar a posição enactiva da consciência de Francisco Varela com a posição dos *new mysterians*. A etiqueta ‘new mysterians’ foi criada por Owen Flanagan (1991), referindo-se aos filósofos que acreditam que o fenômeno da consciência não poderia ser reduzido ao seu funcionamento cerebral, pois haveria um ‘mistério’ na consciência que a tornaria irreduzível à materialidade cerebral, e que, portanto, a metodologia científica do discurso da terceira pessoa seria insuficiente para explicar o fenômeno da consciência. Dentre os que defendem essa posição, estão, por exemplo, Thomas Nagel, Roger Penrose, Colin McGinn, John Searle e Noam Chomsky. No quadro de divisão que propus, a neurofenomenologia de Varela estaria mais próxima da posição defendida pelos *new mysterians* do que da posição dos eliminativistas e dos funcionalistas. Como toda tentativa de sistematização, essa divisão além de simplificar o problema, com certeza, não dá conta dos matizes variados de cada ponto de vista filosófico que, muitas vezes, se distinguem em detalhes nem sempre de relevância teórica. Nesse sentido, nem mesmo David Chalmers poderia estar no grupo dos *new mysterians*, pois há matizes de sua posição que não se adequam ao que defendem os proponentes do Novo Misterianismo. A título de exemplo, em outras sistematizações, D. Chalmers junto com T. Nagel são considerados como adeptos de um certo pampsiquismo que defenderia uma ‘consciência universal’, e se afastam de John Searle e Colin McGinn que são detratores do pampsiquismo. A solução, então, a meu ver, pode ser adotar o esquema proposto pelo próprio Varela em “Neurophenomenology: a methodological remedy to the hard problem” (1996a). Ele propõe uma divisão das abordagens naturalistas da consciência em quatro eixos: 1) o grupo de direita, que resolve o *hard-problem* eliminando o fenômeno da consciência, representado pelos Churchland e F. Crick e Ch. Koch; 2) o grupo de centro direita, que defende a abordagem funcionalista da consciência, reduzindo a consciência a estados mentais funcionais ou intencionais, como D. Dennett, R. Jackendorff e G. Edelman; 3) o grupo de centro esquerda, onde se localizam os *new mysterians* como Nagel e McGinn; 4) e o grupo de esquerda, representados por F. Varela, D. Chalmers, J. Searle, que reconhecem o caráter irreduzível da experiência consciente da primeira pessoa mas que se recusam a conceder uma resposta dualista que endossaria o fosso entre a experiência e a estrutura cognitiva. Varela reconhece que sua posição é próxima da defendida pelos novos misterianos como Nagel e Searle. Ele afirma concordar com o diagnóstico que fazem do problema, ou seja, que a experiência em primeira pessoa da consciência é irreduzível, mas não concorda com a impossibilidade de solução do *gap* entre o fenômeno da consciência e os mecanismos cognitivos que deriva da posição desses autores. Varela propõe como solução a abordagem enactiva ou da cognição situada, na qual o fenômeno da consciência é estudado pela ‘neurofenomenologia’, que conjugaria o fenômeno da experiência de primeira pessoa com as estruturas cognitivas descritas pelas ciências cognitivas. Segundo o autor, o fenômeno da consciência deve ser estudado levando em consideração a experiência do corpo vivo situado no ambiente, e buscando conexões entre a mente biológica e a mente experiencial.

engenharia da mente e descrevam a consciência do ponto de vista da terceira pessoa, o que chama de heterofenomenologia. Para Dennett, a abordagem heterofenomenológica permitiria criar uma organização e catalogação dos dados primários da consciência, viabilizando o desenvolvimento de uma teoria explanatória.

Existem diversas tentativas teóricas de naturalizar a consciência e desmistificar seu caráter ‘misterioso’ que mostram como os aspectos relevantes da consciência podem existir numa concepção puramente fisicalista do mundo (ver U. Kriegel 2005). O projeto de naturalização da mente visa explicar os fatos mentais a partir das ciências naturais, descrevendo estados e processos mentais como aspectos do cérebro e do sistema nervoso central. Dentre os defensores da naturalização da consciência, estão os representacionistas. Examinarei a abordagem representacionista de Fred Dretske (1997) que tenta preservar o caráter qualitativo da experiência subjetiva. Dretske propõe compatibilizar a idéia de que os aspectos qualitativos da experiência são subjetivos e privados com a idéia de que são passíveis de serem determinados objetivamente, ou seja, podem ser capturados por um relato representacional.

Em sua Teoria Representacional da Consciência<sup>16</sup>, o estado de estar consciente é um modo de representar alguma coisa. A consciência fenomenal nos torna consciente das propriedades fenomenais dos objetos e é dada pelos sentidos perceptivos: ver, ouvir, cheirar, provar, sentir. Os estados mentais conscientes são estados que nos tornam conscientes de objetos, propriedades e fatos. São estados que nos tornam conscientes *com* eles e não estados *dos* quais estamos conscientes. O que torna a pessoa consciente de estar tendo a experiência é

---

<sup>16</sup> Gostaria de fazer uma distinção entre três usos da noção de representação: a ‘atividade da representação’, a ‘idéia de representação’, e ‘representação mental’. A ‘atividade da representação’ está associada a alguma atividade mental na qual nos engajamos. A ‘idéia de representação’ geralmente é usada na Teoria Representacional da Percepção como uma interface semântica interna entre nós mesmos e o que estamos pensando ou percebendo, e está associada à idéia de colocar ‘um objeto diante da mente’. Essa noção de representação interna é criticada no Capítulo II, na discussão sobre a noção individualista da mente. O terceiro uso encontramos mais recentemente na literatura após a virada informacional (*the information turn*) e, de certa forma, deriva da idéia precedente da representação como um objeto mental com propriedades semânticas. A noção de ‘representação mental’ é um construto teórico das ciências cognitivas segundo o qual os estados e processos cognitivos são constituídos pela ocorrência, transformação e estocagem (no cérebro ou na mente) de estruturas representacionais que captam informações. Essa noção é o ponto de partida da Teoria Representacional ou Computacional da Mente que encontra em Dretske um dos seus mais promissores defensores. O ponto de partida dessa abordagem é que os estados mentais (pensamentos, crenças, desejos e percepções) estão relacionados a representações mentais e a intencionalidade desses estados deve ser explicada em termos das propriedades semânticas da representação mental. (Para uma melhor discussão ver Pitt, David, ‘Mental Representation’, Stanford Encyclopedia of Philosophy. Electronic Dictionary)

um estado interno que representa as propriedades da experiência. Sem o estado representacional não é possível ter consciência da propriedade que experimentamos.

A experiência consciente, em termos do *'what it is like for the subject'*, pode ser distinguida em dois aspectos: (1) a consciência das propriedades fenomenais dos objetos, que é o caráter qualitativo da experiência que os fenomenólogos chamam de 'presença ao mundo'; (2) o aspecto *'for-me'* (*for-me-ness*, como chama Kriegel, 2005), que é o caráter subjetivo da experiência, ou seja, o que é ter a experiência *para o sujeito*. O caráter subjetivo da experiência consiste num conteúdo representacional particular que se desdobra em uma representação de algum aspecto do ambiente e numa representação interna da representação.

Para um materialista como Dretske, não há fatos mentais que só sejam acessíveis a uma única pessoa. Toda subjetividade faz parte do mundo objetivo. Segundo Dretske, "pode haver circunstâncias de tempo e espaço que capacitem uma pessoa a conhecer alguma coisa que os outros não conhecem, circunstâncias de tempo e espaço que tornem uma pessoa autoritativa, mas não há fatos privilegiados a uma pessoa, fatos que só uma pessoa *pode* conhecer" (1997: 65). Temos informação direta e imediata sobre nossa vida mental, mas não temos acesso privilegiado a informações que os outros não possam ter. Todos podem ter acesso à mesma informação. Isso quer dizer que a vida subjetiva de uma criatura, o que é se sentir como essa criatura (*what it is like to be*), não é inacessível. Contudo, há uma particularidade na experiência de ser um ponto de vista.

Ao contrário de certos relatos naturalistas que explicam o mental a partir do funcionamento da maquinaria cerebral biológica, Dretske concebe a mente como a função de manipulação de representações, a partir da idéia de 'informação'<sup>17</sup>. Ele defende duas teses: (1) todos os fatos mentais são fatos representacionais; (2) todos os fatos representacionais são fatos sobre funções informacionais. A idéia central é a mente é um design projetado pela história evolutiva com a função de trazer a informação sobre o objeto. A mente é um sistema que representa uma dada propriedade e tem a função de indicar e representar a propriedade. A

---

<sup>17</sup> Em seus trabalhos mais recentes, como mostra Pereira (2006), Dretske reformula sua Teoria Semântica Informacional, e propõe uma abordagem representacionista que combina a noção de informação com o elemento teleológico (Teleosemântica). O importante para o argumento aqui desenvolvido é que essa reformulação resulta, como esclarece Pereira (2006), na idéia de que "a principal distinção entre as representações sensoriais e conceituais é feita em termos da distinção entre dois tipos básicos de indicadores de função, sistêmica e adquirida. A representação de que *O é P* é sensorial e não-conceitual quando seu indicador de função é derivado do sistema do qual ele é um estado, e conceitual quando seu indicador de função é adquirido do tipo do qual ele é um estado particular. No primeiro caso, o indicador de função é filogeneticamente determinado pela seleção natural, enquanto no segundo caso ele é adquirido ontogeneticamente pelo aprendizado." (Pereira 2006)

função pode ser sistêmica ou adquirida. As funções sistêmicas são funções filogenéticas dos sistemas sensoriais produzidas pelo processo histórico de seleção natural. As funções adquiridas são funções aprendidas no processo de maturação do indivíduo. Os sistemas perceptivos (incluindo a propriocepção) produzem representações e têm a função biológica de fornecer informação. A tarefa do aparelho sensorial é construir representações mentais sobre o mundo. Todos os estados mentais (crenças, pensamentos e experiências) são representações naturais. Toda representação é representação de um fato, mas nem toda representação é conceitual. A experiência é uma forma de representação não-conceitual e é inerente ao sistema. As crenças e pensamentos são representações conceituais adquiridas ao longo do desenvolvimento ontogenético.

Tanto as atitudes proposicionais (crenças e desejos) quanto a experiência sensorial (o *qualia*) são representações internas. O caráter qualitativo da experiência é identificado às propriedades que os objetos representados possuem, e essas propriedades são, em princípio, passíveis de serem conhecidas por qualquer pessoa. Para Dretske, o caráter qualitativo da experiência perceptiva não deve ser definido funcionalmente, por meio de nossas disposições comportamentais, mas deve ser definido em termos físicos. Chamamos de *qualia* às propriedades fenomenais dos objetos que são experimentadas subjetivamente. Essas propriedades são aquelas que estão representadas nas informações que nossos sentidos têm a função natural de fornecer; e podem ser determinadas assim como determinamos outras funções biológicas dos órgãos corporais.

Os organismos têm funções biológicas específicas com diferentes performances discriminativas e capacidades diferenciadas de representar propriedades. As experiências qualitativas podem ser distintas mesmo que os organismos tenham a mesma capacidade biológica. Portanto, a experiência de ser um ponto de vista subjetivo não pode ser reduzida à função biológica. O ponto de vista subjetivo é determinado também pelas partes e aspectos da experiência do objeto que representamos. Todo e qualquer objeto ocupa diferentes lugares no mundo, portanto, captamos informações diferentes do mundo de acordo com as posições que ocupamos. Contudo, as experiências não diferem subjetivamente simplesmente porque são experiências de diferentes objetos em diferentes perspectivas. Essa diferença é uma experiência que pode ser partilhada ao trocarmos de lugar com o outro e ocuparmos o mesmo ponto de vista do outro. As diferenças nos estados subjetivos que interessam e caracterizam a

singularidade da experiência são as diferenças que resultam não de pontos de vistas diferentes, mas no *modo* como os pontos de vistas são vistos. O que faz diferença no *what it is like for the subject* não são os objetos acerca dos quais é a experiência, mas o *modo* como esses objetos são representados de um dado ponto de vista.

Portanto, como afirma Dretske, o que muda é o *modo* como a experiência é codificada. Ao nível da consciência fenomenal, temos a experiência de uma variedade de propriedades. Ao nível da consciência conceitual, temos experiência de um *tipo* de propriedade, pois há uma categorização da experiência sensorial que é dada pelo aprendizado dos conceitos. Os sistemas sensoriais inatos são funções herdadas filogeneticamente, e não podem ter o caráter representacional da experiência alterado. No entanto, são adaptáveis e recalibráveis. O modo como a *experiência* representa o mundo é fixado pelas funções biológicas do sistema. Porém, a *cognição* é uma função que se desenvolve ontogeneticamente. E o modo como uma *crença* representa o mundo é determinado pelo aprendizado individual da capacidade de conceituar. Podemos ter a mesma experiência e conceituá-la de forma diferenciada. Pelo aprendizado, eu posso alterar o que eu acredito e penso quando experiencio o mundo, mas não posso alterar a qualidade do que experiencio. A consciência da experiência altera aquele que a experimenta, mas não modifica a experiência. A diferença entre as experiências que temos consciência e as que não temos não está na própria experiência, mas naquele que tem a experiência, ou seja, no que a pessoa sabe a respeito da experiência que está tendo.

A aquisição da capacidade de conceituar permite a aquisição de conceitos sobre o eu, o que permitirá fazer referência ao eu e representar a experiência subjetiva. A capacidade de representar a experiência subjetiva altera significativamente o *modo* como a pessoa experimenta seu ponto de vista e permite a emergência da consciência de si como uma criatura com uma perspectiva dentre outras.

## **2.2 A Emergência da Autoconsciência**

A observação do comportamento dos recém-nascidos evidencia aspectos importantes da experiência de ser um ponto de vista distinto dos outros e de se comportar no mundo a partir de uma perspectiva. Até tornar-se uma pessoa e ser capaz de exibir fenômenos da perspectiva forte da primeira pessoa, nós humanos passamos por cerca de 4 anos de desenvolvimento com mudanças qualitativas consideráveis que culminam com a emergência

da capacidade de autoconsciência. Do nascimento até 4-5 anos de idade, são observados diferentes níveis do desenvolvimento da representação do eu que indicam alterações importantes no modo como as informações sobre o eu são representadas. Existem cinco passos que correspondem a diferentes níveis de autopercepção: diferenciação eu-mundo (0-2 meses); agente situado (2-9 meses); agente perspectivo (9-18 meses); agente representacional (18-36 meses); agente meta-representacional (3-4 anos)<sup>18</sup>.

As pesquisas sobre o desenvolvimento do bebê mostram que, desde o nascimento, ele é capaz de perceber seu próprio corpo como uma entidade diferenciada de outras no ambiente, ou seja, capaz de fazer a **diferenciação eu-mundo** (ver Gibson 1979, Neisser 1988, Rochat 2003, Rochat & Striano 2000, Butterworth 1995). A tradição de pensamento gibsoniana busca enfatizar a experiência precoce de diferenciação, opondo-se à idéia hegemônica em psicologia de que o bebê nasceria num estado de confusão com seu entorno, e que suas ações seriam respostas automáticas a estímulos do ambiente – tese defendida por James e Piaget (Butterworth 1995). Essa noção foi amplamente articulada por James Gibson (1979), que preconizava a idéia de que, desde o nascimento, o bebê se comporta como um centro de ação e percepção cuja experiência perceptiva acarreta autopercepção e um sentido implícito do próprio corpo. Como afirma Gibson:

“A informação sobre o self acompanha a informação sobre o ambiente, e as duas são inseparáveis. A ego-recepção acompanha a exterocepção, como se fosse o outro lado da moeda. A percepção tem dois pólos, o subjetivo e o objetivo, e a informação está disponível para especificar ambos. Percebemos o ambiente e co-percebemos a nós mesmos.” (Gibson 1979: 126)

Para Gibson, nascemos biologicamente equipados com as funções proprioceptivas e exteroceptivas, que são funções do sistema perceptivo que se dirigem ao corpo e ao mundo. A propriocepção traz informação<sup>19</sup> sobre “os aspectos mutáveis e persistentes do corpo e a exterocepção dos (com os) aspectos mutáveis e persistentes dos objetos e acontecimentos extracorporais” (Costa 2004: 61). Portanto, todo ato perceptivo é auto-informativo. A concepção ecológica, como esclarece Costa:

---

<sup>18</sup> Esses cinco níveis foram adaptados das classificações apresentadas por Rochat (2003) e Fonagy (2004), que mostram como a autoconsciência, no caso de Rochat, e o sentido da atividade do agente, no caso de Fonagy, se desenvolvem a partir de cinco passos observados na ontogênese.

<sup>19</sup> O termo ‘informação’ usado pelos psicólogos cognitivistas tem um significado diferente da semântica informacional de Dretske. Segundo Neisser (1988), na teoria ecológica da percepção de Gibson, a informação existe objetivamente e especifica as propriedades dos objetos e eventos em virtude dos princípios físicos e é obtida por aquele que percebe.

“defende a idéia de que o comportamento movente e senciante dos animais é a maneira que eles têm de descobrir e utilizar os recursos do ambiente, (...) e, à medida que o indivíduo explora os recursos do ambiente apropriados à sua auto-regulação, auto-reprodução e satisfação, desenvolve a autopercepção de si.” (Costa 2004: 61)

A abordagem ecológica do self criou uma série de experimentos que confirmam a capacidade perceptiva de se discriminar do ambiente e se comportar como um centro de ações intencionais. Nessa acepção funcionalista do self (Eleanor Gibson 1995), privilegiam-se as diferentes informações que o self é capaz de obter por meio da percepção e interação com os objetos físicos e culturais.

Ulric Neisser (1988) chama de *self ecológico* o self diretamente percebido em relação ao seu ambiente físico imediato, capaz de estar consciente (*awareness*), de estar atento. Ao perceber o ambiente, ele é capaz de se situar espacialmente e de agir como um ponto de vista subjetivo em relação ao seu entorno, por meio de padrões perceptivos que especificam sua posição e seu movimento. Desde o nascimento, a partir de seus movimentos e ações no ambiente, o bebê capta informações perceptivas que especificam seu próprio corpo como uma entidade única no mundo. A capacidade inata de diferenciação eu-mundo é atribuída a experiências perceptivas que especificam o corpo do bebê como oposto a outras entidades perceptivas do mundo. A experiência de tocar-se a si mesmo, de ouvir seus próprios sons, de visualizar seus próprios movimentos, são experiências perceptivas básicas que auto-especificam o corpo do bebê (Rochat 2003).

Existem várias pesquisas que mostram empiricamente a capacidade inata de se diferenciar e de se autoperceber<sup>20</sup>. Philippe Rochat e colaboradores (Rocha & Hepos 1997, *apud* Rochat 2003) observaram o comportamento de fuçar de recém-nascidos (reconhecer com o focinho) (*rooting reponse*). Eles compararam o comportamento de fuçar em bebês com 24 horas de vida observando a estimulação tátil tanto originada pelo dedo do experimentador

---

<sup>20</sup> George Butterworth e colaboradores (1995) fizeram uma série de estudos sobre a propriocepção visual dos bebês para testar a capacidade do bebê de discriminar o movimento de seu próprio corpo do movimento dos objetos do ambiente. Esses estudos mostram que muito precocemente, antes mesmo de serem capazes de sentar, os bebês são capazes de ajustar sua estabilidade da postura em relação à estabilidade de seu entorno visual. No estudo do quarto movente (*moving room*), os bebês são presos a carrinhos e colocados num quarto suspenso do chão que se movimenta, de modo que a informação do fluxo óptico capta uma perda de balanço mesmo quando o bebê permanece numa postura estável. Observa-se que o bebê é capaz de distinguir entre o seu movimento (automovimento) e o movimento do ambiente (alomovimento), e perceber a informação proprioceptiva visual como especificando uma mudança na localização espacial do seu corpo. Esse estudo revela que os bebês são capazes de fazer a distinção entre uma alteração de lugar de seu corpo e uma alteração do estado do ambiente.

quanto a auto-estimulação que a criança espontaneamente fazia com um dos dedos de suas próprias mãos em contato com a bochecha. O resultado mostra que os recém-nascidos fuçam muito mais significativamente quando a estimulação é externa do que quando se auto-estimulam. Esse resultado confirma a hipótese de que os recém-nascidos não nascem num estado de confusão com o ambiente, mas, pelo contrário, são capazes de diferenciar entre estimulações e toques que são originados de seu próprio corpo (auto-estimulação) e de fontes externas (alo-estimulação). O bebê é capaz de discriminar entre eventos perceptivos que são gerados tanto pelo eu quanto pelo não-eu.

A partir dos dois meses, o bebê começa a se comportar, mesmo que de forma rudimentar, como um **agente situado** (Rochat 2003) cujos movimentos são ações orientadas por objetivos específicos que se dirigem a aspectos particulares do ambiente, numa performance que mostra capacidade de discriminação e preferência mesmo que rudimentares.

Muito antes de ser capaz de se reconhecer como uma imagem ou rosto, o bebê se percebe como um centro de atividade e se comporta como um agente capaz de controlar suas ações e interações. É capaz não apenas de obter informações de seus estados mentais e corporais, mas de obter informações do sistema como um todo. Ele desenvolve progressivamente a atitude de contemplação (*contemplative stance*) e reciprocidade frente ao mundo e a si mesmo, situando seu corpo em relação a outras entidades no ambiente (Rochat 1999). A atitude contemplativa (Rochat 2001) indica a autopercepção de si como capaz de alterar o ambiente e o início do desenvolvimento da capacidade de controlar e escolher suas ações sobre o seu corpo e o ambiente. O bebê deixa de agir apenas de forma imediata e passa a explorar seletivamente a consequência de suas ações sobre o ambiente. Ao mesmo tempo, com o acúmulo das experiências, aprende um repertório novo de ações que torna sua interação muito mais complexa e o permite escolher entre ações variadas em função do objetivo visado. Ele aprende a controlar suas ações e a ter expectativas dos comportamentos do meio, sobretudo dos seres animados (Edward Reed 1995).

As pesquisas sobre imitação neonatal de Andrew Meltzoff e Keith Moore (1995) mostram que, ao final dos dois meses, com a emergência do sorriso social, das interações face-a-face e das protoconversações, o bebê desenvolve um comportamento imitativo sofisticado e voluntário e se torna capaz, não apenas de discriminar suas ações corporais do modelo que está imitando, como também de mapear seu espaço corporal em relação ao espaço corporal do

outro, situando-se em relação ao modelo percebido que imitam. O bebê imita a dinâmica dos atos que observa. Ele necessita ver o ato em progresso para imitá-lo, pois ainda não é capaz de antecipar a seqüência de movimentos do outro. Por meio das interações face-a-face, se estabelece entre o bebê e o adulto um ‘espelho afetivo’ (*affective mirroring*) (Rochat 2002), que é o fenômeno da ressonância emocional, no qual o adulto responde reciprocamente ao bebê. Este começa a ser capaz de combinar suas habilidades interativas e perceptivas e entrar num modo intersubjetivo de interação e de vida relacional.

A atitude contemplativa é uma condição necessária na ontogênese para que o bebê chegue a desenvolver a perspectiva da primeira pessoa. Observa-se um continuum na experiência do bebê que, entre 2-9 meses, começa, gradativamente e com níveis crescentes de complexidade, a se comportar como um *sistema intencional* (Dennett 1981b), com a emergência das ações intencionais, no sentido dennettiano do termo, com objetivos funcionais que não são apenas aqueles inatos, e a criação de novas ações em função dos objetos encontrados no ambiente próximo. Portanto, o bebê que se recusa a pegar um objeto que está fora de seu campo de ação ou que utiliza seu pé para atingir um móbile suspenso diante de si, pode ser dito capaz não apenas de se diferenciar do objeto, mas também de adotar diferentes estratégias de ação em relação ao objeto e escolher os movimentos mais adequados. Ele explora seletivamente os provimentos dos objetos (*affordances*) (ver Costa 2004) que estão ao seu alcance. Ao final desse período, o bebê já pode ser considerado uma agência autônoma e coordenada (E. Gibson 1995), com capacidade para modular suas ações a partir das expectativas do comportamento de seu meio social.

No decurso do desenvolvimento emocional, é somente quando o bebê desenvolve a capacidade de se perceber como tendo um ponto de vista que ele é capaz de exibir a perspectiva *fraca* da primeira pessoa. Aos nove meses, o bebê se torna um **agente perspectivo** – o que Michael Tomasello (1995) chama de ‘*becoming perspectival*’. Com a emergência da locomoção, surge a possibilidade de se perceber o mundo de várias posições. A função da locomoção está diretamente relacionada com o desenvolvimento da habilidade de se autoperceber como uma perspectiva (Reed 1995), possibilitando um movimento que não se destina a funções específicas, como pegar objetos, mas a explorar o ambiente, e isso permite ao bebê explorar perspectivas alternadas. Surge também a capacidade de apontar, por meio das vocalizações e dos gestos usados para indicar os objetos. Os jogos imitativos também se

tornam complexos e passa a ser suficiente para criança observar o final do ato realizado pelo adulto para imitá-lo, exibindo a capacidade de antecipar o processo. O bebê começa a consultar a face do outro em busca de informações sobre os recursos do ambiente e passa sistematicamente a integrar o olhar do outro em suas explorações pelo meio circundante. Antes de ser capaz de poder se conceber como uma perspectiva, tem a experiência de explorar pontos de vista distintos e promover padrões específicos de atividades com outros. Essa nova função possibilita que partilhe os provimentos dos objetos (*affordances*) do ambiente com os outros, e se engaje em interações ternárias que incluem o bebê, o adulto, o objeto ou outro adulto (Reed 1995).

Portanto, a partir dos nove meses de vida, o bebê exhibe esboços do que vai se tornar futuramente a capacidade de se distanciar de sua própria perspectiva, de ter crenças e desejos sobre suas crenças e desejos. Os *sistemas intencionais de segunda ordem*, tal como concebido por Dennett, são característicos de criaturas que **não** têm o conceito de ‘eu’, mas se comportam como organismos que têm uma perspectiva de si em relação ao seu entorno. É dessa perspectiva egocêntrica que experimentam o mundo. Pode-se dizer que são conscientes de seus estados mentais e corporais, mas não são conscientes de serem possuidores desses estados. Portanto, não são autoconscientes, pois não têm o *conceito* de crenças ou desejos, nem o *conceito* de si mesmo como um sujeito de suas crenças e desejos. O bebê age de sua própria perspectiva sem saber que sua perspectiva difere de outras perspectivas. Autolocaliza-se frente aos objetos e pessoas a partir de sua posição espaço-temporal e age a partir desse ponto de vista tomando a si mesmo como centro. Seus estados são autolocalizados e referidos ao seu ponto de vista. No entanto, ainda não é objeto de seu pensamento e não é capaz de ter auto-representações das quais poderia estar consciente ou pensar sobre<sup>21</sup>, pois ainda não adquiriu a capacidade reflexiva da autoconsciência.

O bebê de 9-12 meses é capaz de dar o passo intencional na direção do que vai se constituir, depois dos 3 anos, a perspectiva forte da primeira pessoa. Ele apresenta a habilidade rudimentar de atribuir tipos de estados mentais – tais como estados de atenção e de emoção – a outros agentes e interpretar o comportamento do agente como racional e dirigido a um

---

<sup>21</sup> Quando abordar a gênese do eu, veremos que estas representações só vão aparecer com a emergência do *self* expandido (*extended self*).

objetivo que o permite prever sua ação futura em direção ao objetivo tendo por base uma nova situação (Tomasello 1995).

Esse processo culmina aos 14-18 meses com o desenvolvimento da capacidade de se reconhecer no espelho (Rochat 2003). A capacidade de identificar a própria imagem no espelho pode ser vista como o primeiro índice de objetificação do eu. É um índice não apenas de reconhecimento da imagem corporal, mas de que aquela imagem o identifica, se refere ao eu. Ou seja, ele não somente diferencia o seu ponto de vista da perspectiva exterior, mas também se torna capaz de perceber essa distinção. É nesse período que surgem condutas de embaraço relacionadas a contextos em que a criança recebe atenção prolongada do adulto sobre si mesmo (Rochat 2002, 2003). O desenvolvimento cronológico mostra que o reconhecimento da imagem no espelho vem antes da utilização da palavra 'eu'. Existe, portanto, uma capacidade gestual implícita de se referir a si mesmo, antes mesmo da capacidade se manifestar linguisticamente, por meio do uso do pronome 'eu'. O reconhecimento da imagem no espelho como a imagem de si mesmo distinta de outras pode ser visto como sinal da capacidade de identificar uma representação de si mesmo, e serve como um marco adaptativo do desenvolvimento do que vai ser chamado depois de autoconsciência. Não obstante, ter um ponto de vista subjetivo e ser capaz de discriminar sua imagem entre outras não é suficiente para exibir a perspectiva *forte* da primeira pessoa.

No segundo ano de vida, com a aquisição da linguagem, a criança desenvolve a capacidade de compreender de um *modo mental* o agir intencional e se autoperceber como um **agente representacional** cujas ações são causadas por estados mentais prévios tais como crenças e desejos. E começa a entender que as ações podem mudar mentalmente assim como as propriedades físicas do mundo. Eles compreendem, por exemplo, que um gesto declarativo de apontar pode alterar o estado de atenção de outro agente (Fonagy 2005).

O desenvolvimento da perspectiva *forte* da primeira pessoa culmina com a aquisição da linguagem e a emergência da perspectiva da teoria da mente. Aos 3-4 anos, a criança desenvolve a habilidade de compreender e atribuir estados emocionais e complexos estados intencionais aos outros, como as falsas crenças, de entrar nos jogos de faz-de-conta e de fazer referência verbal a si mesmo. Surgem também as emoções secundárias como vergonha, orgulho, desprezo e culpa. A compreensão da agência em termos da causação mental passa a incluir a representação epistêmica dos estados mentais, emergindo o **agente meta-**

**representacional** cujas ações são compreendidas como estados mentais intencionais (crenças e desejos) que são construídas como representações conceituais do mundo (Fonagy 2005). A posse das representações e das crenças capacita a criança a poder representar algo como falso ou verdadeiro. Esse processo culmina aos 5 anos de idade com a capacidade de passar no teste da crença falsa, que avalia a habilidade de representar o mundo mental do outro e, conseqüentemente, compreender a perspectiva do outro. Isso permite à criança a compreender a referência a si mesmo (autoconsciência) e a seus estados mentais (autoconhecimento). Ela adquire a capacidade representacional de relatar as memórias de suas atividades intencionais e experiências numa organização coerente de causa e tempo, levando à emergência do *self expandido* no tempo, que é a compreensão da sua história autobiográfica.

A capacidade de pensar em si mesmo como possuidor desses pensamentos traz outras habilidades para a pessoa. A aquisição da autoconsciência nos torna capazes de pensar nos outros como possuidores de seus próprios pensamentos e como pontos de vistas subjetivos diferentes dos nossos. Conseqüentemente, tornamo-nos capazes de refletir sobre nós como possuidores de nossos próprios desejos e ter o que chamamos de desejos ou pensamentos de segunda-ordem sobre nós mesmos. Esta capacidade, como esclarece Baker, é a capacidade de se conceber como ‘si mesmo’, como um ‘eu’ distinto de todo o resto, e de formular um conceito sobre essa distinção, ou seja, é “não apenas ter pensamentos exprimíveis por meio do ‘Eu’, mas também conceber a si mesmo como o possuidor desses pensamentos” (Baker 1998).

No debate atual sobre a noção de autoconsciência, uma das questões que se coloca é se a informação que o eu obtém, desde o nascimento, sobre seus próprios estados pode ser chamada de autoconsciência ou autoconhecimento. Alguns acham que só podemos considerar o fenômeno da autoconsciência quando lidamos com seres de linguagem, capazes de conceituar lingüisticamente sua experiência. Desde que Wilfrid Sellars (1997) propôs distinguirmos entre a consciência discriminativa (*awareness*) (‘ter a experiência de x’) e a consciência reflexiva (*consciousness*) (‘ter o conceito de x’, ‘saber que tipo de coisa é x’), existe em filosofia a tendência a separar a experiência subjetiva pré-lingüística da experiência lingüística reflexiva.

Muitos filósofos, como Dretske, fazem uma clara distinção entre pensamento (julgamento, crença) e sensação (experiência). A experiência de uma sensação (ouvir a Nona Sinfonia de Beethoven) não requer o conceito (saber o que é uma sinfonia) para saber do que

se está consciente perceptivamente. Nos termos de Dretske, a distinção é feita entre a consciência do objeto (a experiência da propriedade x) e a consciência do fato (a crença de que isto é x). Nesse sentido, ter a experiência de K é diferente de ter pensamentos sobre K, assim como ter representações de fenômenos sensoriais difere da representação de fatos conceituais em crenças e julgamentos. A experiência não requer conceito, mas o conhecimento requer conceito. Os sistemas perceptivos têm a função de indicar a qualidade fenomenal da experiência por meio de funções herdadas filogeneticamente, e não podem ter o caráter representacional da experiência alterado. A função de indicar o conteúdo representacional é adquirida ontogeneticamente, por meio do aprendizado individual. O desenvolvimento individual pode mudar não o que a mente filogeneticamente deve representar, mas pode alterar o que a experiência pode representar com a aquisição de novas representações sensoriais e conceituais. Para Dretske, a diferença entre sensação e conhecimento está no modo como a informação deve ser codificada. A condição que uma representação analógica (de uma sensação) deve satisfazer para ser qualificada como experiência é que a informação que ela tem a função de fornecer tem que estar disponível para ser digitalizada, ou seja, para ser representada conceitualmente (em crenças e intenções).

Contudo, esse esquema lógico não explica como se passa do nível lógico da experiência para a reflexão, como se passa da sensação para o conhecimento. Em seu esquema, parece ser apenas a aquisição da linguagem que importaria na mudança de um nível a outro.

A abordagem desenvolvimentista mostra que, para além desse esquema lógico de interpretação, há um continuum de experiências que acontecem e precisam acontecer, para que haja a transformação da consciência discriminativa para a consciência reflexiva. Há uma gradação. Nesse sentido, a história das interações causais (da filogênese e da ontogênese) é fundamental para explicar como o comportamento discriminativo se transforma nas práticas sociais de justificação. Ou seja, como a ‘presença ao mundo’ se torna uma autoconsciência capaz de justificar suas ações.

Todas essas pesquisas mostram que, muito precocemente, os bebês aprendem coisas sobre os objetos que são alcançáveis no mundo físico e sobre suas capacidades em alcançá-los e como ajustar sua configuração e postura corporal. As transações com o mundo físico e social levam ao conhecimento do self como agente intencional e social. Tenderíamos a concordar

com a psicologia gibsoniana e a fenomenologia que essas experiências geram uma auto-informação não-conceitual do eu, tanto de seus estados corporais e mentais quanto do sistema como um todo. Na abordagem ecológica, a percepção do mundo e a percepção do eu são dois aspectos do mesmo processo. Ao percebermos o ambiente, temos também informações sobre o eu. Mas não há ainda representação conceitual da entidade eu, temos apenas informações sobre nossos estados mentais e corporais, da perspectiva fraca da primeira pessoa.

Apesar de fazerem o diagnóstico correto da situação, a conceituam de maneira equivocada. Como vimos, as pesquisas mostram que os bebês obtêm informação sobre seu corpo como uma entidade diferenciada, discriminada de outros objetos e pessoas percebidos no ambiente, e que essa capacidade favorece o desenvolvimento da noção de eu. O bebê aprende com essas informações a modificar seu comportamento e desenvolve outras formas de interação. Mas isso não autoriza o uso de termos como ‘consciência do eu como uma entidade entre outras’ (Rochat 1995), ‘autoconsciência irrefletida’ (Neisser 1988), ‘self como objeto da consciência’ (Butterworth 1995), ou ‘autoconhecimento’ (*self-knowledge*) (Neisser 1988) como os autores desenvolvimentistas costumam sugerir. A não ser que assumamos uma outra noção de autoconhecimento. A crença é uma condição para o conhecimento. Se definimos o conhecimento como dependente de conceitos não é possível falar de autoconhecimento antes de termos conceitos, e o mesmo se aplica para a noção de autoconsciência. O fato do bebê trazer informações analógicas sobre si mesmo que serão futuramente digitalizadas, para usar a expressão de Dretske (1997), e transformadas em representações sobre o ‘eu’, não significa que o bebê tenha consciência de si como uma agência representacional ou meta-representacional.

O fenômeno apontado por essas pesquisas é que, mesmo que convençionemos que o autoconhecimento seja um conhecimento conceitual do eu e de seus estados, e, portanto, dependente da capacidade de usar símbolos, ele é determinado, em alguma medida, pelo modo como o bebê, desde o nascimento, processa as informações perceptivas que capta de seu ambiente físico e social. Os bebês têm informações que representam, de forma não-conceitual, seus estados mentais e corporais, que alteram as ações do eu, mesmo que não sejam capazes de se representar como possuidor desses estados. Há uma pré-história importante que conduz ao desenvolvimento da perspectiva forte da primeira pessoa. Quando os símbolos lingüísticos são adquiridos, a criança já tem uma experiência anterior de

pertencimento a um corpo que interage com objetos físicos e corporais. Somente um corpo que tem a experiência prévia de ser um agente no mundo pode vir a ‘pensar sobre essa experiência’. Antes mesmo de ter conceitos, o bebê experimenta uma série de interações que fornecem informações, que são representações não-conceituais, sobre o eu e seus estados. A perspectiva da primeira pessoa fraca, ainda não objetificada, é um centro de motivação da ação e a criança é capaz de se perceber como esse centro ativo que deflagra ações (E. Gibson 1995).

Sendo assim, a autoconsciência não é algo que irrompe automaticamente e simplesmente com a aquisição da linguagem. A capacidade de conceitualização do eu e de fazer referência a si mesmo emerge progressivamente no desenvolvimento. Existem níveis cumulativos que se consolidam a partir de fundamentos ecológicos do sentido implícito do eu e que, progressivamente, se tornam explícitos com a aquisição da linguagem e da capacidade reflexiva. Há uma transição fundamental, pois a forma como vivenciamos as interações pré-conceptuais imprimem uma marca no modo como as interações lingüísticas serão vivenciadas. Antes de ser capaz de conceitualizar e, portanto, de ter consciência de si como ‘eu’, ele tem acesso a uma série de informações sobre o ‘eu’ que é a experiência subjetiva da primeira pessoa. Tem a experiência de ser um agente que controla suas ações, se situa e se diferencia face ao ambiente, o que pode ser evidenciado por seus comportamentos (imitação, protoconversações, ajustes corporais, localização em relação a outros objetos). O bebê que chora ou grita, ou aponta na presença de dor pode ser dito apresentar um ‘marcador de perspectiva’, ou seja, experimenta o mundo de sua posição encarnada. Porém, não se percebe como possuidor desse estado, pois é um estado que ainda não foi auto-atribuído. Ainda não tem uma perspectiva de si como um ponto de vista dentre outros.

A criança se torna consciente de si mesmo e de suas próprias experiências quando os conceitos necessários à capacidade de autoconsciência são adquiridos. Sem os conceitos, a criança tem a experiência autoperceptiva, mas não tem recursos para compreender o que é estar consciente de si. A autoconsciência emerge a partir da aquisição do conceito de representação e referência de si. Se Baker e Dretske têm razão (e acredito que tenham), não é possível falar de objetificação do eu antes da emergência da capacidade de conceitualização do eu. A conceitualização do eu se expressa pela capacidade de fazer referência a si mesmo, seja por gestos, símbolos, palavras ou frases.

A perspectiva da primeira pessoa fornece a capacidade de pensar em nós mesmos de uma forma única, “de dentro”. Não pensamos em nós mesmos como um objeto ou uma entidade ‘self’ como um corpo ou um eu desencarnado a ser introspectado. O ‘eu’ não se refere a uma entidade a qual tenho acesso direto, refere-se à pessoa em sua “concretude corporal”. Parafraseando Baker (1998), referir-me a mim mesma pelo pronome ‘eu’ não é diferente de referir-me à pessoa Cláudia. O que é particular acerca do ‘eu’ é que nos referimos a nós mesmos de um modo que os outros não podem, **de dentro**. Mas o que é essa capacidade de pensarmos em nós mesmos ‘de dentro’?

### **3. Autoconsciência e Autoconhecimento**

No sentido filosófico, a noção de autoconsciência apresenta distinções importantes que nem sempre correspondem às distinções empíricas observadas em psicologia. O diálogo profícuo entre aquilo que pode ser observado e descrito na psicologia e a necessidade de precisão conceitual na defesa dos argumentos tem estimulado alguns filósofos a proporem uma sistematização do campo de investigação da autoconsciência. Nosso interesse é investigar a noção de autoconsciência e os aspectos que caracterizam esse processo mental, com vistas a uma melhor compreensão do papel que ocupa na constituição do eu, ou seja, de um ser autoconsciente.

Em “Self-Consciousness: A Tutorial” (2006), Uriah Kriegel propõe cinco<sup>22</sup> destas distinções. Para os meus propósitos, detenho-me em duas delas:

#### **1. ‘autoconsciência do estado’ *versus* ‘autoconsciência da criatura’**

#### **2. consciência do ‘self-como-objeto’ *versus* consciência do ‘self-como-sujeito’**

A primeira distinção é entre a consciência de si como self e a consciência de um estado que ocorre no self. A autoconsciência do estado pode ser diferenciada entre a consciência de um estado mental e a de um estado corporal. A consciência de um pensamento em particular

---

<sup>22</sup> Kriegel (2006) propõe duas outras distinções: autoconsciência corporal *versus* autoconsciência mental; e consciência do ‘self-como-autor’ *versus* consciência do ‘self-como-possuidor’. Essas duas distinções serão abordadas no capítulo sobre os modelos constitutivos do eu, pois estão diretamente relacionadas com o processo de desenvolvimento do eu e da autoconsciência. A quinta distinção de Kriegel é a distinção entre a autoconsciência transitiva e a autoconsciência intransitiva. Essa distinção é um desdobramento analítico da distinção da consciência do ‘self-como-objeto’ *versus* consciência do ‘self-como-sujeito’, que expressa a diferença entre afirmar que “Eu estou autoconsciente de pensar que p”, que é a forma transitiva em que a autoconsciência é uma relação entre eu e meu pensamento; e “Eu estou pensando de modo autoconsciente em p”, que é a forma intransitiva na qual a autoconsciência é uma modificação do meu modo de pensar.

pode ser expressa pela frase “Estou consciente de mim mesma pensando que p”, e indica uma versão fraca da autoconsciência cujo objeto é um evento ou estado interno do self. Essa noção é usada para se referir ao conhecimento dos próprios estados mentais da pessoa e tem sido tradicionalmente denominada de **autoconhecimento** (*self-knowledge*).

A versão forte da autoconsciência é a consciência que toma a criatura, o próprio self, como objeto. A **autoconsciência da criatura** (*self-consciousness*) é a capacidade de fazer referência a si mesmo de forma reflexiva e consciente, ou seja, a referência a si mesmo como sujeito. A autoconsciência da criatura está na base do conhecimento da pessoa como um self existente. Distingue-se da capacidade de fazer referência a si mesmo de forma casual ou irrefletida. Posso fazer referência a mim mesmo sem saber que é a mim mesmo que me refiro. O self, ao referir-se a si mesmo, nem sempre sabe que está sendo auto-referido. Por exemplo, na conhecida situação imaginada por John Perry (*apud* Kriegel 2006), eu me encontro num supermercado e vejo um rastro de açúcar no chão deixado por alguém e penso “tem alguém fazendo uma confusão no supermercado”; no instante seguinte, me dou conta de que o saco de açúcar que está na minha cesta está furado e que sou eu mesma a pessoa que está deixando o rastro e, conseqüentemente, fazendo a confusão. Nesse caso, faço referência a mim sem ter consciência de que me refiro a mim mesma.

No **autoconhecimento**, o que está em questão é saber se me encontro em um determinado estado mental descrito por um predicado qualquer, como, por exemplo, na afirmação “creio que eu **tenho fome**”. O autoconhecimento visa o elemento predicativo da consciência, ou seja, como tomo conhecimento que estou com fome. Na **autoconsciência**, o que está em jogo é o elemento que identifica o sujeito, como, por exemplo “creio que **eu** tenho fome”, ou seja, como sei que sou eu que me encontro em tal estado. Ambas se manifestam lingüisticamente por uma identificação e uma predicação, mas a autoconsciência visa a identificação, ou seja, como o elemento referencial é identificado, e o autoconhecimento visa a predicação, ou seja, como o elemento predicativo é conhecido<sup>23</sup>.

A segunda distinção é uma subdivisão da noção de autoconsciência da criatura que pode ser distinguida entre ‘a consciência de si *qua* objeto’ e ‘a consciência de si *qua* sujeito’. Aqui a distinção conceitual a ser feita é entre o self tomado como *objeto* de pensamento e o

---

<sup>23</sup> Essa sistematização das noções de autoconsciência e autoconhecimento foi sugerida por Roberto Horácio Pereira em comunicação pessoal.

self tomado como *sujeito* do pensamento. Embora exista uma única entidade, o eu, existem dois conceitos separados para essa entidade. O eu como objeto do pensamento se refere às várias informações objetivas que identificam o eu (tenho 36 anos, sou mulher, peso 51 quilos, estou sentada, etc.). A autoconsciência como sujeito é o ‘eu’ descrito como aquele que tem o pensamento. A única informação a ser obtida é que o sujeito é aquele que capta a informação, ‘eu sou aquele que realizo o pensamento’. Nessa acepção, o eu não é parte do mundo, ele é o limite do mundo, a perspectiva do mundo. Esse ‘eu’ intangível não pode ser naturalizado, pois não faz parte da natureza do mundo. Essa distinção é aceita por toda tradição filosófica.

Com Kant, a distinção já aparece entre o ‘sujeito transcendental’ e o ‘sujeito empírico’. O ‘sujeito transcendental’ não tem estatuto ontológico e é apenas veículo do pensamento. William James (1992 [1890]) enfatiza que o self possui aspectos distintos, apesar de ser uma unidade<sup>24</sup>. A diferença central é entre o eu que se dá conta da sua existência – o eu da autoconsciência – e o eu que age – o eu empírico. O ‘eu empírico’ (*me*) é o self como ‘objeto que pode ser conhecido’. O ‘eu puro’ (*I*) é o self como ‘sujeito que conhece’. Existem dois modos de apresentação sob os quais a pessoa poderia estar consciente de si, que correspondem a esses conceitos: podemos estar conscientes de nós mesmos sob a descrição do ‘mim’ (eu-como-objeto) e do ‘eu’ (eu-como-sujeito). Na consciência do ‘eu’, penso em mim como aquela que pensa, “estou consciente que penso p”; e na consciência do ‘mim’, “estou consciente de mim mesma pensando p”. Como Kriegel salienta, nas expressões em que usamos ‘mim’, há uma distância conceitual entre a coisa que pensa e a coisa que está sendo pensada, e os aspectos do ‘mim’ parecem posses exteriores. Pois embora eu esteja pensando sobre mim mesma, não estou pensando em mim como aquela que está pensando o pensamento. No primeiro caso, estou pensando em mim mesma precisamente como a coisa que está tendo o pensamento. A diferença entre esses dois conceitos não implica uma

---

<sup>24</sup> Em “The Consciousness of Self” (1992 [1890]), James tenta escapar das metafísicas que concebiam o self da autoconsciência como um ‘olho interno’ que espelha representações privilegiadas, e evitar a armadilha conceitual da divisão kantiana ‘ego empírico’ x ‘ego transcendental’. Mas, guarda a divisão tradicional do eu empírico (o self que é conhecido) e o eu sujeito (o self que é conhecedor), mesmo que seja para questionar os problemas oriundos de tal divisão. A solução encontrada foi postular um fluxo da consciência, na tentativa de afastar qualquer hipótese que propusesse um princípio transcendental de unidade, ou uma ‘alma como meio combinador’, em que pudéssemos ancorar a noção de identidade da pessoa. Mas, se o ‘self conhecedor’ é um fluxo de consciência, como ele explica a identidade do self? Para James, a identidade da pessoa, a experiência de continuidade da pessoa, ou seja, a sensação de que continuamos os mesmos, é dada por um momento em que o self se apropria de um fluxo de consciência produzido anteriormente.

diferença ontológica, pois não acarreta diferença entre as propriedades que poderiam ser descritas por esses conceitos.

O que vai nos interessar no estudo da autoconsciência da criatura é a consciência do ‘self como *sujeito*’. Nessa forma de autoconsciência, não estou apenas consciente de mim mesma, mas estou consciente de mim mesma como um self. Há uma série de características epistêmicas e semânticas atribuídas à autoconsciência. Kriegel (2006) fornece uma lista dessas características, mas, como várias dessas características foram desacreditadas pelas pesquisas contemporâneas, vou me concentrar na análise do privilégio epistêmico do autoconhecimento e na imunidade ao erro por identificação da auto-referência.

### **3.1 Autoconhecimento e Introspecção**

Designamos por autoconhecimento o conhecimento direto que temos de nossos estados mentais expresso pela apreensão espontânea – automática e irrefletida – corriqueira de nossas convicções, intenções e expectativas. É consensual em filosofia a idéia de que há uma assimetria entre a forma como conhecemos nossos pensamentos e a forma como conhecemos os pensamentos dos outros. Desde Descartes, somos levados a considerar que temos acesso privilegiado aos conteúdos de nossos pensamentos, que podem ser conhecidos pela introspecção, e que este conhecimento é infalível e confiável. Descartes defende a autoridade do conhecimento da primeira pessoa de seus próprios estados mentais como um conhecimento de caráter *a priori* e que não precisamos conhecer o mundo externo para obtê-lo. No argumento do *cogito*, defende a tese da indubitabilidade, ou seja, a de que nada poderia abalar sua certeza de que ele está pensando em alguma coisa. Descartes toma como certo e óbvio que temos conhecimento privilegiado e seguro do que se passa na nossa própria mente, e que sabemos diretamente o que estamos pensando. Da evidência de que podemos eventualmente nos enganar quanto à veracidade de nossos pensamentos, Descartes deduz que todos nossos pensamentos podem ser falsos, mas que não podemos nos enganar quanto ao fato de estarmos pensando, pois não podemos ter dúvida dos conteúdos de nossos estados mentais. Portanto, posso me enganar quanto à veracidade do meu pensamento, mas não posso ter dúvidas do caráter seguro e infalível do “eu penso”.

A doutrina do autoconhecimento pode ser assim formulada: “se creio que me encontro no estado M, então sei que me encontro no estado M”. Existem três aspectos que concorrem

para a caracterização da noção de autoconhecimento e que a diferenciam de outras formas de conhecimento: privilégio epistêmico, acessibilidade por introspecção e autoridade da primeira pessoa.

Desde Descartes, os filósofos sustentam a tese do **privilégio epistêmico** do autoconhecimento. Privilégio epistêmico significa que o conhecimento de nossos próprios estados mentais é um evento indubitável. As afirmações mais importantes a favor do privilégio epistêmico do autoconhecimento são as teses da infalibilidade, incorrigibilidade, onisciência e autofamiliaridade (Kriegel 2006).

A doutrina da infalibilidade afirma que a consciência de si é sempre verídica e precisa. Não podemos ter crenças falsas sobre nossos próprios estados mentais e, portanto, não podemos ser corrigidos nas afirmações que fazemos sobre eles. Sempre que estamos conscientes de que pensamos em p, de fato estamos pensando que p. Essa doutrina afirma que: “Se creio que me encontro no estado M, então, (é verdade) que me encontro no estado M”.

A doutrina da incorrigibilidade garante que essas crenças são epistemicamente justificadas, pois o sujeito teria a posse de todas as evidências relevantes. Sendo assim, “Se creio que me encontro no estado M, estou justificado a acreditar que me encontro no estado M”. De acordo com esse postulado, se acredito em alguma coisa sobre minha vida mental, então, minha crença está justificada.

A doutrina da autofamiliaridade afirma que não é possível que algo se passe na minha mente sem que eu tenha consciência disso, sem que eu não perceba, pois os pensamentos são “íntimos ao eu”. Portanto, “se me encontro no estado M, então, creio que me encontro no estado M”. Ou seja, sempre que penso em alguma coisa eu inevitavelmente acredito que estou pensando em alguma coisa.

A doutrina da onisciência da autoconsciência defende uma tese forte do privilégio epistêmico na qual a pessoa sabe tudo o que acontece na sua mente e tudo que não acontece, e pode ser expressa pela tese: “Se me encontro no estado M, então, sei que me encontro no estado M”.

Várias pesquisas em psicologia e principalmente em psicanálise, com a descoberta das crenças e desejos inconscientes, mostram o equívoco de todas essas doutrinas. Atualmente poucos filósofos considerariam essas teses sustentáveis. As objeções à tese da infalibilidade

partem da constatação de que podemos estar certos ou errados a respeito de nossas crenças e que, muitas vezes, nossas auto-atribuições estão sujeitas à correção de outras pessoas.

Muitos filósofos defendem uma versão restrita dessas teses que tornaria o argumento mais plausível. Algumas dessas teses têm sido reformuladas de forma a restringir o privilégio epistêmico a um conjunto específico de estados mentais. A idéia é que haveria um tipo particular de estados que satisfaz às exigências cartesianas, seria o caso de sensações e sentimentos ou estados fenomenologicamente conscientes, passíveis de serem expressos pela proposição “se tiver uma sensação S, estou consciente de ter a sensação S”. Diversos autores, entretanto, mostram a falibilidade dessas teses. Eles contra-argumentam que existem sensações e estados conscientes que ocorrem sem a consciência da pessoa. Mesmo se substituirmos, em todas as afirmações, a expressão ‘estado M’ por sensações – como na frase “se tenho uma sensação S, então creio que tenho uma sensação S” – ou até mesmo por proposições – “se creio que me encontro no estado ‘penso que’, então, é verdade que penso que” –, não temos garantia da infalibilidade dessas teses, pois continuam podendo ser falseadas. Por exemplo, posso ter a sensação consciente do ruído constante do ar condicionado, sem estar consciente disso, ou posso ter a percepção consciente de um polvo de 48 tentáculos quando, na verdade, possui 47 tentáculos (Kriegel 2006).

Existem também tendências a enfraquecer essas teses. Por exemplo, podemos reformular a tese do autoconhecimento como “estou consciente de ter algum pensamento”; ou na tese nuançada “se estou autoconsciente de estar num estado S fenomenologicamente consciente, então, estou num estado fenomenologicamente consciente”. Uma versão alternativa da seguridade epistêmica defende a certeza do autoconhecimento, ou seja, a certeza (relativa ou absoluta) de que se encontra num certo estado mental. Essa tese baseada na justificação, ao contrário da tese da onisciência que se aplica a uma precisão da pessoa como um todo, se aplicaria a uma única crença auto-atribuída. Esses são os tipos de tese que parecem sobreviver às críticas às tradicionais doutrinas do privilégio epistêmico.

Outra característica do autoconhecimento é o fato de termos **autoridade sobre nossos próprios estados mentais**. Nessa acepção, os estados que imputamos a nós mesmos, nossas auto-atribuições, são presumidos como verdadeiros, e não estão disponíveis para serem questionados pelos outros. O que é especial à auto-atribuição é que cada um de nós é uma autoridade, por princípio, de nossos próprios estados mentais. Entender as auto-atribuições

implica aceitar a existência dos auto-atribuidores, ou seja, aceitar que as pessoas tenham autoridade para se auto-atribuir uma crença. Portanto, a autoridade da primeira pessoa não é opcional, pois o fato de nossas auto-atribuições não poderem ser falseadas pelos outros constitui parcialmente os conceitos psicológicos empregados por essas atribuições.

Tomar a autoridade da primeira pessoa como um padrão não significa se comprometer com nenhum fundamento epistemológico. Pois a autoridade da primeira pessoa acerca de seus próprios estados mentais pode estar garantida mesmo quando a pessoa não tem conhecimento desses estados. O que justifica tratar uma pessoa como autoridade de seus próprios estados se ela não se encontra numa posição epistemicamente privilegiada que assegure seu acesso a esses estados? Isso significa considerar, por princípio, todo ser autoconsciente como tendo um ponto de vista privilegiado de seus estados mentais, no sentido conferido por Baker da perspectiva da primeira pessoa.

Outra característica do autoconhecimento é que ele envolve um método especial de conhecimento: **a introspecção**. A introspecção é o único método para determinar os próprios estados mentais e oferece um acesso direto aos conteúdos mentais. No sentido epistêmico, podemos apreender nossos próprios estados mentais sem inferência, sem nos apoiarmos na observação ou evidência. No sentido metafísico, não há nenhum objeto ou estado que medie minha auto-atribuição de crenças (por exemplo, que eu agora estou pensando que devo ir ao banco) e seus objetos (meu pensamento de que devo ir ao banco). Nesse sentido, argumenta-se que minha auto-atribuição deve ser simultânea ao estado atribuído, e que não há nenhum fator epistêmico ou metafísico de mediação entre minha auto-atribuição e seu objeto. Portanto, o acesso aos meus estados é um conhecimento não-inferencial, sem mediação entre o sujeito que pensa e o estado mental do qual tem consciência.

O termo ‘introspecção’ tem sido usado para se referir a um tipo de percepção, de ‘olhar interior’, num sentido não-óptico, de um processo mental de observação e investigação da própria mente. A hipótese tradicional da teoria da introspecção é que o autoconhecimento é promovido por um ‘olho interno’ que observa a mente. Se a introspecção é entendida como um acesso epistêmico direto aos nossos próprios pensamentos, então não é possível nos enganarmos quanto ao conteúdo de nossos estados mentais. Isso levou Descartes a concluir que a acessibilidade epistêmica garantiria a indubitabilidade. Ou seja, se temos acesso direto aos nossos pensamentos não podemos ter dúvida de nossos pensamentos.

A crítica a esse modelo de introspecção visa à idéia de que o autoconhecimento envolveria a observação de estados mentais internos estáveis que estariam disponíveis para serem descobertos pela introspecção. Esse modelo tem sido chamado de concepção ‘ato-objeto’ (*act-object*) do autoconhecimento e está baseado na idéia de que a introspecção é um ato quase-perceptivo de reconhecer um objeto independente. Sidney Shoemaker (1996) mostra como o conhecimento introspectivo de nossos próprios estados mentais tem sido pensado como envolvendo um tipo de percepção e ‘sentido interno’ (*inner sense*), baseado num modelo perceptivo ou observacional. A própria palavra introspecção – que significa literalmente ‘olhar para dentro’ – é impregnada da metáfora visual, da idéia de um ‘olho interno’, um scanner interior que monitora a mente. Muitos filósofos foram atraídos por essa explicação paradigmática do autoconhecimento como visão do interior.

Shoemaker (1996) argumenta que o modelo do ‘ato-objeto’ fracassa em explicar a autoconsciência que ultrapassa a consciência perceptiva. Existem diferenças entre: ter acesso a uma crença ‘Brasília é a capital do Brasil’, ter acesso a uma sensação ‘estou com dor’, e ter acesso a objetos externos ‘vejo um gato’. Quando percebo um objeto externo como um gato, estou perceptivamente consciente do gato, e posso fazer a distinção entre o ato da consciência perceptiva (ver um gato) e o objeto gato que a consciência toma como objeto. Porém, quando tratamos de emoções ou sensações, o modelo não se aplica. Por exemplo, ao tomar consciência de meu rubor não posso fazer a distinção entre o ato de estar autoconsciente do rubor e a sensação do rubor. Não há um ato da autoconsciência e um objeto separado tomado pela autoconsciência. Podemos apreender a nós mesmos com um estado mental particular, como, por exemplo, estar embaraçado, se essa apreensão é entendida no sentido quase-perceptivo, ou se a apreensão for de algum modo fusionada ao estado mental particular e por causa disso assimilada.

Portanto, seria um equívoco considerar a forma como temos acesso ao mundo externo semelhante ao modo como temos acesso a nós mesmos. O acesso privilegiado que temos a nossos estados mentais não é igual ao acesso que temos ao mundo externo. Os estados mentais estão dinamicamente relacionados à reflexão da primeira pessoa. Eles não são entidades estáticas a serem observadas. À medida que consideramos um estado mental como pertencendo a nós mesmos, tornar-se consciente desse estado é já submetê-lo à avaliação racional. Nesse sentido, o autoconhecimento não consiste numa simples observação dos

próprios pensamentos. O que parece ser especial em relação ao método de conhecimento dos meus próprios estados mentais é sua qualidade dinâmica: os estados auto-atribuídos estão disponíveis à avaliação.

Porém, ainda assim, seria correto considerar a introspecção como um tipo de sentido perceptivo? Alguns filósofos acreditam que a introspecção e a percepção partilham aspectos importantes. Mas muitos filósofos que se opõem à visão internalista da mente olhando para si mesma, tentam formular um conceito de introspecção que preserve a autoridade da primeira pessoa, e explique a assimetria entre essas duas formas de conhecer: conhecer nossa mente e conhecer a mente dos outros. Apesar do termo introspecção conotar ‘olhar para dentro’, tem-se argumentado que o autoconhecimento requer, ao contrário, um olhar para o exterior. Isso significa afirmar que nos atribuímos nossos pensamentos ao ‘olhar para fora’, para os estados do mundo que esses pensamentos representam. Mas afirmar que para formar crenças precisamos considerar as evidências oferecidas pelo mundo, o que qualquer externalista aceitaria, não chega a ser um argumento contra a introspecção e o autoconhecimento.

No debate contemporâneo, o ceticismo quanto ao autoconhecimento é amplamente negado pela maior parte dos filósofos. Porém, não há um consenso sobre os critérios que distinguem o autoconhecimento do conhecimento do pensamento dos outros. Alguns filósofos, como certos behavioristas, negam até mesmo a possibilidade de que exista um método especial para determinar os próprios estados da pessoa. Ryle (1984), por exemplo, nega que o autoconhecimento seja verdadeiramente específico e que envolva um processo epistêmico direto e único. Ryle critica a noção cartesiana de que os estados mentais são privados e internos e que são diretamente acessíveis. Para ele, é um absurdo considerar que tenho acesso privilegiado e autofamiliaridade com meus estados mentais, denominando essas teses de dogma do *fantasma na máquina*. Há um erro categorial no modo cartesiano de categorizar os estados mentais como não-físicos e internos à mente. Ryle nega que as auto-atribuições gozam de uma seguridade epistêmica especial. A diferença entre o conhecimento de nossas mentes e o conhecimento da mente dos outros é uma questão de grau, e que o autoconhecimento é baseado na observação que fazemos de nossos próprios comportamentos. Nega que haja um acesso privilegiado por meio da introspecção, pois imagina que a apreensão imediata de um primeiro estado pressuporia um estado mental de segunda-ordem que o apreenderia, um tipo de reflexividade que tenderia ao regresso infinito. Como alternativa Ryle propõe que o

autoconhecimento seria um estado de estar atento (*to be aware*) ao que se está fazendo ou pensando.

Alguns externalistas, que adotam a atitude behaviorista, argumentam contra a idéia de que o autoconhecimento seja especial em relação ao conhecimento de outras mentes, pois afirmam que os conceitos associados aos estados psicológicos são entidades teóricas, e, como tal, partilhadas publicamente. Na démarche behaviorista, o autoconhecimento é obtido pela hetero-atribuição, ou seja, a referência e o predicado são atribuídos do ponto de vista da terceira pessoa por meio da inferência e da observação. Portanto, tanto as auto-atribuições quanto as hetero-atribuições são realizadas do ponto de vista da terceira pessoa por meio da inferência de comportamentos observados. Contudo, a adoção do behaviorismo não explica a assimetria observada entre o conhecimento de nossa mente e o conhecimento da mente dos outros.

Como resolver esse problema sem abandonar a idéia da autoridade da primeira pessoa? Como explicar o conhecimento direto e imediato que temos de nossos próprios estados que não é obtido nem pela auto-observação, nem pela conclusão de uma inferência?

Quando apresentarmos o externalismo perceptivo de Davidson, veremos que ele propõe uma forma de manter a autoridade epistêmica da primeira pessoa sem se comprometer com as teses fortes cartesianas. Ele propõe abrandar essas proposições e reformular a tese do autoconhecimento de forma a incluir os falibilismos de nossas auto-atribuições. A tese deflacionada do autoconhecimento poderia ser assim formulada: “se creio que me encontro no estado mental M, então, é possível que me encontre no estado mental M”. Para Davidson, se admitimos a autoridade da primeira pessoa então temos que admitir que existe um modo de conhecimento que mostra uma diferença sistemática em relação ao modo como conhecemos a mente dos outros. O ponto de vista da primeira pessoa não pode ser separado da sua experiência. Do ponto de vista empírico, não há uma desconexão entre minha experiência subjetiva e a compreensão do comportamento do outro. Isso ocorre simultaneamente. Mas não é a partir da minha experiência interna que conheço o mundo, a mente do outro ou minha própria mente. É por meio de minha interação com o mundo que me conheço. Para Davidson, como veremos, o que garante a relação de veracidade entre o eu e seus estados é a relação triangular entre eu, intérprete e mundo.

### 3.2 Autoconsciência e Eu-Mentalês

Em relação à autoconsciência da criatura, a primeira questão que se coloca é acerca da possibilidade ou não do ‘eu’ fazer referência a si mesmo. O “eu” tem sido considerado, por alguns, como inteiramente não-referencial e, por outros, como a única forma de referência.

A tradição cartesiana concebia o ‘eu’ como passível de ser ‘objeto’ de contemplação para si mesmo. Descartes acreditava que haveria um ‘sentido interno’, uma experiência privada a qual só o eu teria acesso. Teríamos uma percepção introspectiva do eu e seria esta a fonte de todo nosso autoconhecimento.

Contudo, a consciência introspectiva do eu vem sendo negada desde Hume. Ele nega a possibilidade do eu estabelecer uma relação perceptiva consigo mesmo. Segundo Hume, não é possível perceber o eu por um ‘sentido interno’. Ele nega a existência do eu *qua* sujeito da experiência e a possibilidade de se confrontar com o eu ‘como objeto’ da introspecção.

A refutação de Hume está expressa na conhecida passagem: “For my part when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other (...). I never can catch myself at any time without a perception, and never can observe anything but the perception.” (Hume 1888, *apud* Shoemaker 1996). Hume afirma não só que não é possível captar a si mesmo como um único estado mental, mas ele exclui qualquer possibilidade de uma consciência do eu, exclui mesmo a possibilidade da autoconsciência da criatura, deixando apenas a possibilidade da autoconsciência dos estados mentais.

Para Hume, todos os estados mentais são relacionais, portanto, ter um estado mental sempre consiste em ter uma relação com um certo tipo de percepção. Ele imagina uma forma perceptual de apreender o self, um encontro direto com o self como um olho da mente que se volta para o interior. Em sua acepção, admitir a autoconsciência seria admitir que o self poderia ser um objeto interno a ser observado. Como afirma Kriegel, Hume concebe um modelo introspectivo da autoconsciência, que envolve o emprego de um ‘sentido interno’, “um mecanismo interno cuja operação seria análoga em aspectos essenciais à operação do sentido externo” (2006: 5). Sua concepção do que seria o ‘eu *qua* sujeito mental’ implica que o eu não poderia ser percebido introspectivamente.

Portanto, Hume rejeita a idéia de que exista uma entidade ‘self’ que seja mais do que um fluxo de consciência ou um conjunto de experiências. Ele recusa a reificação do conceito de self como um substrato que sustenta as experiências internas. O que Hume nega encontrar

quando se volta para si mesmo é o self substantivo. No modelo humeano, a autoconsciência seria um objeto interno que requer a existência do self substantivo para ser observado. Isso o levou a deduzir que seria um erro imaginar que o self possa ser apreendido internamente pela consciência. Da premissa ontológica que ‘o eu não é uma substância imaterial’, ele deriva a premissa epistemológica que ‘não se pode conhecer o eu’. A autoconsciência como objeto interno seria um emaranhado de percepções e não o eu. A tese da intangibilidade do eu (*elusiveness thesis*) proposta por Hume afirma que não nos apresentamos na introspecção como entidades corporais, pois quando estamos autoconscientes de nossos pensamentos, crenças e sentimentos não nos apresentamos a nós mesmos como um eu corporal físico, portanto não nos apresentamos a nós mesmos como um ‘objeto’ (ver Shoemaker 1996).

Desde a negação de Hume sobre a introspecção do eu, importantes discussões psicológicas e filosóficas têm sido travadas sobre as implicações concernentes à natureza do eu e do autoconhecimento. A tradição filosófica se divide entre aqueles que supunham o ‘eu’ como um ‘arranjador’ dos processos mentais e capaz de fazer referência a si mesmo e aqueles que defendiam a intangibilidade do eu, o ‘eu’ como uma construção lógica. A tese humeana foi usada como argumento *prima facie* para a visão de que o self seria uma construção lógica. Segundo Wittgenstein e Anscombe (Kriegel 2006), por exemplo, a palavra ‘eu’ não tem referente, não se refere à coisa alguma, e não existe o ‘eu’ como o sujeito das experiências.

Contudo, como mostra Shoemaker, o modelo humeano vem sendo questionado. Muitos filósofos que aceitaram a refutação de Hume não consideram que o conhecimento direto que as pessoas podem ter dos fatos sobre elas mesmas sejam construções lógicas. Alguns filósofos aceitam o modelo perceptivo da introspecção, mas negam que o self estaria entre os objetos percebidos por esse ‘sentido interno’.

A versão oferecida por Shoemaker escapa das ameaças que a versão humeana oferece à autoconsciência. Apesar de restringir as possibilidades da autoconsciência, ele tenta garantir a sua plausibilidade propondo um outro modelo. Ele propõe que a autoconsciência seja analisada semanticamente a partir das frases que usamos para fazer referência a nós mesmos e a nossos estados mentais, ou seja, das afirmações em primeira pessoa nas quais usamos o pronome pessoal ‘eu’, os *I-thoughts* (Shoemaker *apud* Kriegel 2006), ou o eu-mentalês como propõe chamar Kriegel (2006). Segundo Shoemaker, a particularidade da auto-referência lingüística (o conceito de ‘eu’) deriva de aspectos da autoconsciência da criatura e da auto-

referência mental, ou seja, do dispositivo mental (o eu-mentalês) usado para se referir a mim como mim mesma. Este modo de nos referirmos a nós mesmos não permite equívoco. Shoemaker designa essa peculiaridade da semântica da autoconsciência de “imunidade ao erro por identificação relativa ao pronome de primeira pessoa”.

Nossos pensamentos podem ser representados como tendo uma estrutura composta que pode ser dividida em dois componentes separáveis: identificação e predicação. Quando penso em alguma coisa distinta de mim mesma posso me enganar de duas formas. Conforme o exemplo de Kriegel (2006), imaginemos que considero que o meu vizinho do lado é uma pessoa simpática. Posso me enganar quanto ao fato do meu vizinho ser uma pessoa simpática e posso me enganar quanto a quem é o meu vizinho do lado. Nesse caso, tanto o elemento predicativo (que ele é uma pessoa simpática) quanto o elemento identificador (quem é o meu vizinho do lado) podem ser falseados. Posso me equivocar tanto quanto à predicação como quanto à identificação.

Nesse sentido, podemos afirmar que há 3 tipos de privilégio semântico: (1) pensamentos imunes ao erro por predicação; (2) pensamentos imunes ao erro por identificação; (3) pensamentos imunes ao erro por predicação e por identificação. Para Shoemaker, apenas o tipo dois pode ser satisfeito quando falamos de autoconsciência.

Dentre os pensamentos autoconscientes, existem dois usos do pronome pessoal ‘eu’ que devem ser distinguidos na auto-referência: como ‘sujeito’ e como ‘objeto’. Por exemplo, se afirmo “acho que está fazendo calor lá fora”, posso me equivocar quanto ao fato de estar fazendo calor, mas não posso me equivocar que sou eu quem está pensando este pensamento. Não há sentido em me perguntar se estou certa que sou eu quem está pensando que está fazendo calor. O princípio da imunidade se aplica ao uso do ‘eu’ como sujeito, pois o uso do eu como sujeito não está disponível para identificação. Shoemaker chama esse tipo de pensamento de *I-thought*. Ele considera que certos *I-thoughts* – direcionados à mente e à vida mental e opostos ao corpo e à vida corporal<sup>25</sup> – são imunes ao erro por identificação. Pode haver um erro por identificação de um estado corporal, mas não de um estado mental. Quando o predicado é psicológico ou mental não é possível erro, pois a identificação é garantida pela

---

<sup>25</sup> A autoconsciência pode assumir diferentes versões de acordo com as definições de self (consciência do self como um corpo, consciência do self como uma mente, consciência do self como agente). Uma das teses atuais, defendidas por Q. Cassam (1998), é que não pode haver uma consciência da mente sem uma consciência do corpo. Sua contribuição não será investigada nessa tese.

forma como o eu constrói relatos sobre si mesmo. Não posso duvidar que sou eu quem está sendo tomado como referência do *I-thought*.

Existem duas versões do princípio da imunidade:

1. A infalibilidade da identificação, segundo a qual os *I-thoughts* não podem ser tornados falsos pelo erro de identificação porque as identificações que eles envolvem estão sempre e necessariamente corretas. Nessa acepção, os *I-thoughts* têm a mesma estrutura de composição que outros pensamentos, mas há algo de particular nos *I-thoughts*, que é o componente da identificação, que os torna imune ao erro. A identificação nunca falha.
2. A referência sem identificação, segundo a qual os *I-thoughts* não podem ser falseados pelo erro de identificação, pois não envolvem uma identificação. Nessa versão, os *I-thoughts* são estruturalmente diferentes, a eles faltam um componente de identificação, ou seja, são atos de referência que dispensam a identificação. Dito de outro modo, quando me auto-refiro, não desenvolvo um processo cognitivo no qual tento conectar minha experiência em primeira pessoa com algum critério conhecido por meio do qual avalio se a experiência me pertence ou não. O meu acesso a mim mesma na experiência da primeira pessoa é imediato e não-observacional e não envolve nenhum ato reflexivo ou perceptivo da consciência.

Shoemaker defende a segunda versão que postula a dispensabilidade da identificação. Ele argumenta que se toda autoconsciência requeresse auto-identificação, na forma como propõe a versão da infalibilidade da identificação, isso levaria a um regresso ao infinito. Pois, nessa versão, toda auto-referência requer uma auto-identificação. Então, toda vez que afirmo “estou triste”, meu pensamento tem a estrutura interna de “eu sou  $\emptyset$  e  $\emptyset$  está triste”. Isso acarretaria que para identificar quem é  $\emptyset$ , eu precisaria antes de um outro pensamento que identificasse a mim mesma com  $\emptyset$ , e assim por diante.

Para Shoemaker, a imunidade ao erro por identificação é uma característica<sup>26</sup> da criatura autoconsciente e não do estado autoconsciente, pois envolve essencialmente a consciência de si, e não apenas a consciência de um pensamento particular de alguém. É também uma peculiaridade *semântica*, e não epistêmica, central da autoconsciência, pois diz

---

<sup>26</sup> Existem alguns filósofos, como D. Rosenthal (Kriegel 2006) que tentaram se opor a esse argumento mostrando que existem alguns pensamentos autoconscientes que conduziram ao erro por falsa identificação, e que, portanto, a imunidade ao erro por identificação seria um aspecto e não uma particularidade da autoconsciência. Existe uma rica discussão a respeito da pertinência dessa tese e, se pertinente, a que estados de consciência essa característica se aplicaria. O mais importante para o meu argumento é que, de forma geral, a tese da imunidade ao erro por identificação tem sido considerada um aspecto semântico central da autoconsciência.

respeito a como o “eu” se conecta à referência. A especificidade da auto-referência lingüística deriva de aspectos correspondentes da autoconsciência e mais especificamente da auto-referência mental. Uma forma de entender a tese de Shoemaker, segundo Kriegel, é imaginar que a auto-referência não emprega nenhum estabilizador de referência como os designadores rígidos de Kripke (ver a discussão no Capítulo II). Ou seja, a auto-referência está livre do estabilizador de referência. Portanto, os *I-thoughts* se conectam ao sujeito sem que o sujeito realize a identificação, implicitamente ou explicitamente. Isso significa afirmar que a auto-referência é não apenas direta, mas inteiramente não mediada.

A tese de Shoemaker é que o componente da identificação está ausente. Quando afirmo “creio que eu tenho dor”, faço referência direta a mim mesma, e, portanto, não há um componente ao qual me identificar, pois não estou identificando a mim mesma com alguma coisa que reconheço, conheço, ou em cuja existência acredito, a qual o predicado da minha afirmação se aplica. Quando uso a palavra “eu” nesse contexto proposicional ela sempre se refere àquele que profere o ato de fala. O eu que acredita não pode ser distinto daquele que é tomado como tendo dor. Há um conjunto de estados conscientes em relação aos quais apareço para mim mesma como o único sujeito que está tendo esses estados. Estou consciente de mim mesma como eu mesma, sem que isso seja inferido de algum aspecto meu em particular. Shoemaker argumenta que faço referências que não são identificadoras de mim mesma. Sei coisas sobre mim, mesmo sem conhecer as propriedades que me identificam. Estou e devo estar consciente de mim mesma sem inferir isto de nenhuma coisa que conheço de mim mesma. Se devo saber certas coisas sobre mim mesma sem inferir isto de qualquer outra coisa que eu saiba sobre mim mesma, meu conhecimento de que é de mim mesma que estou consciente é independente de conhecer qualquer coisa sobre mim mesma (ver Brook 2005). Portanto, há uma referência a mim a partir de características e propriedades, mas não há auto-identificação. É por isso que Andrew Brook (2005) sugere que a chamemos de ‘auto-referência não-atributiva’, ao invés de ‘auto-referência sem identificação’.

Para Shoemaker, em relação às características da autoconsciência, não há acesso privilegiado e nem o conhecimento é infalível. Pois, posso, por exemplo, ruborizar diante de uma situação embaraçosa sem perceber o meu rubor. O que caracteriza a autoconsciência é a imunidade ao erro por identificação. A imunidade repousa no modo pelo qual a informação é obtida. Por exemplo, na afirmação “Estou com vergonha”, o modo de obter a informação é

imune ao erro por identificação, pois o próprio estado é auto-informativo. O estado “estar envergonhado” é auto-informativo, no sentido que traz a informação de que o eu se encontra nesse estado. É um conceito que adquire de mim mesma de forma independente de minhas propriedades objetivas<sup>27</sup>.

### 3.3 Autoconsciência e externalismo

De todos os aspectos analisados da autoconsciência e do autoconhecimento, podemos dizer que o que resta como critério definidor do conceito é o acesso epistêmico privilegiado aos nossos próprios estados mentais, na versão deflacionada proposta por Davidson (ver discussão no capítulo II), e a imunidade ao erro na consciência de si mesmo como sujeito proposta por Shoemaker. Examinemos agora quais os problemas que tanto a autoconsciência quanto o autoconhecimento podem causar numa abordagem externalista do mental.

O externalismo (ver discussão no capítulo II) é a visão de que algum elemento crucial no conteúdo de nossos estados representacionais é determinado por algo que é exterior a esses estados e exterior à pessoa que possui esses estados, portanto, não está localizado na pessoa. Esse elemento determinante pode ser a conexão causal com o contexto físico (para Putnam), com o contexto social (para Burge), com o contexto perceptivo (para Davidson) e o contexto biológico (para Dretske). Um dos problemas do externalismo é como compatibilizar o fato de que os estados psicológicos aparecem em nossa mente com a idéia de que são parcialmente determinados pelo contexto externo.

Uma das soluções apontadas pelo externalismo – e com isso Dretske e Davidson concordam – é que não aprendemos sobre o que nossos estados mentais representam ao olharmos para dentro de nós mesmos como as teorias internalistas da introspecção afirmam. A abordagem externalista propõe uma outra visão da introspecção. Para Dretske, o que conhecemos por introspecção são os fatos sobre a vida mental, ou seja, fatos mentais. Contudo, os objetos e eventos que percebemos no aprendizado desses fatos mentais não são

---

<sup>27</sup> Outro aspecto semântico da autoconsciência é a impossibilidade de exprimir certos pensamentos sem o emprego de um *I-thought*. Existem certos fatos que só podem ser expressos pelos *I-thoughts*. As crenças autolocalizadoras (“eu estou em frente à janela”) não podem evitar o emprego do dêitico, pois não há nenhum outro modo de pensar o mesmo pensamento a não ser empregando o vocabulário mental do *I-thought*, ou seja, há um vocabulário mental que aprendemos para expressar nossos estados mentais. O que caracteriza os pensamentos formulados pelo emprego do ‘eu’ como um dêitico essencial é que só podemos conhecer a referência se conhecemos o contexto do enunciado da fala. A referência é determinada pelo contexto do proferimento. (Kriegel 2006)

internos nem mentais. Os fatos mentais são construídos pelas relações que os eventos internos estabelecem com o mundo exterior. Se isso é verdadeiro, como é possível conhecer por introspecção aquilo que temos em mente?

Dretske tenta compatibilizar duas teses: 1) de que a experiência perceptual existe dentro da pessoa; 2) de que nada dentro da pessoa tem as propriedades das quais se tem consciência. Ele admite que existe um conjunto de conteúdos da mente e da consciência que permanecem estáveis mesmo quando alteramos nossos contextos de vida. Somos capazes de preservarmos nossas memórias, afetos e crenças em contextos muito diversos. Como isso é possível se o conteúdo de nossos estados mentais é, em parte, determinado por elementos externos à pessoa? E como podemos garantir a seguridade epistêmica do conteúdo de nossos estados mentais se ele é, em parte, determinado pelo mundo externo?

Como vimos na Teoria Representacional de Dretske, um estado consciente é um modo de representar alguma coisa. Os estados conscientes são simplesmente os estados com os quais nos tornamos conscientes das coisas. Esses estados são diáfanos, transparentes ao eu, no sentido de que permitem vermos através deles. Quando tenho uma experiência sensorial visual, por exemplo, eu vejo a propriedade do mundo fenomenal. Não obstante, se os estados conscientes são transparentes, nossa consciência deles deve ser deslocada para o que não é transparente, isto é, para o que os estados estão representando. Nesse sentido, a introspecção, entendida como um ‘deslocamento perceptivo’, não é um processo de olhar para dentro da mente.

Essa idéia em Dretske (1997) está subsumida pela tese da “percepção deslocada” (*displaced perception*). Na tese da percepção deslocada, a consciência de nossos estados mentais está deslocada para o que não é transparente, isto é, para o que os estados estão representando, ou seja, para a representação. Dretske defende a tese da introspecção como uma forma de ‘percepção deslocada’, que significa afirmar que o conhecimento de fatos internos se dá por meio da consciência (*awareness*) de objetos externos. O ‘deslocamento perceptivo’ ocorre quando o fato perceptivo está deslocado do objeto. Isto é, quando vejo que o objeto *k* é *F*, não por olhar para *k*, mas por olhar para o objeto *z*. Isso ocorre quando a representação conceitual do objeto não corresponde à representação sensorial. O exemplo de Dretske é que vejo o peso do meu corpo ao olhar para a balança e não para o meu corpo. O

deslocamento perceptivo requer uma crença que esteja numa conexão apropriada entre as propriedades do objeto percebido *z* e as propriedades do objeto alvo *k*.

Se o conhecimento introspectivo é o conhecimento direto da própria mente, pode se dizer que é o conhecimento direto acerca dos fatos sobre as aparências fenomenais. O conhecimento do sentido da experiência é o conhecimento de como as coisas estão sendo representadas pela pessoa na qual elas estão sendo representadas. Mas esse conhecimento não se dá de forma indireta e inferencial, é um conhecimento direto e imediato sobre os fatos representacionais de si mesmo. A introspecção é uma instância da percepção deslocada. É um processo pelo qual o sistema obtém informação sobre si mesmo, ao perceber, não a percepção, mas outra coisa. Existe uma experiência que é interna, mas não é olhando para dentro de nós mesmos que aprendemos o que nossos estados representam. Portanto, como afirma Dretske, a introspecção é um processo pelo qual as informações sobre coisas internas são obtidas pelo ato de representar objetos externos. O que aprendemos de nós mesmos são fatos constituídos externamente, ou seja, são fatos **sobre alguma coisa**.

Para Dretske, o conhecimento introspectivo é sempre conhecimento de alguma coisa acerca da mente, isto é, de fatos mentais. Levando em consideração que os fatos mentais são fatos representacionais, então o conhecimento introspectivo é uma representação conceitual de um pensamento acerca de alguma coisa ou uma representação da experiência de alguma coisa. Nesse sentido, conclui Dretske, é uma *metarepresentação*. As metarepresentações são representações de representações como representações. Para Dretske, é necessário ter crenças para ser capaz de meta-representar e de representar a si mesmo<sup>28</sup>. E essa é a fonte da autoridade da primeira pessoa.

A autoridade da primeira pessoa parece ser ameaçada quando concebemos os fatos mentais que são conhecidos pela introspecção como fatos constituídos por relações externas, históricas, existentes entre o que está na cabeça e o que está fora da cabeça. Contudo, a externalidade da mente é um problema apenas quando imaginamos que conhecemos os fatos extrínsecos olhando para dentro de nós mesmos e que é assim que experimentamos sobre o que são esses fatos. Como afirma Dretske:

---

<sup>28</sup> O modelo introspectivo de Dretske exclui as máquinas, os animais e os bebês da lista de seres autoconscientes capazes de exibir autoconhecimento. E mesmo os adultos não têm o conhecimento de alguns de seus estados representacionais.

“se compreendemos que a introspecção é um processo pelo qual a informação sobre as coisas *internas* é obtida no ato de representar (perceber) os objetos *externos*, então, o fato de que o que aprendemos (de forma introspectiva) sobre nós mesmos são fatos constituídos na externalidade deixa de ser um problema.” (1997: 54)

Toda mente representacional ao representar o mundo tem toda informação que necessita para saber como ela está representando o mundo. Contudo, alguma coisa mais do que a informação é necessária para se falar em autoconhecimento. Qual a informação que deve estar disponível à mente representacional para ser introspectada? Para Dretske toda informação fenomenal que precisamos é o acesso fenomenal ao que está sendo representado. Se sabemos o que estamos pensando e experimentando não sabemos isso por introspecção. Necessitamos de algum tipo de informação fenomenal sobre a representação. Temos acesso ao veículo da representação e não apenas ao que estamos representando.

Para Dretske, os elementos que determinam qual é a informação que a representação traz são externos à representação. Mesmo que os estados conscientes não tragam informação sobre o elemento de seu conteúdo que é externo, mesmo assim esse elemento pode ser conhecido pelo sistema. Mas, para ser conhecido, esse conteúdo deve estar acessível ao sistema do *modo* correto. A representação pode não trazer a informação sobre sua função, mas o sistema como um todo possui a informação. O sistema tem que ter a informação para saber qual é a informação que a representação é suposta trazer. Portanto, não é importante se a informação que o sistema representa vem do meio externo, o importante é que essa informação esteja acessível do *modo* correto.

Nesse sentido, o autoconhecimento não colocaria problema para o externalismo. Como argumenta Brook (2005), se a consciência fosse apenas representacional não existiria nenhuma ameaça ao externalismo, pois a consciência de nossos próprios estados mentais começaria com a consciência do elemento externo do conteúdo, então, não haveria nada não-representacional na mente. Para Brook, esse argumento só se sustenta na defesa do autoconhecimento dos estados mentais, não na autoconsciência da criatura, do sujeito de tais estados mentais.

Investiguemos agora o segundo aspecto da autoconsciência, a imunidade ao erro (ou auto-referência não-atributiva), tal qual a tese proposta por Shoemaker segundo a qual não me equivoco nunca que sou eu quem tem dor nas afirmações “Eu creio que eu tenho dor”. Se a

referência do “eu” é determinada, em parte, pelo ambiente externo, como posso ter certeza que não me equivoco na referência do eu?

O externalismo de Dretske oferece uma resposta que tenta compatibilizar autoconsciência e externalismo. A autoconsciência para Dretske é uma forma de auto-representação. O conceito de eu é uma função adquirida no desenvolvimento humano. Antes de adquirir a função de representar o eu, podemos dizer que o sistema traz a informação sobre os estados do eu e a informação sobre o sistema, mas não sobre o conteúdo representacional do conceito de ‘eu’. É o meio que atribui a noção de eu. Com a aquisição do conceito de eu, o estado adquire a função de informar o próprio sistema, adquire a função de se auto-representar como um conceito. A função que o eu já tinha de representar seus próprios estados se transforma em representação do eu. A vantagem dessa nova função é trazer uma dimensão objetiva do eu e do mundo, pois a autoconsciência permite a objetificação não somente do eu mais do mundo também.

Na auto-referência não-atributiva, como afirma Brook (2005):

“se eu sou, e para certos propósitos devo ser, consciente de mim mesmo como mim mesmo, sem estar inferindo isto de qualquer outra coisa que eu saiba de mim mesmo, então, o meu conhecimento de que é de mim mesmo que eu estou consciente é independente de eu saber qualquer outra coisa acerca de mim mesmo.” (Brook 2005)

Nesse caso, como vimos, ser eu mesma não é uma propriedade minha, não preciso conhecer nenhuma propriedade minha quando me refiro a mim mesma, ou seja, é uma referência que não atribui nada ao eu.

Como compatibilizar a auto-referência não-atributiva com o externalismo? De onde vem a consciência de que é você quem está representando? Só posso saber que sou eu que estou representando alguma coisa se realmente conheço a mim mesma como mim. A autoconsciência não pode ser explicada pela percepção deslocada. Como afirma Brook, ter a informação de que alguma coisa poderia ser suficiente para eu inferir que alguma coisa tem essa informação, isso não seria suficiente para saber que sou eu que tenho essa informação, pois a consciência de mim como mim mesmo não se dá pela consciência das propriedades de mim mesmo.

Para Dretske, a autoconsciência deve ser pensada como uma consciência factual, pois o eu tem consciência de fatos sobre o eu, de proposições sobre o eu: “Eu creio que eu sinto frio”.

Nessa proposição, o eu não é um objeto para o próprio eu. Portanto, não me relaciono com nenhuma entidade ‘eu’, me relaciono comigo mesma a partir de uma proposição, e não de um objeto mental no qual sou tematizado. O self não é objeto para o próprio eu quando a criatura tem consciência de si mesma. O que Baker e Dretske mostram, e, de certa forma, as teses empíricas sobre o desenvolvimento vão confirmar, é que não é possível ter autoconsciência antes da capacidade de conceitualização nem antes da aquisição do conceito de ‘eu’. O que temos antes são noções não-objetivas de si.

No entanto, se por um lado o relato representacionista de Dretske permite compatibilizar externalismo e autoconsciência, por outro, sua naturalização da autoconsciência parece deixar de fora aspectos centrais da perspectiva da primeira pessoa. A proposta naturalista de Dretske visa reduzir os conceitos a informações. Então, a capacidade de se referir ao eu e de possuir conceitos de ‘eu’, para ser uma função do sistema, precisa ser reduzida a termos mais básicos, como a informação conceitual do eu. Nessa acepção, ter autoconsciência é ter informação do eu.

Como mostra Baker, existe um aspecto da autoconsciência que resiste ao naturalismo robusto. A naturalização robusta, que tem em Dretske um dos seus mais proeminentes defensores, exclui a possibilidade de que a referência ao eu seja construída a partir de um conceito semântico ou psicológico. Baker concede que a autoconsciência seja concebida no relato do naturalismo pragmático (ver Ramberg 1997), que significa afirmar que o ‘eu’ é uma parte do mundo e que se refere à existência material encarnada da pessoa, e que a autoconsciência é uma habilidade adquirida por nossa espécie em sua história evolutiva. Mas para Baker, é impossível explicarmos a autoconsciência levando em consideração apenas os aspectos físicos ou informacionais da mente. A autoconsciência emerge da semântica partilhada, pois não há semântica sem interação externa.

Pois mesmo que o sistema adquira, na interação com o ambiente, a função de auto-representar-se, como poderia saber que a representação é dele e não de outra pessoa? Como poderia saber que é ele que está nesse estado?

Argumento aqui a favor da tese de Baker que defende o caráter **relacional** e **indispensável** da perspectiva da primeira pessoa. Segundo Baker, a perspectiva da primeira pessoa é algo que se constitui ao longo da vida de interação do indivíduo. A possibilidade de nos concebermos como um ‘eu’ só é possível na presença do conceito de outras coisas que

permite nos distinguirmos de coisas que são diferentes de nós mesmos, pois “não se pode ter conceitos de outras coisas sem a presença de outras coisas” (Baker 1998). (ver no capítulo 3 a elaboração de Winnicott e Costa sobre essa idéia). É somente na relação de oposição às coisas do mundo que podemos nos constituir como uma perspectiva de primeira pessoa. Portanto, a perspectiva da primeira pessoa depende das interações do indivíduo com coisas diferentes de si mesmo. Como afirma Baker, “x tem uma perspectiva de primeira pessoa somente se x está inserido (*embedded*) no mundo das coisas a partir das quais x pode se distinguir como si mesmo” (Baker 1998).

Baker rejeita a reificação cartesiana da interioridade, pois considera que o que é específico sobre ser uma pessoa não precisa ser assegurado por uma entidade privada lógica a qual ninguém mais tem acesso a não ser eu. O internalismo cartesiano pode ser evidenciado nas Meditações, onde não há nenhuma interrogação acerca de onde vêm os conceitos empíricos das coisas usados em seus pensamentos e nem o conceito requerido para afirmar “Estou certo que eu existo”. O cartesianismo contemporâneo defende a possibilidade de que a pessoa possa adquirir conceitos sem nenhuma interação com coisas diferentes dela mesma. Baker defende uma tese externalista, em oposição ao internalismo em filosofia da mente. Para ela, é implausível que os conceitos sejam inatos ou construídos de forma solipsista. É na interação que adquirimos conceitos e inclusive o conceito de ‘eu’. Quando imaginamos um único indivíduo no universo a fazer inferências sobre as coisas, não sabemos quais seriam os critérios que usaria para fazer uma inferência e não outra. Não há relato cartesiano que possa explicar como uma pessoa adquire um conceito, ou faz uma inferência, ou adquire o conceito de crença. Como mostrou Wittgenstein (*apud* Baker 1998), a aquisição de conceitos depende de fatos contingentes e de fatos não contingentes acerca do que significa ter um conceito empírico. E não há como compreendermos isso sem avaliarmos o contexto, pois os aspectos interativos do ambiente são parcialmente determinantes dos conceitos que possuímos.

A aquisição da perspectiva da primeira pessoa traz uma nova concepção de si. Baker mostra como essa distinção aparece gramaticalmente no uso do pronome ‘eu’ nas atribuições da referência à primeira pessoa nos discursos indiretos das orações subordinadas (*that-clause*). Os discursos indiretos são seguidos por verbos psicológicos “eu penso que eu ...”, “eu acredito que eu ...”, “eu gostaria que eu ...”, nos quais o segundo “eu” marca a distinção entre a primeira e a terceira pessoa, e o primeiro ‘eu’ marca a conceitualização dessa distinção. Como

afirma Baker, “a habilidade de atribuir a si mesmo a referência da primeira pessoa no discurso indireto é um sinal dos fenômenos da primeira pessoa forte.” (Baker 1998). O sujeito da perspectiva forte da primeira pessoa não só é capaz de ter pensamentos de primeira pessoa (‘Eu estou com fome’) como é capaz de se atribuir esse pensamento de primeira pessoa (‘**Eu\*** creio **eu** que estou com fome’), e de conceber a si mesmo como si mesmo por meio da habilidade lingüística de atribuir a si mesmo a referência em primeira pessoa ‘eu’.

A perspectiva da primeira pessoa não depende da linguagem natural, mas é, de forma geral, manifestada pelo uso de frases com o ‘eu-mentalês’. O discurso indireto expresso pelas frases com o ‘eu-mentalês’ difere do discurso direto das frases com ‘eu’. Afirmar “Eu creio que penso p” difere de afirmar “Eu penso p”. O “eu” do discurso direto pode ser eliminado. Os pensamentos que expressamos no discurso indireto, nas frases com ‘eu-mentalês’, não poderiam nunca ser expressos na ausência da perspectiva da primeira pessoa ou na perspectiva da terceira pessoa. O uso do ‘eu-mentalês’ indica exatamente perspectiva da primeira pessoa. O discurso indireto não pode ser expresso numa perspectiva não pessoal, na qual o eu seria substituído por um nome próprio ou um predicativo. Portanto, o discurso indireto do ‘eu-mentalês’ (‘eu creio que estou pensando p’) não pode ser reduzido a frases que dispensem a perspectiva da primeira pessoa, e não pode ser parafraseado por frases que dispensem o uso do ‘eu-mentalês’, como nas frases “Claudia crê que pensa p”, “A escritora dessa tese acredita que pensa p”. Continua Baker, “se atribuo a referência em primeira pessoa a mim mesmo, minha frase não pode ser adequadamente parafraseada por qualquer outra frase que fracasse em atribuir a referência em primeira pessoa a mim: a atribuição da referência em primeira pessoa a mim mesmo parece ser ineliminável.” (Baker 1998)

Outro modo pelo qual a perspectiva da primeira pessoa não pode ser dispensada é na explicação psicológica. Algumas explicações psicológicas exigem a perspectiva forte da primeira pessoa para que elas atribuam atitude ao agente. Por exemplo, o sentimento de culpa e a emoção da vergonha são sentimentos que não podem ser imaginados num agente que não se perceba como uma perspectiva de primeira pessoa. Por exemplo, o que motivou Édipo a cegar-se é a autoconsciência de ser ele a mesma pessoa que executou o parricídio, ou seja, a crença de ter sido o responsável pela ação (‘Eu creio que fui eu quem matou meu pai’). Não é possível que o agente exprima seu sentimento de culpa sem ser por meio das frases do ‘eu-mentalês’, pois são estados psicológicos que requerem a habilidade de pensar em si mesmo

como um self. O sentimento de culpa só pode ser motivado pela perspectiva forte da primeira pessoa. É necessário ter os recursos conceptuais necessários para ter pensamentos da perspectiva da primeira pessoa forte. Ou seja, ser capaz de pensar em si mesmo como um self.

Veremos no próximo capítulo que a solução ternária de Davidson permanece ainda a melhor forma de compreendermos como construímos os relatos de auto-referência e como surge a capacidade de autoconsciência. O modelo de Dretske e boa parte da psicologia cognitiva oferecem uma solução dual para o problema. Ou seja, estaria na relação organismo-ambiente, mente-mundo toda informação necessária para adquirir autoconsciência. A lógica ternária de Davidson – mente-mundo-intérprete – permite explicar o caráter relacional e indispensável da perspectiva da primeira pessoa forte de Baker.

## CAPÍTULO II

### EXTERNALISMO DO MENTAL

#### 1. O Debate Individualismo *Versus* Externalismo

Um dos grandes problemas atuais da filosofia e da psicologia é a idéia de *mente*. No debate sobre a vida mental, duas posições se destacam: o individualismo e o externalismo. A noção central neste debate é da **determinação mente-mundo**. Afirmar simplesmente que os estados mentais estão na cabeça, que não haveria outro lugar para eles estarem, parece ser um argumento trivial contra o externalismo. Do mesmo modo, dizer que existem muitos fatores ambientais que determinam os estados mentais do indivíduo parece também tornar banal o externalismo.

Tanto o externalista quanto o individualista reconhecem que o ambiente do indivíduo tem um papel causal em seus pensamentos, e conseqüentemente, na sua mente. Ambos consideram que o agente humano interage causalmente com o mundo obtendo informação sobre o mesmo pela percepção sensorial e pelas trocas comunicativas com os outros. Nessa interação, tanto o indivíduo quanto o ambiente sofrem modificações e um é causa da transformação do outro. O conteúdo da mente não é causalmente isolado do ambiente em que vivemos, pois o mundo participa causalmente do que percebemos, desejamos e pensamos. Não obstante, individualistas e externalistas discordam sobre o modo pelo qual a mente é determinada pela natureza do mundo. Para analisar a idéia de que a mente é em parte determinada pelo ambiente do agente é preciso explorar as condições sob as quais uma diferença no mundo implica uma diferença na mente.

Os individualistas defendem o que Robert Wilson (2004) chama de *determinação metafísica* da mente. O que ocorre *dentro da cabeça* do indivíduo determina metafisicamente a natureza dos estados mentais do indivíduo, pois tais estados mentais são autônomos em relação ao mundo externo. A mente é determinada pela própria constituição física do indivíduo. A relação de determinação metafísica é uma relação de *superveniência*, na qual as propriedades físicas do indivíduo determinam as propriedades mentais, ou seja, fixadas as propriedades físicas subvenientes, as propriedades mentais supervenientes estão também fixadas. Na relação de *superveniência*, se dois indivíduos são fisicamente idênticos em suas

estruturas, então, devem ter os mesmos estados psicológicos, pois não há diferença mental sem diferença física. Para os individualistas, uma diferença no mundo pode ocasionar uma mudança correspondente dentro da fronteira do indivíduo, caso afete suas propriedades.

Já os externalistas defendem o que Wilson (2004) chama de relação de *determinação causal*. Os externalistas não negam que os eventos mentais estão *dentro da cabeça*, mas consideram que o que está *fora da cabeça* é importante para a taxonomia psicológica, pois os tipos psicológicos que um indivíduo tem ou instancia não são metafisicamente determinados pelo que está *dentro da cabeça*. Para os externalistas, dois indivíduos idênticos em todos seus aspectos físicos podem ter estados mentais com conteúdos diferentes, ou seja, seus estados mentais podem variar mesmo que sejam fixados os estados físicos internos. Isso implica que as propriedades físicas não determinam os estados mentais do indivíduo.

Essas diferenças na determinação mente-mundo estão intimamente ligadas às diferentes formas com que ambas as démarches caracterizam o mental. Individualistas e externalistas discordam sobre como os estados psicológicos devem ser taxonomizados ou individuados. Isso acarreta diferenças no modo como concebem o tipo de propriedade que define os estados mentais. O que está em jogo é a controvérsia sobre qual seria a tese adequada sobre a individuação dos estados e eventos psicológicos presentes nas explicações do comportamento intencional, ou seja, a tese sobre como identificamos os ‘tipos’ de eventos mentais.

O individualismo afirma que os estados e eventos psicológicos a que nos referimos nas explicações causais do comportamento devem ser individuados de forma que apenas sejam consideradas as propriedades físicas internas do organismo. Para o individualismo psicológico<sup>29</sup>, as condições do ambiente externo do indivíduo podem ou não afetá-lo em seus estados internos. Se não afetam, não têm efeito sobre o comportamento do indivíduo e são irrelevantes do ponto de vista explicativo; se afetam, seus efeitos sobre o comportamento são imediatos, em função dos efeitos nos estados internos. O modo como as condições externas afetam os estados internos é parcialmente determinado pela economia dos estados internos do indivíduo. A mesma condição externa pode ter diferentes efeitos sobre os estados internos de diferentes indivíduos ou do mesmo indivíduo em momentos diferentes. Não são as condições

---

<sup>29</sup> Para uma defesa da abordagem individualista em psicologia ver David J. Buller (1997).

externas que explicam o comportamento, mas os estados internos por elas afetados. Mesmo que as condições externas afetem os estados internos, a referência a essas condições não oferece informações sobre a explicação. A explicação causal do comportamento faz referência apenas aos estados que sobrevivem aos estados físicos internos do indivíduo.

Já a démarche externalista apresenta uma visão pluralista da taxonomia psicológica. Ela percebe a taxonomia e a explicação científica como sensíveis a uma série de fatores relacionados às propriedades históricas e relacionais. Seria mesmo possível a aceitação do lema “nenhuma diferença psicológica sem uma diferença física”, porque as diferenças físicas também repousam além das fronteiras do indivíduo. Nesse sentido, para conhecermos o conteúdo de nossos estados mentais deveríamos investigar nosso ambiente. Para o externalismo, o conteúdo mental (das crenças e de outras condições psicológicas que dependem da linguagem) é, ao menos em parte, dependente da determinação da referência no contexto particular das palavras usadas nas sentenças ou na expressão da crença. Assim sendo, a referência depende de fatores que são externos ao corpo e ao cérebro do falante.

O que interessa, para os propósitos de investigação do *autoconhecimento e da constituição do eu*, é que a defesa da tese taxonômica individualista ou externalista dos conteúdos mentais acarreta modificações consideráveis na forma como concebemos o mental e os processos introspectivos.

O individualismo estuda os fenômenos mentais a partir do que acontece com a pessoa e como ela responde ao seu ambiente físico sem referência ao que acontece no contexto social em que está situada. Isso nos leva a considerar que aquilo que uma pessoa pensa depende do que ocorre ou aparece na sua mente, ou melhor, depende de suas disposições internas. O individualismo postula a autoridade do conhecimento da primeira pessoa, de seus próprios estados mentais, como um conhecimento *a priori*, que para obtê-lo não precisaríamos investigar o ambiente ou conhecer o mundo externo.

Na perspectiva externalista, acredita-se que se um pensamento é identificado pela relação entre o indivíduo e alguma coisa fora da sua cabeça, então, ele não está completamente na minha cabeça. Se ele não está completamente na minha cabeça, ele não pode ser apreendido pela minha mente no modo que é requerido pela autoridade da primeira pessoa. Portanto, ao adotarmos o externalismo, parece que não podemos conhecer os conteúdos de nossos estados mentais simplesmente pela introspecção. Pois se os conteúdos mentais são

determinados pelo ambiente, então, a investigação introspectiva não pode garantir a apreensão desses conteúdos. Isso parece significar, em alguma medida, uma renúncia ao autoconhecimento como um conhecimento *a priori* de nossos estados mentais.

Nesse capítulo, apresentarei as teses individualistas da mente de inspiração cartesiana e argumentarei contra essas teses a partir das críticas antiindividualistas que indicam que os estados intencionais não satisfazem as exigências taxonômicas do individualismo. Em seguida, traçarei a história conceitual endógena das teses externalistas e de seus modelos, restringindo-me às implicações dessas teses no que concerne ao campo da filosofia da mente, em particular à questão do autoconhecimento.

## **2. Individualismo**

### **2.1 Descartes e o Individualismo**

A filosofia da mente cartesiana afirma que a natureza dos estados mentais é essencialmente determinada por propriedades e características ‘internas’ à mente individual. Essa forma de conceber a mente foi denominada de *individualismo* por Tyler Burge (1988 [1979]). O individualismo em psicologia defende a tese que afirma que os estados psicológicos deveriam ser construídos sem referência a qualquer coisa além da fronteira do indivíduo (ver Wilson 2004). Burge identifica o individualismo com uma série de concepções da mente que prevalecem na filosofia moderna desde Descartes, a partir da segunda metade do séc XVII. Dentre elas, as concepções que interessam investigar são o dualismo cartesiano e o cartesianismo materialista das teorias da identidade.

O *locus classicus* do individualismo da obra de Descartes se expressa na sua constatação de que a natureza e o conteúdo dos estados mentais permanecem inalterados e estáveis apesar das variações no mundo dos objetos aos quais esses estados se referem. Haveria uma inadequação entre os conteúdos internos e o mundo externo que tornariam falsas as convicções ou percepções. Portanto, mente e mundo não co-variam e não co-dependem. Não somente os conteúdos de nossa mente se mantêm estáveis, independentemente de suas ‘relações exteriores’ variáveis, como também possuímos um acesso privilegiado a esses conteúdos. Conhecemos o conteúdo de nossos pensamentos com a certeza e autoridade da primeira pessoa, ou seja, de forma independente do mundo, sem investigar o ambiente. Vejamos essas idéias nas palavras de Descartes:

“fecharei os olhos, taparei os ouvidos, distrairei todos os meus sentidos, até apagarei de meu pensamento todas as imagens das coisas corporais, ou, pelo menos, porque é difícil fazer isso, eu as reputarei como vãs e falsas; e, assim, entretendo-me somente comigo mesmo, e considerando meu interior, tratarei de tornar-me pouco a pouco mais conhecido e mais familiar a mim mesmo. Sou uma coisa que pensa, ou seja, que duvida, que afirma, que nega, que conhece poucas coisas, que ignora muitas, que ama, que odeia, que quer, que não quer, que imagina também e que sente. Pois, assim como observei anteriormente, embora as coisas que sinto e imagino talvez não sejam absolutamente nada fora de mim e em si mesmas, estou certo, não obstante, de que essas maneiras de pensar a que chamo de sentimentos e imaginações, somente na medida em que são maneiras de pensar, residem e se encontram certamente em mim. E nesse pouco que acabo de dizer, creio ter relatado tudo o que sei verdadeiramente, ou pelo menos tudo o que até agora notei que sabia.” (Descartes 2000: 57)

Na filosofia cartesiana estão implícitas duas premissas, conforme explicitadas por Wilson Mendonça (1997a): (1) *os estados mentais são imediatamente acessíveis ao sujeito de tais estados mentais*; (2) *os estados mentais são radicalmente independentes do contexto natural e social da mente*.

A primeira premissa postula o acesso privilegiado do sujeito ao conteúdo de seus estados mentais. Como vimos, para Descartes, o autoconhecimento é incorrigível, infalível e seguro, pois as propriedades e características determinantes da natureza e dos conteúdos dos estados mentais são imediatamente acessíveis ao sujeito em tais estados mentais. O critério usado por Descartes para definir o mental é a *indubitabilidade*. Este critério se apóia no fato de que as dores, os pensamentos e as convicções são de tal ordem que o sujeito não pode duvidar que as tem, ao passo que é possível duvidar de tudo que é físico (ver Rorty 1988). Os objetos mentais, como as sensações brutas, seriam entidades privadas naturalmente dadas das quais teríamos um conhecimento introspectivo e incorrigível. A tese sobre esse fundamento baseia-se na presunção de que estamos certos acerca do conteúdo de nossas mentes, o que não aconteceria com nossas crenças sobre o mundo e outras mentes. Pois seria a partir do conhecimento de nós mesmos que poderíamos aceder à realidade exterior.

Dado o impacto dessa premissa no discurso da primeira pessoa, ela foi analisada separadamente no capítulo precedente, dedicado à discussão das noções de autoconsciência e autoconhecimento e os problemas levantados por essa premissa nas abordagens externalistas.

A segunda premissa cartesiana postula que os estados mentais são independentes do contexto natural e social da mente. A filosofia cartesiana advoga um tipo de realismo indireto, no qual o mundo existe independentemente do sujeito que só tem acesso a ele por

intermediários entre a mente e o mundo<sup>30</sup>. Não há contato direto entre a mente e o mundo externo. Não percebemos a realidade imediata, mas as qualidades secundárias<sup>31</sup> tal como aparecem às nossas mentes. Descartes concebe os objetos da percepção como internos. A imagem mental representa e faz a tradução entre duas substâncias: *res cogitans* e *res extensa*. Tudo o que conhecemos do mundo exterior depende do que observamos a partir das representações interiores.

Este modelo cartesiano está baseado na teoria dos *sense data*. Segundo essa teoria, há um mundo de objetos reais independentes da mente que causa nossas experiências. Porém, não percebemos diretamente esses objetos externos. O que percebemos são os efeitos desses objetos sobre nós. É o objeto mental, denominado *sense datum*, que é apreendido diretamente. Essas experiências sensoriais ocorrem *dentro da cabeça* e refletem nossa capacidade de sermos afetados passivamente de diferentes formas pelo objeto. Como mostra Hilary Putnam (1999), nesse modelo perceptual, a percepção envolveria sempre uma *interface* imaterial *dentro da cabeça* – impressões, idéias, representações – uma ponte entre a mente e os objetos externos percebidos. A mente seria o *teatro interno ou a tela de cinema interior*, como chama Wilfrid Sellars (1997)<sup>32</sup>, na qual estariam projetadas as imagens mentais. Nessa concepção, o *qualia*<sup>33</sup> não é uma propriedade dos objetos externos, mas uma propriedade das imagens na *tela mental interna*. Não percebemos diretamente as propriedades dos objetos externos, temos consciência direta apenas das ‘imagens mentais’. Os objetos que vemos são apenas suas representações mentais. As representações internas são supostas estarem relacionadas a pessoas e objetos materiais apenas como efeitos internos de causas externas.

Esse modelo da mente deu origem ao que Donald Davidson (2001 [1988]) chama de *mito do subjetivo*, no qual os estados mentais são entidades privadas ‘diante da mente’. Para ele, essa forma de conceber a relação entre a mente e a natureza reproduz a dicotomia

---

<sup>30</sup> Desde Descartes, a concepção das relações entre a subjetividade e o mundo objetivo está baseada na postulação de entidades mediadoras entre o sujeito e o mundo. A postulação de intermediários epistêmicos pressupõe a possibilidade de distinguir entre conceitos e um conteúdo não contaminado por eles. Ver Rorty (1988), Costa (1994), Evnine (1991), Ferreira (2000).

<sup>31</sup> Na metafísica tradicional, haveria uma dicotomia entre as ‘qualidades primárias’ que representam os modos como as coisas reais são nelas mesmas, e as ‘qualidades secundárias’, que representam o modo como as propriedades do mundo real aparecem na experiência fenomenal. As qualidades secundárias teriam propriedades relacionais como disposições que afetam nossa mente ou cérebro. Ver Putnam (1999).

<sup>32</sup> Para uma crítica da noção da mente como *tela de cinema interior* ver Sellars (1997).

<sup>33</sup> O primeiro filósofo a usar o termo ‘qualia’ foi D. Lewis como um termo derivado do latim para ‘qualidades’. O *qualia* se refere às qualidades mentais que estamos conscientes na experiência fenomenal. Para maiores detalhes da história e do uso da noção em filosofia da mente ver “Are appearances ‘qualia?’” em Putnam (1999).

esquema-conteúdo. Haveria uma divisão entre a experiência não-interpretada e o esquema conceitual que a organizaria. A teoria dos *sense data* pressupõe a divisão da experiência cognitiva em dois tipos<sup>34</sup>: *os conteúdos*, que são os objetos imediatos (como os *sense data*, as impressões, sensações, ou aparências) dados à consciência e não categorizados; e *os esquemas*, que são as formas, construções e interpretações que representam a atividade do pensamento. Segundo essa perspectiva, haveria diferentes esquemas lingüísticos que forneceriam conceitos variados cuja tarefa seria organizar a experiência sensorial em objetos, eventos e estados. Os conteúdos não-interpretados da experiência seriam invariantes. Essa idéia emerge da imagem da mente como um observador passivo que avalia o espetáculo do teatro interno.

A segunda premissa, que postula a independência contextual dos estados mentais, acarreta dois problemas: *o solipsismo perceptivo e a noção de 'estado interno'*. Começo pela crítica ao solipsismo.

O solipsismo perceptivo é característico do cartesianismo e do empirismo que tomam como ponto de partida o indivíduo e sua sensorialidade, entendidos como eventos isolados e irreproduzíveis. A mente solipsista é concebida como algo que emerge da experiência produzida pelos sentidos ou como o resultado de estimulações neuronais. A mente cartesiana é solipsista. Nela, os estados psicológicos pressupõem apenas a existência do indivíduo que é o sujeito do próprio estado mental conhecido. Nada existe a não ser a própria pessoa e os conteúdos de sua mente.

A analogia entre a forma como percebemos os objetos e o modo como adquirimos nossas crenças deu origem ao modelo perceptual do conhecimento. Esse modelo fundamenta-se na idéia de que o processo cognitivo começaria com o autoconhecimento de nossas sensações e pensamentos, depois se desenvolveria na direção do conhecimento do mundo

---

<sup>34</sup> Wilfrid Sellars (1997) critica a concepção de mente cartesiana que resulta na distinção entre conhecimento inferencial e conhecimento não-inferencial. O 'mito do dado' baseia-se na idéia de que as sensações cruas – os *sense data* – teriam o estatuto de representações privilegiadas sobre as quais se fundaria o conhecimento inferencial baseado em crenças, ou seja, as sensações cruas seriam 'objetos diante da mente' dos quais teríamos uma consciência imediata e direta e que essa consciência acarretaria o conhecimento empírico desses objetos. Sellars mostra como a mente cartesiana pressupõe que todo estado mental é um estado, incorrigível e transparente, conhecido imediatamente como algo que nos é dado. O 'mito do dado' é a idéia de que haveria um tipo de consciência com duas propriedades: as sensações cruas e as crenças justificadas. E as sensações cruas teriam o status de representações privilegiadas sob as quais se fundamentaria o conhecimento inferencial. Sellars propõe o abandono da distinção dado-versus-não-dado e da idéia de que há uma questão de relação especial entre idéias (ou palavras) e objetos. Para uma discussão mais detalhada sobre o empirismo e o 'mito do dado' ver Rorty (1988), Cavell (1993).

externo objetivo, e, por último, dirigir-se-ia para o conhecimento de outras mentes. Ou seja, o conhecimento progrediria do subjetivo para o objetivo<sup>35</sup>. Esse solipsismo resultou, como salientam Putnam (1999) e Davidson (2001 [1988]), na dúvida quanto à existência do mundo externo e de outras mentes. Pois se nossas convicções, expectativas e percepções fossem construídas sem referência a qualquer coisa além das fronteiras do indivíduo, como poderíamos ter acesso ao mundo externo?

Um dos primeiros a discutir esse problema foi Wittgenstein. Seu conhecido ‘argumento da linguagem privada’ tornou-se exemplar para a compreensão da questão. O caráter público e interpessoal da linguagem garante que frequentemente possamos atribuir mente aos outros de forma correta por sabermos, na maior parte das vezes, o que eles pensam. Portanto, o acesso a outras mentes é dado pela linguagem. A evidência que usamos no caso do conhecimento de outras mentes é de caráter público. Como afirma Davidson, “os pensamentos são privados no sentido em que pertencem à pessoa, porém, os pensamentos fazem necessariamente parte de um mundo público” (2001 [1987]: 23). Se as palavras e os pensamentos são acerca de tipos de objetos e eventos, não há espaço para a dúvida cartesiana sobre a existência independente de tais objetos e eventos.

O que Davidson (2001 [1988]) procura mostrar é que o fato de que possamos nos enganar às vezes sobre a natureza do mundo, não significa que nos enganemos sempre. Ao endossar o caráter essencialmente público da linguagem, Davidson chama atenção para a

---

<sup>35</sup> O solipsismo perceptivo é característico do empirismo que defende a idéia de que podemos construir um mundo tal qual ele se apresenta aos sentidos perceptivos. O empirismo pressupõe que a mente organiza a realidade dos sentidos segundo certos esquemas conceituais. Esse seria o terceiro dogma do empirismo criticado por Davidson (*apud* Evnine 1991) ao abordar a obra de Quine. Como mostrei em outro trabalho (Ferreira 2000), “a tese dualista do empirismo ainda persiste na concepção filosófica naturalizada de Quine. Para Quine, a impressionante riqueza do nosso aparato conceitual tem sua origem em determinadas excitações das superfícies dos órgãos sensoriais, e a tarefa da epistemologia consiste em explicar o processo pelo qual se chega das excitações aos conceitos e teorias. As contribuições sensoriais dão origem ao esquema conceitual, e o conteúdo sensorial se distribui vagamente através da estrutura de conceitos e juízos, sendo difícil estabelecer uma distinção clara entre: enunciados analíticos, cuja verdade dependeria unicamente do significado dos termos empregados neles; e enunciados sintéticos, cuja verdade dependeria de sua confrontação com a experiência sensorial. É o sistema conceitual como um todo que se confronta com o tribunal da experiência. Então, diante da questão de como chegamos à verdade dos significados de uma sentença, ou melhor, de como uma teoria do significado das sentenças de uma linguagem pode ser empiricamente testada, Quine propõe como resposta sua teoria da tradução radical. De forma resumida, para Quine, a tradução depende do conjunto de estimulações sensoriais que concordam ou discordam da sentença dada, ou seja, as estimulações são tratadas como evidência para essas sentenças. Dessa forma, Quine tenta isolar o conteúdo das sentenças sobre o mundo introduzindo uma mediação sensorial. Ao invocar a necessidade de um nível de estimulação sensorial agindo como intermediário epistemológico entre o mundo e a verdade das sentenças, Quine garante a manutenção do dualismo esquema-conteúdo.” Ferreira (2001: 73). Sobre a obra de Wilard Quine ver Coutinho (1994), Costa (1994), Evnine (1991).

importância da natureza externalista de nossas crenças. Os significados das frases são dados pelas situações que nos levam a manter nossas crenças como verdadeiras ou falsas. E isso faz parte da própria definição do que é ter uma crença. *Possuir uma crença é enunciar frases que consideramos verdadeiras ou falsas.*

A imagem solipsista da mente desconsidera que qualquer conhecimento proposicional, seja de dentro ou de fora, exija a posse do conceito de verdade objetiva. A tese de Davidson é que o conceito de *verdade*<sup>36</sup>, fundamental para aquisição de conteúdos mentais intencionais, é acessível somente a criaturas em comunicação com outras. Nesse sentido, o conhecimento de outras mentes é imprescindível aos outros tipos de conhecimento. Portanto, como veremos ao abordar o externalismo de Davidson, o conhecimento emerge de forma holística e interpessoal desde o primeiro momento e a mente faz parte com o mundo de um único esquema conceitual de matriz e desenvolvimento *intersubjetivo*.

Wittgenstein, como Davidson, também mostrou que ‘nosso conhecimento de outras mentes’ é um aspecto essencial de nosso uso de predicados mentais que aplicamos a outros com base na evidência do comportamento. Contudo, sua resposta é insatisfatória para o problema do autoconhecimento. Para ele, o conhecimento de nossa própria mente se basearia também na evidência ou observação. O sujeito trataria seus próprios pensamentos da mesma forma como trata os dos outros. Davidson (2001 [1987]) mostra como essa tese, apesar de evitar o solipsismo cartesiano e empirista, não explica a assimetria observada entre o conhecimento de si mesmo e o conhecimento do outro. Veremos, ao argumentar a favor da conciliação do externalismo com o autoconhecimento, como a concepção intersubjetiva da mente de Davidson oferece uma explicação para essa assimetria observada por toda tradição filosófica desde Descartes.

## 2.2 Superveniência e Relevância Causal

Outra consequência da premissa individualista cartesiana é a postulação de *estados internos*. Segundo Sellars (1997)<sup>37</sup> e Putnam (1999), foi com a noção de *estado psicológico interno* que começamos a ver o retrato cartesiano da mente como *teatro interno*. A visão

---

<sup>36</sup> Para uma discussão detalhada sobre a noção de verdade adotada pelo pragmatismo de Davidson, ver Costa (1994, 1995), Eynine (1991).

<sup>37</sup> Não apresentarei as críticas de Sellars. Para uma discussão sobre a noção de *episódio interno* no cartesianismo ver Sellars (1997) Rorty (1988).

cartesiana trata os estados mentais como estados da substância imaterial e os torna propriedade de algo que está dentro da mente. A idéia de que existem estados internos foi amplamente adotada pelas concepções pós-cartesianas do mental. As teorias materialistas, que surgiram em resposta às críticas ao dualismo cartesiano<sup>38</sup>, acabaram por reproduzir essa noção. O cartesianismo materialista, ao reduzir o mental a um evento físico cerebral e não mais a um efeito imaterial do cérebro, evita o problema de explicar como o evento cerebral pode causar o evento imaterial na mente. Não obstante, mantém a noção internalista do mental. Portanto, tanto a concepção imaterial da mente quanto a fisicalista derivam da mesma concepção individualista da mente. Na aceção individualista, os eventos psicológicos são *estados internos* que estão presentes aqui e agora no espaço psicológico do indivíduo. Como mostra Wilson (2004), muitos filósofos foram atraídos pelo individualismo devido à possibilidade de compatibilizar a noção de *estados psicológicos internos* e o fisicalismo.

O cartesianismo materialista é a combinação do fisicalismo com os *estados internos* cartesianos. Os estados mentais internos, antes imateriais, passam a ser concebidos como materiais. Na ontologia fisicalista, as explicações físicas são a explicação última para os fenômenos. Não há nenhuma diferença psicológica sem uma diferença correspondente dos estados físicos do indivíduo. A mente é concebida como um órgão, e as sensações e pensamentos como processos cerebrais.

O cartesianismo materialista emprega a mesma filosofia da percepção baseada na teoria dos *sense data* para compreender o funcionamento cerebral. Apesar da teoria dos *sense data* ter caído em desuso, suas premissas solipsistas estão subjacentes às teses perceptivas das teorias materialistas, por meio da idéia de que existe um centro de experiências, de ‘impressões’ – as chamadas ‘representações mentais’, no caso das ciências cognitivas – localizado no cérebro. O cerne da idéia é que haveria um conjunto de experiências sensoriais, ocorrendo dentro do cérebro, que interconectaria a mente e o mundo externo. Nesse modelo, o cérebro processa as representações mentais que são os outputs dos processos perceptivos. Estas representações estão conectadas aos objetos do ambiente apenas causalmente e não cognitivamente. Ou seja, não temos acesso às propriedades dos objetos, elas são traduzidas em

---

<sup>38</sup> A filosofia da mente tem repudiado a tese cartesiana do dualismo mente-corpo. Concordo inteiramente com as críticas ao dualismo, mas não discutirei o problema mente-corpo. Estou interessada em investigar a concepção individualista cartesiana do mental e suas conseqüências para o externalismo do conteúdo mental. Encontramos uma excelente sistematização das teorias corpo-mente em Kim (1998) e Fissette & Poirier (2000).

experiências qualitativas – o *qualia* – que são identificadas a eventos físicos. Portanto, as propriedades dos objetos externos não fazem parte do processo cognitivo, elas são traduzidas em representações mentais internas num centro que processa as informações a partir do qual resultam os processos cognitivos. Isso significa que o ambiente externo tem um papel causal deflagrador, incitando o processo perceptivo, mas não participa do processo cognitivo. Como afirma Putnam:

“os inputs perceptivos são considerados o limite exterior de nosso processo cognitivo, tudo que repousa além desses inputs está conectado aos nossos processos mentais apenas de forma causal e não cognitiva. Mesmo se os ‘inputs’ são colocados tão longe quanto na superfície de nosso corpo (como Quine faz com seu relato de ‘neurônios de superfície’ e ‘irritações de superfície’), tudo que estiver do lado de fora de nossa pele está também do lado de fora de nosso processo cognitivo”. (1999: 16).

Nesse sentido, as experiências sensoriais *qua* processos cerebrais também seriam como *interfaces* entre nós e o mundo por meio das quais teríamos acesso ao mundo externo.

Segundo Putnam, existem duas teorias que identificam os estados psicológicos internos a processos cerebrais: a *identidade teorética*, baseada na identidade entre teorias; e a *identidade anômala*, baseada no monismo anômalo de Davidson<sup>39</sup>. Considerarei somente a primeira, pois a teoria de Davidson será abordada em conjunto com as abordagens externalistas. A identidade teorética é peculiar às teorias materialistas. As posições materialistas, ao identificarem a mente ao cérebro, operam uma redução na qual cada propriedade mental interna depende da propriedade física<sup>40</sup>. As teorias da identidade físico-mental defendem, portanto, a tese individualista de que os estados psicológicos devem ser individuados de forma que sejam consideradas somente as propriedades físicas internas do

---

<sup>39</sup> A teoria do monismo anômalo deriva de três premissas: “(1) **Os eventos mentais são causalmente relacionados com os eventos físicos.** Há uma *interação causal* entre os eventos físicos e mentais, independente da forma como são descritos – um evento físico pode causar um evento mental, e vice-versa. (2) **As relações causais singulares instanciam leis estritas.** Ou seja, se dois eventos estão causalmente relacionados eles instanciam leis deterministas, e, portanto, podem ser previsíveis, controláveis. (3) **Não há leis estritas psicofísicas.** Ou seja, não é possível estabelecer leis de conexão entre o físico e o mental, inviabilizando qualquer projeto reducionista fiscalista do mental. É impossível a criação das chamadas *‘bridging laws’*, que são leis que conectam duas propriedades expressas por “para cada propriedade mental *m* do evento *x*, segue-se sempre a propriedade física *f* do mesmo evento *x*”. (Ferreira 2000: 22).

<sup>40</sup> Essas teorias são chamadas teorias da identidade *type-type* pois estabelecem a identidade entre tipos de eventos. Elas são definidas por duas teses: (1) todo evento mental é idêntico a um evento físico no cérebro; (2) a identidade é estabelecida pela identidade entre *tipos* de eventos, ou seja, um tipo de evento mental corresponde a tipos de eventos cerebrais caracterizados fisicamente. (Ver Dennett 1986, 1996)

organismo. O individualismo das teses materialistas se apóia em dois argumentos: *superveniência* e *causalidade*<sup>41</sup>.

Para ilustrar a tese da superveniência, vou tomar o funcionalismo de Jaegwon Kim, que é uma das teorias mais bem formuladas em defesa do individualismo, posto que Kim leva em consideração várias das críticas anti-individualistas já formuladas.

A teoria funcionalista de Kim é a conjugação de três teses (ver Putnam 1999):

1. *A tese da superveniência*<sup>42</sup>, que afirma que todo estado psicológico interno de um organismo é superveniente a seu estado físico interno sincrônico.
2. *A tese da dependência*, que afirma que toda propriedade mental depende da sua propriedade física correspondente.
3. *A tese explicativa*, segundo a qual os estados psicológicos internos são os únicos estados que a teoria psicológica precisa invocar ao explicar o comportamento humano.

Kim defende uma versão forte da tese da superveniência, na qual as propriedades psicológicas internas sobrevêm às propriedades físicas específicas. Esta tese é a adaptação a uma classe particular de estados psicológicos da tese clássica da correlação psicofísica<sup>43</sup> que postula a identidade entre o físico e o mental. Ele acredita que essa correlação só pode ser aplicada a estados que são *estritamente internos*. Esses estados estritamente internos são chamados ‘estados psicológicos no sentido estreito’ (*narrow content*). Kim admite a existência de estados psicológicos não-internos (*wide content*) desde que possam ser fracionados em um *estado interno* e um conjunto de relações históricas e cognitivas externas. Assim, a relação de superveniência se mantém para a parte interna de tais estados, ou seja, para os estados de *conteúdo estreito*.

---

<sup>41</sup> Cf. Mendonça (1997b).

<sup>42</sup> Um dos problemas do fisicalismo, que identifica o mental com estados cerebrais, é encontrar um lugar para a mente num mundo fundamentalmente físico. O problema mente-corpo vem tentando encontrar respostas filosóficas a essa questão. Duas noções têm sido invocadas para descrever a relação entre mente e corpo: superveniência e realização (ver nota 22). A noção de **superveniência** é a relação entre dois conjuntos de propriedades – as propriedades supervenientes e as propriedades de base – que pode ser assim definida: “as propriedades mentais sobrevêm às propriedades físicas, no sentido que necessariamente duas coisas (no mesmo ou diferente mundo possível) indiscerníveis em todas as suas propriedades físicas são indiscerníveis em suas características mentais”. (Kim 1998: 10).

<sup>43</sup> A Tese da Correlação Psicofísica pode ser assim definida: para cada evento psicológico M há um evento P tal que, como se fosse uma lei, um evento de um tipo M ocorre no organismo num momento se e somente se um evento do tipo P ocorre a ele no mesmo momento (Putnam 1995).

No funcionalismo, as entidades são definidas por suas funções, por aquilo que são capazes de realizar. Os estados psicológicos são estados múltiplemente realizados<sup>44</sup>, que podem ter diversas realizações (implementações) físicas ou biológicas em diferentes espécies e diferentes estruturas (ver Kim 1998; Putnam 1992). Os estados internos são estados psicológicos concebidos como *estados funcionais*, postulados para explicar o comportamento, cuja função é fazer a mediação entre os inputs sensoriais e os outputs comportamentais de modo apropriado. (Grosso modo, os inputs são identificados em termos das irritações das superfícies do organismo e os outputs em termos dos movimentos de seu corpo). Estar num estado psicológico é estar num estado com uma função causal específica. Isso significa afirmar que para cada um desses estados funcionais deve haver um estado físico com o mesmo papel explicativo. O estado psicológico é realizado por uma condição física particular, a mesma condição na qual o organismo se encontra toda vez que estiver nesse estado. Porém, não é necessário conhecer a condição física para identificar o estado psicológico. Para identificá-lo é necessário somente conhecer a função desempenhada, cujo conhecimento é originado das definições dadas por uma teoria psicológica em particular. A concepção funcionalista, portanto, pressupõe a existência de uma teoria psicológica que trate as condições psicológicas como entidades a serem identificadas a estados físicos do organismo.

Em geral, o funcionalismo sustenta que os estados psicológicos dos quais tratamos na psicologia cotidiana (*folk psychology* ou *psicologia vernacular* como quer Putnam) são estados funcionais idênticos aos estados fisiológicos identificados por meio de seus papéis causais explicativos. Isso significa tratar a relação entre os estados psicológicos do mesmo modo que são tratadas as relações causais entre os estados físicos determinadas por leis estritas, o que acarretaria numa teoria psicológica cuja *raison d'être* seria gerar explicações e predições causais nomológicas. O *caráter nomológico da causalidade* significa que as

---

<sup>44</sup> Opto pela tradução, corrente na literatura filosófica, da noção de ‘multiple realization’ por ‘realização múltipla’. Um dos sentidos em inglês para “to realize” é “to realize something is to cause it to be real or to exist or happen in fact”, e “realization” seria o ato derivado desse verbo. Isso me parece semelhante com o sentido em português de realização como o “ato de tornar-se realidade no tempo e/ou no espaço, ocorrer, ter lugar”, ou seja, “realizar” no sentido de “implementar” que significa “pôr em execução, pôr em prática (plano, programa ou projeto)”. A palavra “realization”, portanto, usada na literatura filosófica, tem o sentido de “implementação” ou “realização”. Traduzo-a como *realização*, pois me parece mais adequado para exprimir a novidade do conceito. A noção de *realização*, introduzida por filósofos materialistas da filosofia da mente, tem sido muito usada para falar sobre “mental states as being realized by physical states of the brain” e tornou-se um conceito amplamente invocado pelos autores. Na relação de *realização*, os estados físicos intrínsecos do indivíduo, de seu sistema nervoso central, são *realizações* físicas dos estados mentais do indivíduo, e esses *realizadores* são considerados suficientes para garantir a presença dos estados que eles realizam (ver Putnam 1999, Wilson 2004).

generalizações podem, em princípio, ser expressas em leis estritas. Ou seja, se dois eventos estão causalmente relacionados eles instanciam leis deterministas, e, portanto, podem ser previsíveis, controláveis.

Kim propõe abandonar essa teoria psicológica, porém, sugere que seja mantido o papel normativo da psicologia cotidiana nas avaliações das ações intencionais. Ele imagina poder compatibilizar sua concepção funcionalista dos estados psicológicos com o papel normativo da psicologia cotidiana. Como, a partir do uso de uma atitude normativa utilizada na individuação de estados psicológicos se poderia chegar à predição causal desses estados do mesmo modo que se chega à predição das relações causais de estados físicos? A proposta do funcionalismo de Kim é identificar os estados psicológicos com estados computacionais do cérebro. Os estados psicológicos (crenças, intenções ou sensações) são estados funcionais identificados a processos físicos caracterizados computacionalmente (sintaticamente)<sup>45</sup>. Para tal, faz-se necessário produzir uma redução teórica, na qual um estado computacional é identificado a uma atitude proposicional. Se num primeiro momento, Kim reconhece que as crenças e determinadas atividades mentais, como o lembrar, são exemplos de estados psicológicos derivados da psicologia cotidiana, em seguida, ele reafirma que são estados psicológicos internos implementados por estados físicos. É exatamente essa oscilação do argumento que será alvo das críticas mais severas de Putnam.

Putnam mostra que as propriedades dos estados psicológicos são muito diferentes das propriedades que o formalismo em computação possui. Ele concorda com Davidson (1995) de que o fato de haver leis não implica que possamos formulá-las em qualquer tipo de vocabulário. Já que uma lei é uma formulação lingüística, incluída em um sistema conceitual.

---

<sup>45</sup> O Funcionalismo opera com duas metáforas computacionais: a analogia entre as distinções hardware-software e mente-cérebro; e a idéia de que os estados mentais são funções multiplamente realizadas, ou seja, são instanciadas mesmo se os realizadores físicos variam (Wilson 2004). A teoria cognitiva computacional é originária das ciências cognitivas. A nomenclatura de ‘ciências cognitivas’ (Varela 1994 e 1996; Andler 2004) designa um novo campo interdisciplinar que abarca diferentes disciplinas como neurociências, psicologia cognitiva, filosofia e lingüística. A história das origens das ciências cognitivas (Dupuy 1994) remonta ao início do movimento cibernético entre 1943 e 1953 tendo como marco as ‘conferências Macy’. A cibernética reunia várias técnicas matemáticas que buscavam suplantar a concepção metafísica do sujeito presente nas psicologias existentes na época, e oferecer uma descrição fisicalista baseada em mecanismos materiais e lógicos que encarnariam a mente. A mente é comparada à máquina lógica de Turing e o cérebro é descrito como uma máquina cibernética (Changeaux 1983). A tese principal da cibernética é a idéia de que o funcionamento do sistema nervoso central é assimilável àquele de uma rede elétrica de comunicação entre neurônios. A associação dos estudos em cibernética aos avanços em neurobiologia resultou, nos anos 1980, na criação do cognitivismo e da vertente que o prossegue, o conexionismo, dando origem aos estudos em inteligência artificial. O cognitivismo pode ser definido como a aplicação do modelo computacional à cognição, e o conexionismo se caracteriza pela concepção que os processos cognitivos emergem de uma rede funcional auto-organizada.

Portanto, quando Davidson afirma a impossibilidade de formular leis psicofísicas estritas que conectem tipos de predicados mentais com tipos de predicados físicos, ele está afirmando a impossibilidade de que pelo menos um desses tipos de vocabulário seja adequado para a formulação de leis. As descrições physicalistas são sistemas conceituais onde é possível formular leis propriamente ditas, leis deterministas. Outros sistemas, como o da psicologia cotidiana, por suas próprias características, impedem a formulação de leis, podendo gerar, no máximo, generalizações com algum grau razoável de previsibilidade.

Portanto, pode-se afirmar que a forma proposta pela teoria computacional para individuar estados psicológicos é utópica. Nenhuma atitude psicológica é individuada dessa forma. Não há como reduzir as atitudes proposicionais a estados computacionais sem tornar o mental epifenomenal. Isso acarreta tal restrição dos conteúdos mentais que acaba por deixar de lado os aspectos explicativos relevantes do comportamento intencional.

Outra dificuldade é a forma como o funcionalismo concebe a *individualização dos estados psicológicos*.

O programa funcionalista de Kim pretende restringir a psicologia a ‘estados internos’ e ‘movimentos corporais básicos’. Isso depende de encontrar um modo de individuar as atitudes proposicionais sem apelar para fatores externos. Pois, na opinião de Kim, a igualdade ou a diferença de dois estados internos não dependem de algo externo ao organismo e nem mesmo do momento no qual o organismo se encontra em dado estado. O estado psicológico interno realmente relevante para o funcionalismo é ter uma crença com um *conteúdo estreito* que deve estar numa relação de dependência com estados físicos especificados não-intencionalmente<sup>46</sup>. Nessa acepção, as crenças têm seus conteúdos fixados de forma independente do contexto e os significados são fixados pelo estado interno no qual se encontra o indivíduo.

A suposição de que existem crenças e desejos individuados como o *conteúdo estreito* de um estado interno computacional levou Kim a propor uma variação do argumento do sócia proposto originalmente por Putnam (1995). Kim imagina um indivíduo e seu sócia, internamente idênticos, com a mesma estrutura física e as mesmas propriedades estruturais e disposicionais, que se encontram no mesmo estado cerebral. Mesmo que seus conteúdos amplos sejam diferentes, quando os dois indivíduos estão no mesmo estado cerebral, suas

---

<sup>46</sup> A causa proximal, relevante para explicação do comportamento, deve estar localizada no interior do organismo, pois o que está ‘fora da cabeça’ é considerado impotente causalmente e, portanto, sem valor explicativo.

crenças têm o mesmo *conteúdo estreito*. Para Kim, o argumento mobilizado na interpretação do argumento do sócia pode ser corretamente subsumido sob a tese da superveniência forte. Se abstrairmos todas as atitudes proposicionais cuja atribuição depende de eventos objetivos no meio ambiente do indivíduo (como ‘saber que p’, ‘lembrar que p’, ‘reconhecer que p’), todas as propriedades e relações relevantes para explicações do comportamento devem ser comuns aos sócias (ver Mendonça 1997b).

Essa distinção entre *conteúdo estreito e conteúdo amplo* é característica da tese forte da superveniência. A manutenção da distinção se tornou uma tentativa de resistir ao externalismo, a partir da idéia de que haveria alguma noção de conteúdo que seria imune aos argumentos externalistas. Os defensores do individualismo exigem mais do que a superveniência fraca do psicológico sobre a base do físico, eles exigem a superveniência forte no sentido de uma relação de dependência entre propriedades psicológicas e propriedades fisiológicas especificáveis de forma individualista. Essa concepção dispensa a necessidade de considerarmos a situação do organismo num contexto natural e social mais amplo, pois os estados internos *a la* Kim excluem todas as propriedades relacionais do organismo. Isso faz com que individualistas se oponham contra uma forma de superveniência que se estenda a eventos fora do organismo. Como os estados mentais de *conteúdo amplo* (atitudes proposicionais) não sobrevivem aos estados internos, os individualistas concluem que eles devem ser objetados e desconsiderados da perspectiva científica da psicologia.

A crônica das Terras Gêmeas de Putnam (1995), como veremos a seguir, torna ilógico o argumento de Kim. Como mostra Putnam, não existem estados internos<sup>47</sup> que possam ser identificados como crenças, pois os significados de nossas crenças são, em parte, individuados por fatores externos. A forma que usamos para determinar se duas crenças têm o mesmo significado ou significados diferentes é feito de modo holístico. Individuar crenças é uma tarefa interpretativa intimamente relacionada com individuar os conteúdos das palavras e frases que usamos para comunicar e descrever crenças. Individuamos crenças – especificamos seu conteúdo – ao determinar o que o portador da crença deve possuir para que a crença em questão possa fazer sentido ou se estabelecer como uma conexão cognitiva relevante. Dado o

---

<sup>47</sup> Vale à pena ressaltar que Putnam (1999) guarda um sentido inteligível para a noção de estado psicológico interno. Há uma diferença entre ‘ter uma dor’ e ‘pensar sobre’. A dor não se enraíza em nenhum objeto externo, nem está fora do tempo ou do espaço. Ela pode ser dita interna, no sentido que é privada ao indivíduo. Mas ‘imaginar que’, ‘acreditar que’ pensar sobre’ são atividades que só podem ser definidas a partir do contexto em que se expressam.

holismo do mental<sup>48</sup>, uma atitude proposicional não pode ser atribuída isoladamente; só faz sentido atribuí-la dentro de um conjunto de outras crenças, desejos, intenções, que se articulam, a partir de princípios normativos que dão coerência e consistência à trama de crenças e desejos. A atribuição de crenças a uma pessoa envolve a atribuição de outras redes de crenças. Portanto, o conteúdo de uma crença depende sempre de seu contexto. Se isso é verdade, então, a idéia de que o significado de uma crença é que ‘*a pessoa está em algum estado interno fixo*’ é falsa.

A outra noção em que se apóiam os argumentos individualistas é a idéia de que somente os conteúdos estreitos poderiam preservar a *relevância causal do mental*.

Jerry Fodor (ver Putnam 1999, Wilson 2004, Mendonça 1997b) é um dos autores que aceitam a tese forte da superveniência e propõe uma relação de determinação metafísica. Fodor argumenta que, *a priori*, os estados psicológicos devem ser taxonomizados ou individuados de forma a sobrevirem aos estados físicos dos indivíduos que implementam esses estados. Defende a versão forte da superveniência sobre a base da fisiologia, na qual o poder causal de uma entidade sobrevém à outra, e propõe individuar os estados mentais pelos poderes causais (*causal powers*) de suas propriedades. Isso significa considerar apenas as relações proximais de causa e efeito dos comportamentos a partir da identificação dos mecanismos que implementam fisiologicamente as variações nos poderes causais. As causas do comportamento devem ser buscadas nos eventos proximais que devem sobrevir aos estados cerebrais, ou seja, à micro-estrutura local (Mendonça 1997b). Nesse tipo de relato, todo e qualquer poder causal de eventos externos distais são excluídos, e os estados mentais individuados como conteúdos amplos são considerados inertes causalmente, pois só teriam valor explanatório se fossem descritos na gramática formal da psicologia computacional.

Fodor oferece também um argumento metodológico a favor do individualismo. Ele sugere que a psicologia adote a atitude do *solipsismo metodológico* de modo a lidar apenas com estados internos, estados psicológicos verdadeiramente subjetivos isolados de suas relações com o mundo exterior. Pois acredita que seja possível ajustar as atitudes proposicionais de forma a isolar o elemento puramente subjetivo. O solipsismo metodológico é a idéia de que em algumas situações o comportamento pode ser explicado sem levar em consideração o ambiente externo, mas apenas alterações dos estados internos do indivíduo.

---

<sup>48</sup> Sobre o holismo do mental ver Davidson (2001), Evnine (1991), Costa (1994), Land (2001).

Isso requer uma abstração do ambiente do indivíduo para taxonomizar ou individuar seus estados psicológicos. Essa suposição levou à identificação do *conteúdo estreito* com o componente semântico de uma gramática internalizada. A *Psychosemantics* de Fodor (Mendonça 1997b) apresenta o *conteúdo estreito* como estruturas conceituais da mente, e concebe a mente como processos de manipulação sintática de símbolos e a atividade do pensamento como mera manipulação de objetos sintáticos em resposta aos inputs perceptivos (Wilson 2004).

Como veremos, os defensores do externalismo, como Putnam (1995 [1975], 1999) e Burge (1988 [1979], 1988), afirmam que não existe uma gramática universal como um conjunto fixo de dados capaz de gerar o componente semântico. Os critérios de uso das palavras, além de estabelecidos pelo ambiente, são ajustados para se adaptarem ao contexto. Ao negar as complexas propriedades relacionais entre o organismo e o ambiente, a psicologia computacional deu origem a um mecanismo explicativo individualista baseado nas propriedades sintáticas, deixando de lado os aspectos semânticos essenciais à psicologia cotidiana. Porém, não é possível fazer psicologia sem determinar as propriedades semânticas dos eventos internos, e essas propriedades não podem ser determinadas de forma individualista. A taxonomia da psicologia científica, e das ciências em geral, é relacional. As ciências não taxonomizam pelos poderes causais, elas individualizam suas entidades de forma relacional levando em consideração as relações históricas que determinam as entidades estudadas.

Nas tentativas de reduzir o mental ao físico, os individualistas consideram o esquema de individuação não-individualista para estados psicológicos incompatível com os princípios da superveniência forte e da eficácia causal. A solução, portanto, parece ser que ou adotamos o eliminativismo<sup>49</sup> e deixamos de lado as propriedades semânticas dos estados mentais relevantes para a compreensão da psicologia humana, ou nos submetemos ao dualismo e postulamos uma alma imaterial como fator causal independente. Ambas as soluções terminam por tornar irrelevante o poder explicativo do mental. As démarches anti-individualistas oferecem uma solução que tenta compatibilizar relevância causal do mental com uma visão de mundo materialista. Eles tentam preservar a identidade físico-mental, pois consideram que o

---

<sup>49</sup> Refiro-me ao materialismo eliminativista dos Churchland, que propõe não apenas a redução ontológica dos fenômenos mentais aos fenômenos cerebrais, mas a conseqüente eliminação do papel explicativo do vocabulário intencional da *folk psychology*. (Andrieu 1998)

externalismo é compatível com a tese fraca da superveniência. A superveniência não seria em si um problema para o externalismo. Como veremos no externalismo das Terras Gêmeas, o argumento do sócia é compatível com a versão fraca da superveniência segundo a qual as propriedades psicológicas sobrevivem às propriedades físicas do organismo *mais* o ambiente natural e social (ver Mendonça 1997b). O problema, para individualistas como Kim e Fodor, é que os externalistas reivindicam um *papel causal* para os estados identificados de forma não-individualista que incluem na sua determinação eventos de fora do organismo.

### 3. Externalismo

A finalidade dessa sessão é explorar as conseqüências das teses externalistas para a filosofia da mente. Para os propósitos desta tese, o *punctum saliens* do externalismo é a afirmação de que a determinação do conteúdo mental envolve o concurso de fatores que podem não ser acessíveis ao próprio sujeito do pensamento. Ou seja, se aceitamos o externalismo parece que devemos abandonar as teses tradicionais sobre autoconsciência e autoconhecimento e mesmo a autoridade epistêmica da primeira pessoa. Aparentemente, então, a tese externalista seria incompatível com a autoridade da primeira pessoa.

O externalismo, também chamado de anti-individualismo, tem como tese central a idéia de que os conteúdos dos estados mentais são constituídos, ao menos em parte, pelas relações entre nós mesmos e o ambiente natural e social. O externalismo do mental emergiu como a extensão, para a totalidade do conteúdo proposicional, da intuição fundamental contida na análise dos termos singulares em semântica: “que a capacidade de pensar em um objeto – e, portanto, de fazer referência a ele através do uso de um termo singular apropriado – é independente da capacidade de identificar descritivamente esse objeto; e que, em troca, fatores contextuais, que podem não ser epistemicamente acessíveis ao sujeito, concorrem para fazer com que meu pensamento seja sobre tal objeto particular, e não sobre qualquer outro” (Faria 2006: 3). Ao estender as teses externalistas ao conteúdo intencional em geral, a nova teoria da referência culminou numa concepção externalista da intencionalidade<sup>50</sup>, na qual o conteúdo mental é uma propriedade relacional e não-intrínseca dos estados mentais.

A idéia de que nossos pensamentos e os significados do mundo são fixados por fatores externos que ignoramos sugere que podemos *pensar* sem saber exatamente *sobre o que*

---

<sup>50</sup> Sobre as diversas concepções e usos da noção de intencionalidade ver Pierre Jacob (2004).

estamos pensando. Isso nos leva a considerar errônea a idéia cartesiana de que a única coisa que podemos estar certos é o conteúdo de nossa própria mente. Portanto, o externalismo cria restrições ao postulado da indubitabilidade, tal como formulado pelo individualismo cartesiano. A imunidade à dúvida de proposições da forma *'penso que'* consiste em tomar uma classe privilegiada de proposições como conhecimento certo e indubitável. Para Descartes, os conteúdos proposicionais seriam transparentes ao sujeito do pensamento que os observaria na tela mental através de seu olho interno. E não haveria erro ou ignorância acerca dos conteúdos mentais. O postulado da indubitabilidade, também chamado 'postulado da transparência do conteúdo proposicional' (Faria 2006), afirma que posso desconhecer o valor de verdade de meus pensamentos, ou seja, a relação de adequação entre meus conteúdos mentais e a realidade externa que torna minha crença verdadeira, mas não seu conteúdo, como diz Descartes 'que é isso e não outra coisa que estou pensando'. O externalismo impugna esse postulado. Ao defender a determinação contextual do conteúdo das atitudes proposicionais, coloca em dúvida a suposição cartesiana do senso comum de que cada sujeito tem acesso epistemicamente privilegiado aos conteúdos de seus próprios pensamentos. Essa observação resume a disputa contemporânea sobre a compatibilidade entre externalismo e autoridade da primeira pessoa.

Ao examinar os modelos de externalismo, investigarei a restrição ao problema do autoconhecimento. Existem diferentes tipos de externalismo. O externalismo do conteúdo mental é central para a discussão sobre autoconhecimento e autoridade da primeira pessoa.

### **3.1 Os Modelos de Externalismo**

Há um grande número de visões que têm sido chamadas externalistas em filosofia da mente. As idéias de que há uma dimensão social de ter uma mente e que a cooperação social ocupa um papel central na determinação do que o indivíduo pensa vêm sendo debatidas por vários autores. Há uma longa tradição na filosofia e ciências humanas, que, ao invés de se focalizar na análise do indivíduo, tenta compreender o impacto das instituições sociais em moldar o indivíduo e seus conteúdos mentais. Um dos modos pelo quais a sociabilidade tem sido pensada permear o psicológico é pela idéia de que o mental é normativo. A idéia é que o mundo social – as relações interpessoais, os papéis institucionais, convenções e instituições – tem implicações normativas, relativas às práticas sociais de justificação, racionalidade, que

participam da determinação do mental. De forma resumida, a normatividade do mental está baseada na idéia de que os estados mentais (atitudes proposicionais) fornecem razões para agir. A tradição de pensamento que vai de Wittgenstein ao Pragmatismo e às visões interpretacionistas de psicologia cotidiana representadas por Davidson e Dennett, e até mesmo a escola de pensamento enraizada no ataque de Wilfrid Sellars ao Mito do Dado, conectam a normatividade do mental ao externalismo da mente e, de certa forma, podem ser vistas como antecipadoras do debate individualismo-externalismo. O que essas visões partilham é a idéia de que para se possuir certos tipos de estados mentais carregados de significados deve-se estar sujeito a regras, padrões e convenções (ver Wilson 2004).

Não vou discutir nessa tese todas as abordagens externalistas. Vou me concentrar na visão que é motivada pelo argumento das Terras Gêmeas: o “*Twin Earth externalism*” – um artifício de pensamento que se passa no nosso mundo e num mundo contrafáctico. Os principais filósofos responsáveis por esse argumento são Hilary Putnam e Tyler Burge. As teses que Putnam e Burge discutem são a respeito de como individuamos os estados e eventos psicológicos a que nos referimos em descrições, explicações e previsões do comportamento intencional, quer dizer, como identificamos os tipos de atitudes proposicionais. Ambas as teses recorrem a afirmações sobre a intencionalidade e o conteúdo que podem ser aplicadas aos estados psicológicos em geral. Putnam e Burge acreditam que o caráter social do conteúdo mental se estende virtualmente a todas as expressões de nossa linguagem e todo aspecto de nossos pensamentos, e que, portanto, para entender o significado de um conteúdo é necessário entender a linguagem dos parceiros da comunidade lingüística. Apresentarei as teses desses autores e os desdobramentos de tais teses para a filosofia da mente, em particular para o problema do autoconhecimento. Depois apresentarei o externalismo de Davidson que oferece respostas a alguns problemas levantados pelo externalismo das Terras Gêmeas, e tenta compatibilizar externalismo e autoconhecimento.

### **3.2 O Externalismo Semântico de Putnam**

As teorias externalistas surgiram em meados da década de 60 como reação às teses descritivistas em semântica sobre a referência dos termos. Na perspectiva descritiva, os termos são tratados como descrições e a referência do termo como uma coisa do mundo que satisfaz

as várias descrições às quais associamos o termo. Portanto, o sentido e a intensão<sup>51</sup> determinam a referência. Esta visão descritiva da referência foi o alvo principal das críticas de Saul Kripke<sup>52</sup> e Hilary Putnam, que, por meio de uma teoria causal da referência, buscaram oferecer um outro modo de pensar nosso uso da linguagem em nossa interação com o mundo.

O externalismo emerge historicamente na forma da doutrina da Referência Direta cujo representante mais significativo é Saul Kripke<sup>53</sup>. Kripke tenta fornecer uma visão externalista para o modo como construímos a referência. Sua teoria causal da referência sustenta que os significados dos nomes próprios fracassam na construção de uma referência descritiva, pois esses termos têm denotação, mas não conotação. Os nomes próprios são chamados por Kripke de *indicadores rígidos (rigid designator)*. O indicador rígido é o termo que designa a mesma entidade em todos os mundos possíveis. Na nova teoria da referência, o significado de um nome próprio é determinado pelas relações de seu usuário com o ambiente e com a comunidade de usuários do nome. Para Kripke, há uma cadeia histórica de usos do nome próprio, mas não é o nosso conhecimento dessa cadeia que valida o nosso uso do nome, mas a existência da cadeia<sup>54</sup>. Putnam (1995 [1975]) estende essa tese da designação rígida aos termos das *espécies naturais (natural kinds)* – como, por exemplo, água, ouro e tigre. Ele

---

<sup>51</sup> Como esclarece Costa (1995): “dois termos têm o mesmo referente se possuem as mesmas extensão e intensão (com “s”). A extensão de um termo é o conjunto de coisas ao qual o termo se aplica; a intensão é aquilo que circunscreve a variação do sentido em função do contexto. Por exemplo, a palavra rato, em português, pode ser usada para aludir aos “animais roedores”, aos “ladrões” ou, ainda, a “pessoas moralmente desprezíveis”. No primeiro caso temos uma intensão de rato diferente da intensão expressa nos últimos casos. A intensão especifica o sentido, ou seja, indica como a extensão muda de acordo com as circunstâncias lógica ou empiricamente possíveis.” (Costa 1995: 67)

<sup>52</sup> Considera-se o marco clássico de surgimento do externalismo as conferências proferidas por Saul Kripke em Princeton em 1970, publicadas em Kripke (1972), *Name and Necessity* (*apud* Putnam 1995)

<sup>53</sup> Paulo Faria (2001, 2006) apresenta um bom resumo da história conceitual do externalismo e uma exposição e apreciação sucinta das duas principais doutrinas sobre a referência – a saber, a doutrina ‘descritivista’ do significado de Frege, Russel e Carnap, e a doutrina da Referência Direta de Saul Kripke e Keith Donnelland (que surge como oposição ao descritivismo) – e a consequência dessas teses semânticas para a epistemologia. Apesar da importância dessas teses para o externalismo do mental, não as explorarei aqui, pois exigiria uma análise abrangente do campo da semântica e das teorias da linguagem que fugiria muito aos meus propósitos explicativos. Paulo Faria (2001, 2006) apresenta também uma boa referência bibliográfica dos autores e textos mais importantes nessa discussão.

<sup>54</sup> Como esclarece Peter Geach: “Eu penso, de fato, que para o uso de uma palavra como nome próprio alguém deve estar no começo familiarizado (*acquainted*) com o objeto nomeado. Mas a linguagem é uma instituição, uma tradição; e o uso de um dado nome para um dado objeto, como outros aspectos da linguagem, pode ser transmitido de uma geração a outra; a familiaridade (*acquaintance*) requerida para o uso de um nome próprio pode ser mediata, não imediata. Platão conheceu Sócrates, e Aristóteles conheceu Platão, e Teofrasto conheceu Aristóteles, e assim por diante em sucessão apostólica até a nossa época, é por isso que podemos legitimamente usar ‘Sócrates’ como um nome do modo como o fazemos. Não é o nosso conhecimento dessa cadeia que valida nosso uso, mas a existência da cadeia; assim como, de acordo com a doutrina católica, um homem é um verdadeiro bispo se há, de fato, uma cadeia de consagrações que remonta aos Apóstolos, não se *sabemos* que há.” (*apud* Faria 2006: 4)

afirma que a relação de designação entre “água” e água é rígida, pois o que chamamos “água” não poderia ser uma substância distinta em outro mundo possível, pois se fosse uma substância distinta não seria água. Portanto, a referência de um termo é fixada pelas condições históricas e naturais da coisa que o termo designa.

O *locus classicus* do externalismo é o artigo “The meaning of ‘meaning’” (1995 [1975]) de Putnam. Nesse artigo, Putnam está interessado em discutir a *noção de significado*. Ele critica as principais intuições das concepções tradicionais das teorias da referência, em particular as intuições de Frege e Carnap<sup>55</sup>, e apresenta as modificações propostas por ele e Kripke. (Concentro-me nas críticas anti-individualistas relativas à noção de *estado psicológico* oriundas da nova teoria causal da referência)

Nas teorias descritivas da referência, como Frege e Carnap as formularam, a referência de um termo é fixada ou metafisicamente determinada pelas descrições que os falantes conectam ao termo. Contrários às concepções semânticas tradicionais que concebiam o significado como uma entidade mental, eles propunham que o significado seria uma propriedade pública, partilhada por várias pessoas e em momentos diferentes, e que os conceitos seriam entidades abstratas e não mentais. Contudo, mantinham que a apreensão dessas entidades abstratas seria *um ato psicológico individual*. Nesse sentido, a intensão de um termo seria uma propriedade a ser apreendida, e apreender uma intensão seria associar a intensão a um termo. Portanto, compreender uma palavra (ou seja, conhecer sua intensão) seria estar num certo estado psicológico.

As teses das teorias tradicionais da referência em linguagem natural assentam-se em duas suposições acerca do *significado*: (1) o significado de um termo (sua intensão) determina sua referência (extensão); (2) conhecer o significado de um termo é estar num certo estado psicológico particular. Putnam foi o primeiro a mostrar a incompatibilidade dessas duas teses. Ele argumenta que não há noção de *significado* que possa conjuntamente satisfazer essas duas teses. Vejamos cada uma delas separadamente.

A primeira sustenta que o *sentido* e a *intensão* do termo determinam sua *referência*. A referência e a extensão do termo são fixadas pelo conjunto de descrições que a pessoa conecta ao termo como parte da compreensão do sentido. E como essas descrições, assim apreendidas,

---

<sup>55</sup> Imagino haver alguma imprecisão em juntar ambos autores, Gottlieb Frege e Rudolph Carnap, como pertencendo a uma única escola de pensamento. Mas devido ao meu desconhecimento desses autores e das teorias da linguagem natural, apenas reproduzo as idéias que os autores que utilizo imputam a esses autores.

*estão na cabeça*, pode-se dizer que a extensão do termo passa a ser determinada por algo que estaria *dentro da cabeça* do indivíduo. Nessa acepção individualista, o conhecimento do significado está metafisicamente determinado pelo que está *dentro da cabeça*. Para Putnam, isso não se aplica à linguagem natural e nem à linguagem de forma geral, pois tornaria a linguagem dependente do universo privado do indivíduo.

A segunda tese opera com a noção de estado psicológico *no sentido estreito* pressuposta pela noção de ‘apreender uma intensão’ da semântica tradicional. Considera-se que o estado psicológico de *conteúdo estreito* determina a intensão o qual por sua vez determina a extensão dos termos. Nesse sentido, ‘apreender uma intensão’ equivaleria a encontrar-se num certo estado psicológico *no sentido estreito*. Como esclarece Putnam, “segue-se disso que se dois seres humanos estão no mesmo estado psicológico total, então, eles necessariamente designam a mesma extensão a todo termo que empregam.” (1995 [1975]: 264). Putnam considera que isso não se aplica à linguagem natural, porque a extensão é, em parte, determinada socialmente e não pela competência individual.

O alvo da crítica de Putnam é a suposição implícita nessas teses de que não existiria o mundo além do indivíduo usuário da linguagem, e de que os ‘estados internos’ determinariam o que o indivíduo tem em mente com o seu pronunciamento lingüístico. Ele mostra que ela é incompatível com a suposição de que a intensão do termo determina sua extensão. Os estados internos não podem determinar o que o indivíduo tem em mente com seu pronunciado lingüístico.

Ao estender a teoria da referência de Kripke aos termos de *espécies naturais*, Putnam busca reconstruir a noção de *significado* de forma a preservar a idéia de que o significado possui um *componente indexical (indexicality)*, ou seja, “nossos termos se referem a coisas que são similares, de um certo modo, a coisas que designamos rigidamente” (1995 [1975]: 246). Apontando para o caráter indexical do significado, ele deseja chamar atenção para a existência da relação contextual que conecta o termo à sua referência. A extensão de um termo é determinada socialmente pelo que Putnam denomina de *divisão do trabalho lingüístico*, um complexo processo cooperativo de fixação da referência feito pela comunidade lingüística, incluindo seus especialistas. As essências subjacentes dos referentes de expressões de espécies naturais, como ‘água’ ou ‘ouro’, são descobertas *a posteriori* por meio da metodologia científica. Nesse sentido,

“o significado de uma tal expressão, quando proferida por um indivíduo qualquer, não somente inclui a espécie correspondente como sua extensão, mas, além disso, inclui esta espécie como ela é determinada pelos melhores métodos científicos disponíveis. (...) O que alguém significa (tem em mente, dá a entender), portanto, com uma expressão deste tipo é determinado através de um fato cientificamente fundado sobre o mundo natural e não pode mais ser remediado exclusivamente a propriedades intrinsecamente individuais”. (Mendonça 1997: 3)

Putnam constrói um experimento de pensamento para elucidar o equívoco. Vejamos como ele exemplifica sua tese no *Gedankenexperimente* dos *doppelgängers* apresentado de forma sucinta e esclarecedora por Mendonça.

O argumento do sócia (...) baseia-se na idéia de uma Terra Gêmea (*Twin Earth*) na qual tudo acontece como na nossa Terra. Somente o líquido, que lá também é chamado de “água” e não se distingue “macro-estruturalmente” da nossa água – também e sobretudo nos seus efeitos sobre os órgãos do corpo humano –, tem uma outra micro-estrutura, diferente de H<sub>2</sub>O, uma micro-estrutura que Putnam abrevia com a expressão “XYZ.” A palavra “água,” como utilizada por um habitante qualquer da Terra, Oscar<sub>1</sub>, significa (se refere a) H<sub>2</sub>O, ao passo que a mesma palavra (o mesmo som) na boca de Oscar<sub>2</sub>, um sócia de Oscar<sub>1</sub>, idêntico a este molécula por molécula e em todas as propriedades intrinsecamente individuais, incluindo nisso as propriedades fenomenológicas, se refere a XYZ. (A circunstância de que Oscar<sub>2</sub> não saiba exatamente como é fixada a extensão da palavra “água” em sua linguagem é irrelevante. Tampouco Oscar<sub>1</sub> conhece necessariamente os fatos correspondentes no seu mundo.) Para Putnam, a cadeia causal entre o proferimento de “água” por parte de Oscar<sub>2</sub> e os eventos micro-estruturais do líquido denominado de “água” no seu mundo explica porque a sua palavra “água” se refere a XYZ. Esta cadeia causal é evidentemente muito complexa, na medida em que é mediada pela atividade dos especialistas no seu mundo social. O mesmo vale, *mutatis mutandis*, para Oscar<sub>1</sub> e H<sub>2</sub>O. Se considerarmos, portanto, os fatos que vão além do indivíduo – de tal forma que sejam levadas em conta não somente as relações complexas com o mundo natural, mas sobretudo também a interação social mediada lingüisticamente – então poderemos reconhecer que a palavra “água” tem um caráter essencialmente indexical e se refere *rigidamente* (no sentido de Kripke) a *essa* substância. Assim, de acordo com Putnam, o que Oscar<sub>2</sub> quer dizer (tem em mente, significa) com seu proferimento “água” não é água, isto é, H<sub>2</sub>O, mas sim uma outra substância.” (1997: 4)

A lição que se tira desse exercício de imaginação é que o significado de uma expressão é ao menos em parte constituído pelo ambiente físico e social do usuário da língua. Putnam desenvolveu o argumento das Terras Gêmeas para contestar a tese individualista de que os estados psicológicos internos determinam a extensão dos termos. Ele mostra que os aspectos da história natural de como aprendemos a usar uma palavra altera o que a palavra significa.

Portanto, duas pessoas podem estar em estados físicos idênticos e ainda assim querem dizer coisas diferentes com as mesmas palavras. A extensão de uma palavra não é determinada pelos estados psicológicos. “Cut the pie any way you like, “meanings” just ain’t in the head!” (1995 [1975]: 227). A consequência dessa afirmação é que se as pessoas podem expressar seus pensamentos em palavras, então, seus pensamentos devem, em parte, ser identificados por eventos e objetos fora da pessoa. É possível imaginar dois falantes que estejam exatamente no mesmo estado cerebral e no mesmo estado psicológico no *sentido estreito*, embora a extensão do termo A no dialeto de um seja diferente da extensão do termo A no dialeto do outro. Certamente, possuem os mesmos estados psicológicos no *sentido estreito*, mas não necessariamente os mesmos significados.

A noção de conteúdo estreito (*narrow content*) foi introduzida por Putnam (1995 [1975]) ao se referir a estados psicológicos concebidos como estados físicos, cerebrais ou computacionais, em oposição aos conteúdos amplos (*wide content*), como uma crença que inclui a referência dos termos envolvidos, a qual não é fixada internamente. Nesse sentido, os *estados psicológicos estreitos* não corresponderiam às atitudes proposicionais normalmente identificadas pelas relações com o ambiente social.

Putnam argumenta contra a idéia de que *estados psicológicos* como ‘ter crenças’ e ‘saber o significado de uma palavra’ são, ao mesmo tempo: (1) internos (no sentido que não pressupõem a existência de um indivíduo outro que não o sujeito ao qual o estado é atribuído); (2) e identificados como atitudes proposicionais. A seu ver, nenhum estado pode satisfazer essas duas condições, pois normalmente identificamos e individualizamos eventos mentais e significados pelas relações com outros objetos e eventos que não apenas o indivíduo.

Não há consenso em filosofia da mente sobre a distinção inteligível entre *estados de conteúdos estreitos* e *estados de conteúdos amplos*. Mas parece que todos concordam com a advertência de Putnam de que não existem estados que satisfaçam as duas condições. O próprio Putnam já expressava desconfiança quanto à futura utilidade da distinção no mesmo artigo no qual a propôs<sup>56</sup>. Em suas críticas a individualização dos estados mentais, Putnam propõe o abandono da distinção que, como mostrei, será recuperada como argumento contra o externalismo.

---

<sup>56</sup> “A noção de conteúdo estreito (sob a terminologia “estado psicológico no sentido estreito”) foi introduzida por mim em “The Meaning of ‘Meaning’” (embora eu tenha expressado lá minha dúvida de que ela se mostraria útil em psicologia)” (Putnam 1995 [1975]: 121).

Em sua primeira apreciação do externalismo, Putnam restringe sua análise à noção de referência dos termos da linguagem natural e à relação da semântica desses termos com o que está dentro da cabeça<sup>57</sup>. No início, ele tinha, como admite, uma concepção funcionalista do mental. Precisou estender sua crítica antiindividualista da referência das espécies naturais aos conteúdos dos estados mentais para chegar à posição que nega também o individualismo da mente. Como ele mesmo admite em “The Twin Earth Chronicles”:

“negar que os significados estejam na cabeça deve ter conseqüências para a filosofia da mente, mas à época em que escrevi essas palavras eu não sabia ao certo que conseqüências eram essas. Afinal de contas, *conhecer o significado de uma palavra e usar uma palavra com significado* são ‘habilidades mentais’ paradigmáticas; entretanto, eu não sabia ao certo, quando escrevi “The Meaning of ‘Meaning’”, se a moral daquele ensaio deveria ser que não devemos conceber o significado das palavras como algo que está na mente ou se (como John Dewey e William James) deveríamos parar de pensar na mente como algo que está “na cabeça”, e concebê-lo, antes, como um sistema de capacidades e interações que envolvem o ambiente. No fim, terminei oscilando entre essas duas posições.” (*apud* Faria 2006: 6)

Ao abrir mão da noção de conteúdo estreito e das teorias da identidade físico-mental, Putnam cria outra forma de conceber o mental. Ele passa a rejeitar a noção de que nossas características psicológicas são como ‘estados internos’ e que *qua* estados internos devem estar relacionados com o que acontece dentro de nosso corpo. As características psicológicas são individuadas de modo a serem sensíveis a fatores externos: a natureza dos objetos que percebemos e sobre os quais agimos, os fatores sociais e históricos e nossas práticas lingüísticas comunicativas. A distinção entre, por exemplo, imaginar, acreditar ou lembrar é sempre dependente do contexto no qual a afirmação é feita e não de um suposto *estado estreito* que fixaria o conteúdo da crença. A condição neurológica do indivíduo não é suficiente para determinar se o indivíduo se refere a uma crença ou outra. Como alternativa a noção de estado mental interno, Putnam considera mais pertinente o uso do conceito de ‘evento mental’, adotando a concepção de evento *à la* Davidson, pois o mesmo não se refere a

---

<sup>57</sup> Na conclusão de “The meaning of ‘meaning’”, ele apenas faz menção ao problema do individualismo da mente em filosofia visto como uma “tendência a tratar a cognição como uma questão puramente *individual* e uma tendência a ignorar o mundo, na medida em que o *mundo* consiste em mais do que ‘observações’ dos indivíduos. Ignorar a divisão do trabalho lingüístico é ignorar a dimensão social da cognição; ignorar o que chamamos de *indexicalidade* de muitas palavras é ignorar a contribuição do ambiente. A filosofia tradicional da linguagem, como a maior parte de filosofia tradicional, deixa de fora outras pessoas e o mundo; uma melhor filosofia e uma melhor ciência devem abarcar ambos” (1995 [1975]: 271).

algo estritamente interno. O evento mental, quando aplicado ao organismo, num tempo e espaço determinado, depende conceitualmente do que acontece em outro tempo ou outra existência das coisas fora do organismo. Vale ressaltar que isso não significa uma adesão de Putnam ao monismo anômalo de Davidson<sup>58</sup>.

Em *Three-fold core: mind, body and world*, como o título sugere, Putnam reformula sua teoria da mente levando em consideração as conseqüências externalistas do mental<sup>59</sup>. Ele concebe a mente como um conjunto de capacidades que envolvem o mundo e é dependente do contexto daquele que pensa. Para Putnam, o mental é um modo de descrever o exercício de certas habilidades que possuímos que sobrevêm sobre atividades de nossos cérebros e sobre várias transações com o ambiente, mas que não devem ser explicadas de forma redutível pelo vocabulário das ciências duras, mas por uma pluralidade de fontes conceituais.

### 3.3 O Externalismo Social de Burge

Em “Individualism and the Mental”, Tyler Burge defende a tese antiindividualista de que os conteúdos intencionais de nossos pensamentos e a forma como os taxonomizamos **não**

---

<sup>58</sup> A teoria da identidade *token-token* de Davidson foi discutida por mim em Ferreira (2000). “No binômio monismo - anômalo, certamente a tese do monismo foi a mais atacada, porque dependia de uma pressuposição fundamental, assim formulada: 'um evento-token mental x é idêntico ao evento-token físico y'. (...) O problema de Davidson era, então, defender um monismo não reducionista, o que o obrigou a formular uma versão não reducionista de uma teoria de identidade entre o físico e o mental. Para postular a identidade entre dois eventos, ele propôs, em uma das versões da teoria, que o evento-token mental x seria idêntico ao evento-token físico y, se e somente se eles possuísem a mesma causa e produzissem o mesmo efeito. Porém, o critério escolhido não se sustentou por muito tempo. Como nos mostra Putnam (1999), Quine demonstrou a **circularidade** inerente ao critério escolhido. Para saber que um evento-token mental x teve a mesma causa que o evento-token físico y e que ambos produziram o mesmo efeito, temos que primeiro ter um critério para individualizá-los. Ou seja, temos que ter um critério para separá-los na continuidade fenomênica do mundo, sendo este critério necessariamente lingüístico. Como Davidson descartou, por premissa, critérios definicionais (analíticos) ou nomológicos para estabelecer a identidade, geralmente utilizados em teoria *type-type*, restou-lhe apenas postular a identidade de causas e efeitos sem perceber que ela dependia primeiro da capacidade de reconhecer que os eventos eram idênticos. A circularidade era, portanto, a seguinte: só sabemos que os eventos têm as mesmas causas e produzem os mesmos efeitos se já sabemos que esses eventos são idênticos. Logo, o critério escolhido não serve para a finalidade pretendida. As outras tentativas de Davidson para individualizar eventos também não foram bem sucedidas, estando as justificativas do monismo em aberto, do ponto de vista teórico, para outras propostas e formulação.” (Ferreira 2000: 21) Apesar do monismo anômalo escapar aos problemas apresentados pelas teorias da identidade *type-type* que adotam a versão forte da tese da superveniência, o relevante para o argumento atual é mostrar que Davidson também não resolve o problema.

<sup>59</sup> Atualmente, existe uma forma mais radical de externalismo, chamada de externalismo localizacional ou ativo (Clark & Chalmers 1998), que leva em consideração não apenas o papel dos ambientes físico e social na individuação dos estados mentais, mas enfatiza os aspectos arbitrários da cabeça e da pele para limitar as fronteiras dos processos cognitivos na vida real. O que os autores propõem não se limita à discussão sobre relatos ou visões sobre a cognição, eles tentam construir uma ontologia da mente externa. (Wilson & Clark 2006; Wilson 2005)

são determinados pelas propriedades intrínsecas do indivíduo considerado em isolamento de seu entorno social.

Para Burge, na maioria das visões dominantes sobre o papel da linguagem na constituição do indivíduo, há uma negligência dos fatores sociais. Embora o uso social da linguagem e de sua função de *comunicação* entre indivíduos não ter sido ignorado pela filosofia e a psicologia, os aspectos sociais têm sido tratados de forma derivada, dando origem a uma semântica individualista. Burge acredita que isso é a consequência da não consideração da função primária de comunicação da linguagem e do fato da investigação da mente nos levar, intuitivamente, para *dentro da cabeça*. Nas ciências cognitivas, por exemplo, como vimos, a mente tem sido concebida como uma forma de computação, na qual os aspectos sociais são, de algum modo, secundários ou derivados. Burge argumenta contra essa abordagem individualista que explica os comportamentos intencionais do indivíduo face ao seu ambiente físico sem fazer referência ao contexto social em que o mesmo está situado.

O antiindividualismo de Burge está baseado na análise de estados mentais com conteúdos proposicionais. Ele está interessado em analisar os discursos intencionais que compõem a *psicologia cotidiana* e suas conexões com os fatores externos. Os discursos intencionais<sup>60</sup> são os que possuem aspectos semânticos que indicam a atitude proposicional do estado ou evento mental que caracterizam a perspectiva epistêmica da pessoa. Como esclarece Mendonça:

“Convicções, desejos, intenções e outras atitudes proposicionais são estados intencionais com um conteúdo proposicional que é especificado, na atribuição normal, através de uma *oratio obliqua*. A intencionalidade do mental significa, neste caso, que na especificação do conteúdo algumas expressões linguísticas são utilizadas *obliquamente*, isto é, de tal forma que não podem, via de regra, ser substituídas *salva veritate* por expressões extensionalmente equivalentes. (...) Uma convicção tem, portanto, propriedades semânticas que estão numa conexão estreita com as propriedades semânticas da sentença com a qual seu conteúdo é especificado, o que, por sua vez, está conectado ao fato de que o proferimento desta sentença pode funcionar como a expressão da convicção em questão. Quem conhece o conteúdo de uma convicção conhece também o

---

<sup>60</sup> Os pensamentos e conteúdos mentais têm sido divididos em dois tipos de discurso mental: (1) o que faz referência a estados ou eventos mentais em termos de expressões sentenciais como ‘Eu acredito que p’, são as chamadas *that-clauses*, que são orações subordinadas que compõem o discurso mentalista (pensamentos, crenças, intenções), e que podem ser julgadas como verdadeiras ou falsas; (2) e outro que faz referência a estados mentais em termos de expressões como ‘Eu sinto dor’ (dores, sentimentos, cócegas). O discurso mentalista que contém expressões de ocorrência oblíqua tem sido chamado de *discurso intencional*. O discurso intencional é composto pelas *that-clauses* que são ‘orações subordinadas que complementam uma atitude proposicional’ (Essa tradução foi sugerida por Roberto Horácio em comunicação oral), que são expressões de ocorrência oblíqua cuja função é fornecer o conteúdo dos estados ou eventos mentais (Ver Burge 1988 [1979]).

estado de coisas que determina o valor semântico da convicção. Isto vale, *mutatis mutandis*, para todas as atitudes proposicionais. O conteúdo de uma atitude proposicional (sua “intensão”) determina, neste sentido, a extensão. A identificação de uma atitude proposicional (quanto ao tipo) por meio do conteúdo implica, porém, que *orationes obliquae* – que, por causa da ocorrência opaca de expressões extensionalmente equivalentes, não são intersubstituíveis *salva veritate* – contribuem para a individuação de diferentes atitudes proposicionais.” (1997a: 3)

Como mostra Burge, por exemplo, dado o fato que água é H<sub>2</sub>O, uma pessoa pode ser da opinião de que ‘água é própria para beber’, mas, ao mesmo tempo, não ser da opinião de que ‘H<sub>2</sub>O é próprio para beber’. Em outras palavras, quando uma expressão como água é usada numa oração subordinada (*that-clause*) ela não é livremente modificada por todas as expressões extensionalmente equivalentes (água, H<sub>2</sub>O). Essas expressões não são intercambiáveis, pois ‘água’ ocupa um papel em caracterizar um estado mental diferente do estado mental caracterizado por H<sub>2</sub>O em ocorrências oblíquas. Considera-se que as expressões que ocorrem obliquamente nas orações subordinadas são um meio primário de identificar os estados e eventos intencionais da pessoa. Essas frases corretamente atribuídas a uma pessoa identificam diferentes eventos mentais. Portanto, diferenças nas extensões implicam diferenças nos conteúdos mentais que implicam em diferentes estados mentais.

Para compreender como é que os conteúdos mentais podem variar em função de diferenças no contexto social, Burge usa uma variação do experimento de pensamento das Terras Gêmeas de Putnam. Em suas versões do argumento do sócia, Burge generaliza os argumentos de Putnam sobre o modo como construímos a referência dos termos das *espécies naturais* para todos os contextos sociais de uso da linguagem e enfatiza o papel desempenhado pelo ambiente social na constituição dos estados mentais do indivíduo. A contribuição de Burge para a discussão é que seu argumento não pressupõe nenhum tipo de essencialismo científico sobre a semântica de termos de espécies naturais, tal como o *externalismo semântico* de Putnam.

Como mostrei, a visão de Putnam supõe uma forma de externalismo no qual os conteúdos mentais são afetados não apenas pelo estado físico e mental em que a pessoa se encontra, mas pela natureza de seu *ambiente físico*. É o caráter *físico* do mundo – a natureza das coisas físicas, como a água ou o ouro – que, em parte, determina os conteúdos mentais da pessoa. No argumento de Putnam, o uso rígido de um indicador, que designa uma espécie

natural, envolve pressuposições empíricas sobre as propriedades estruturais relevantes da substância em questão, e tais propriedades estruturais devem especificar e explicar o modo como os componentes internos da substância se combinam para gerar suas propriedades macro-estruturais. (Putnam 1995 [1975]; Mendonça 1997b).

O *externalismo social* de Burge distingue-se da versão de Putnam, pois Burge estende o argumento a expressões não naturais – como no caso da palavra artrite que designa somente um conjunto de sintomas ao qual se dá convencionalmente uma etiqueta verbal (Burge 1988 [1979]). Segundo ele, o caráter social do mundo e a natureza da comunidade lingüística da pessoa determinam o significado. Burge parte da constatação de que temos uma compreensão incompleta de muitas coisas acerca das quais temos pensamentos e para as quais temos expressões na linguagem natural. Em muitas situações é possível atribuir um estado ou evento mental cujo conteúdo envolve uma noção que o sujeito compreende de forma incompleta. Segundo Burge, “uma série de erros conceituais é comum entre nós (...) e tais erros não impedem sempre ou automaticamente a atribuição de conteúdos mentais fornecidos por todos os termos que são compreendidos de forma incompleta ou aplicados de forma equivocada” (1988 [1979]: 34). A ‘compreensão parcial’ é comum na maioria das expressões de nosso vocabulário.

Em suas versões da crônica das Terras Gêmeas, Burge propõe imaginarmos um doente imaginário chamado Bertrand que sofre de artrite e seu sócia que vive no mundo contrafáctico. Ambos são idênticos no que diz respeito às propriedades intrínsecas e à experiência fenomenal. A contrafactualidade refere-se às alterações no ambiente social do indivíduo. Bertrand, apesar de dominar o conceito de artrite, comete um erro ao se queixar ao seu médico que sua artrite se deslocou do joelho para a coxa. Como artrite ocorre somente nas juntas, atribui-se a Bertrand uma crença falsa baseada na interpretação de suas palavras. Agora, imaginemos o sócia de Bertrand. O sócia não exprime no mundo em que vive uma crença falsa quando profere a mesma crença, pois a palavra ‘artrite’ é usada na sua comunidade lingüística como denominação de todas as doenças semelhantes ao reumatismo. Nesse sentido, *Bertrand1* tem atitudes proposicionais que têm ‘artrite’ como objeto intencional que *Bertrand2* não pode ter (Mendonça 1997b). As experiências qualitativas internas, seus estados e eventos fisiológicos, seus comportamentos e disposições para agir permanecem constantes, enquanto suas atitudes proposicionais variam. As diferenças em seu conteúdo são atribuíveis a

diferenças em seu ambiente social. De fato, atribuímos pensamentos com diferentes conteúdos a Bertrand e a seu sócia, e parece que fazemos isso apenas por causa dos aspectos da linguagem da comunidade na qual ele está situado.

A estratégia da crônica das Terras Gêmeas é manter constante a história dos movimentos corporais da pessoa, a superfície das estimulações e a fisiologia interna. Ao variar o ambiente no qual a pessoa interage enquanto mantém constante os efeitos moleculares do corpo da pessoa, pode-se mostrar que alguns de seus pensamentos variam. A variação dos pensamentos é indicativa da existência de princípios subjacentes de individuação de tipos mentais.

A conclusão que Burge chega é que as atitudes proposicionais não são fixadas somente pelo que se passa com o indivíduo, mas elas são fixadas por fatores externos e sociais que determinam os conteúdos da mente da pessoa. Isso o leva, assim como a Putnam, a argumentarem contra as teorias individualistas de individuação das atitudes proposicionais. Putnam e Burge enfatizam a idéia de que a mente é social por natureza, ou seja, a cognição está situada (*embeddedness*) no ambiente social<sup>61</sup>. Para ambos, a dimensão social da mente demanda o abandono dos modelos individualistas da mente. Eles rejeitam tanto o cartesianismo e o behaviorismo quanto o funcionalismo. (Mesmo Putnam rejeita o funcionalismo no seu externalismo revisado). Assim como Putnam, Burge reconhece várias deficiências da concepção funcionalista do mental. Além da negligência de fatores sociais relevantes na determinação do mental, Burge condena também o modo como o funcionalismo propõe individuar os eventos mentais. Para Burge, o funcionalismo e todas as teorias da identidade físico-mental falharam em oferecer tipos de identidades entre estados funcionais e estados mentais.

Como vimos, as teorias materialistas da identidade fornecem um relato em que os conteúdos mentais são descritos em termos não-intencionais negligenciando o papel semântico e intencional dos estados e eventos mentais. As teorias materialistas da identidade postulam

---

<sup>61</sup> O movimento da cognição situada também adotou esse tipo de visão da mente situada (*embeddedness*) no mundo externo. Os trabalhos sobre a cognição situada têm geralmente focalizado como os indivíduos se tornam estreitamente acoplados a seu ambiente físico, mais do que social, concentrando-se no papel que o ambiente físico tem ao realçar e mesmo constituir a performance do indivíduo. Pessoas, artefatos, instituições, práticas e interações, ambos criados e herdados, são os aspectos mais significantes de qualquer ambiente individual para a vida mental. (Ver Clark 1996).

que o conteúdo do evento mental pode variar enquanto os eventos físicos permanecem fixos. Burge argumenta contra esse modo de individuação dos eventos mentais.

Para Burge, as teorias da identidade físico-mental foram tornadas implausíveis pelos experimentos de pensamento não-individualista. No argumento das Terras Gêmeas, o evento de pensamento não é idêntico a nenhum evento descrito em termos físicos que ocorra no sujeito. Nenhuma ocorrência de um pensamento pode ter um conteúdo diferente e ser o mesmo evento. Nesse sentido, se dois eventos mentais têm diferentes conteúdos eles são seguramente eventos diferentes. O fato de que os estados e eventos mentais ordinários sejam individuados em termos de suas relações com o mundo exterior serve como argumento para desacreditar as teorias da identidade físico-mental como tal.

O argumento das crônicas das Terras Gêmeas mostra que se dois eventos mentais têm diferentes conteúdos eles são seguramente eventos diferentes. Se levarmos em consideração as relações intrinsecamente individualistas, mas também as causais-históricas, então os estados mentais dos *doppelgängers* são distintos porque os estados físicos globais nos mundos respectivos são distintos. Veremos no *externalismo perceptivo* de Davidson que “uma psicologia não-individualista não pode mais ser eliminada sob a suspeita de não ser materialista em sua orientação básica. Uma suposição materialista implica a afirmação de que os eventos físicos são fundamentais no sentido em que determinam tudo mais no mundo” (Mendonça 1997b). Portanto, apesar de negar a validade das teorias da identidade físico-mental, o argumento das Terras Gêmeas não nos obriga a abrir mão da dependência do mental ao físico. Ele mostra, simplesmente, que muitos estados intencionais não podem ser remetidos aos estados físicos puramente internos do organismo ao qual são atribuídos. Reconhece-se as relações de dependência entre esses estados intencionais e fatores físicos fora do organismo.

Na crônica das Terras Gêmeas, fica claro que o significado de nossas palavras é fixado em parte pelas circunstâncias nas quais aprendemos e usamos as palavras. Todo pensamento e linguagem estão fundados em conexões históricas que restringem o significado dos conteúdos mentais. O que é importante reter nos argumentos de Putnam-Burge é a idéia de que nós, *language users* e seres psicológicos, dependemos e contamos com o outro de um modo que é refletido nos nossos modos de pensar sobre a linguagem e o pensamento do nosso senso comum cotidiano. Existe um aspecto social da natureza do significado e do pensamento em ambas as visões e isto justifica etiquetá-los de antiindividualistas. Segundo Burge, estamos

circunscritos a dar às palavras de uma pessoa e a interpretar suas atitudes proposicionais a partir do significado dado a essas palavras na sua comunidade lingüística. Portanto, o uso lingüístico social dita o que o falante quer dizer com suas palavras ou o que são os seus estados psicológicos no sentido estreito.

Devemos ter em mente um aspecto social da visão de Putnam: *a divisão do trabalho lingüístico*. Para Putnam, as propriedades do mundo natural determinam o significado e o conteúdo das atitudes proposicionais, não diretamente, mais sempre por meio da mediação do mundo social. O que permite aos indivíduos o uso de termos de espécies naturais para se referirem a objetos no mundo é a sua habilidade para tomar de empréstimo as descrições de identificação de outros especialistas, que são capazes, de forma confiável, de escolhê-las por meio do conhecimento deles. A originalidade do argumento de Burge, como afirma Mendonça, é mostrar que:

“a contribuição da sociedade para a determinação dos estados mentais do indivíduo não precisa necessariamente aparecer *por trás* da contribuição do mundo, como imagina Putnam. O pertencimento de um indivíduo a uma comunidade lingüística por si mesmo pode desempenhar o papel constitutivo na psicologia do indivíduo”. (1997b)

### **3.4 O Externalismo das Terras Gêmeas e o Autoconhecimento**

O externalismo das Terras Gêmeas afirma que os estados mentais, ao menos as atitudes proposicionais, são, em parte, identificados pelas relações com a sociedade e o resto do ambiente, relações que podem, em alguns aspectos, não ser conhecidas pela pessoa que possui esses estados. Essa posição externalista cria um conflito para a existência da autoridade da primeira pessoa. **A démarche externalista uma vez que identifica os estados mentais, em parte, por seus fatores sociais e que os torna passíveis de serem publicamente observados, reintroduz o problema da autoridade da primeira pessoa.**

O externalismo do mental tem sido muitas vezes usado para atacar a autoridade da primeira pessoa tal como formulada por Descartes. Os externalistas incompatibilistas argumentam que os métodos usados ao interpretarmos o outro, o intérprete da terceira pessoa, determinam o que estimamos ser os conteúdos mentais do agente. Porém, o que determinamos como seus conteúdos mentais pode ser desconhecido do próprio agente. Se o agente pode ter conteúdos mentais que ele mesmo ignora, então, não há garantias para autoridade da primeira

pessoa. Muitos defensores do externalismo não chegaram a considerar isso um problema. De início, o behaviorismo, por sua ênfase no comportamento público e observável do indivíduo, foi invocado para mostrar como é possível uma pessoa conhecer o que outra tem em mente. Porém, ao mesmo tempo, foi rejeitado, pois não explicava o fato de que os estados mentais são conhecidos pela pessoa que os possui sem fazer apelo à evidência comportamental.

Para aqueles que, como Putnam ou Fodor, aceitam a distinção entre os conteúdos da mente determinados internamente e as crenças, desejos e intenções, às quais atribuímos uma base social, o problema da autoridade da primeira pessoa não pode ser resolvido, e, na verdade, nem se preocupam em explicá-lo. Putnam, por exemplo, dá pouca importância ao debate sobre a questão da autoridade da primeira pessoa. Uma resposta, portanto, é manter a distinção entre os conteúdos da mente determinados subjetivamente e internamente e as crenças, desejos e intenções, as quais atribuímos uma base social e conexões externas.

Para Burge, é possível compatibilizarmos a concepção cartesiana do conhecimento privilegiado de nossos próprios pensamentos e a concepção não-individualista de individuação dos pensamentos proposta pelo externalismo. Descartes sustentava a tese que conhecemos nossos eventos mentais de modo direto, autoritativo e não empírico. Burge concorda com essa tese e acredita que Descartes estava certo em afirmar o caráter especial de certeza e diretividade do autoconhecimento. Contudo, Burge afirma que Descartes hipostasiou o poder do conhecimento autoritativo e extraiu conclusões metafísicas da tese do autoconhecimento. Da constatação de que podemos conhecer nossos pensamentos de forma privilegiada e duvidar da existência do mundo físico, ele inferiu que nosso pensamento deve ser independente de qualquer relação com o mundo físico.

Se, para Burge, a individuação dos tipos mentais da pessoa, incluindo pensamentos sobre objetos físicos e propriedades, é necessariamente dependente das relações que a pessoa estabelece com seu ambiente físico e social, e se cada pensamento que a pessoa tem é dependente da relação que estabelece com o meio ambiente, como podemos afirmar que conhecemos nossos eventos mentais de forma direta e não empírica?

Sabemos que uma pessoa não precisa investigar o ambiente para conhecer seus pensamentos. Podemos conhecer o que são nossos próprios eventos mentais mesmo sem conhecer os fatos relevantes sobre as condições de individuação desses eventos. É absurdo imaginar que, para conhecermos ‘o pensamento que estamos pensando’, deveríamos investigar

as condições empíricas que determinam nossos pensamentos de tal modo que pudéssemos distinguir nossos pensamentos.

O argumento de Burge é que não é necessário investigar o ambiente para conhecermos nossos próprios pensamentos. Vejamos como ele desenvolve o argumento.

A crônica das Terras Gêmeas mostra que existem condições externas que determinam nosso pensamento. Digamos que penso que “a água é líquida”. Chamarei este pensamento entre aspas de pensamento de primeira ordem. Dentre as condições que determinam os conteúdos do pensamento de primeira ordem, algumas só podem ser conhecidas por meios empíricos. Nesse pensamento, a referência é a substância água. Para pensar em algo como água devemos estar numa certa relação causal com água, ou seja, num meio ambiente onde exista uma substância aquosa que permite pensar em algo como água. Contudo, estas condições não precisam ser conhecidas, mas apenas pressupostas. Para ter um pensamento a pessoa não precisa conhecer as condições que a capacitam a pensá-lo, apenas que essas condições sejam satisfeitas, e algumas dessas condições só podem ser satisfeitas por meios empíricos. Para conhecer o que estamos pensando quando temos pensamentos sobre entidades físicas, são pressupostas algumas condições que determinam os conteúdos dos pensamentos empíricos daquilo que sabemos que estamos pensando.

Agora, digamos, que penso “creio que a água é líquida”. Chamarei esse pensamento de pensamento de segunda ordem ao qual se refere à noção de autoconhecimento. O autoconhecimento não requer o conhecimento das condições que capacitam o sujeito a ter o pensamento. O autoconhecimento consiste num juízo reflexivo que envolve pensar sobre um pensamento de primeira ordem, sobre o qual é o juízo. Esse juízo reflexivo herda o conteúdo do pensamento de primeira ordem. Não há nenhum critério de identificação entre dois tipos de pensamento. Conhecemos nossos próprios pensamentos apenas exercitando os poderes das atribuições de segunda ordem, ou seja, simplesmente pensando o pensamento de modo reflexivo. Como esclarece Burge:

“‘individuamos’ nossos pensamentos, ou os discriminamos de outros, ao pensar neles e não em outros de forma auto-atributiva (*self-ascriptively*). Grosso modo, nosso conhecimento de nossos próprios pensamentos é imediato, não discursivo. Para sua justificação, o simples autoconhecimento não necessita, de nenhum modo, de suplementação de investigações discursivas ou comparações.” (1988 [1979]: 119)

Burge apresenta algumas características que distinguem o autoconhecimento. Concentro-me em duas delas. A primeira característica é o aspecto da primeira pessoa do autoconhecimento. O estatuto epistêmico especial desses pensamentos depende deles serem feitos simultaneamente a partir e sobre o ponto de vista da pessoa. O pensamento reflexivo identifica o ponto de vista. Portanto, o caráter da primeira pessoa é fundamental para o seu estatuto epistêmico. Ao se ter um pensamento reflexivo, esse pensamento é simultaneamente objeto da referência e constituinte de um ponto de vista.

A segunda característica salientada por Burge é o aspecto auto-referencial do autoconhecimento, ou seja, ele assegura que o objeto da referência seja somente o pensamento que está sendo pensado. No pensamento de primeira ordem, a referência do pensamento é a entidade que está sendo pensada. Se penso “a água é líquida”, água é a referência. No pensamento de segunda ordem, “creio que água é líquida”, a referência do pensamento é “meu pensamento que ‘a água é líquida’”. Portanto, o pensamento de segunda ordem é auto-referencial, pois se refere a algo do próprio mundo de pensamento do sujeito. Desse modo, o conteúdo proposicional do pensamento de primeira ordem ‘a água é líquida’ é incorporado ao pensamento de segunda ordem ‘creio que a água é líquida’. Como afirma Burge:

“No autoconhecimento básico, alguém simultaneamente pensa por meio do pensamento de primeira ordem (que a água é líquida) e pensa nesse pensamento “como sendo seu próprio. O conteúdo do pensamento de primeira ordem (contido) é fixado pelo background de condições não-individualistas. E pelo seu caráter reflexivo e auto-referencial, o conteúdo do pensamento de segunda ordem está logicamente bloqueado (de forma auto-referencial) pelo conteúdo de primeira ordem que ele tanto contém como toma como assunto.”  
(1988 [1979]: 122 )

Portanto, para Burge, ainda que os conteúdos sejam determinados em parte externamente, para saber (segunda ordem) o que estou pensando (primeira ordem) quando penso em "água", não preciso fazer observações do ambiente que me circunda, ou inferências a partir de observações. O conteúdo de primeira ordem fixa a referência do conteúdo de segunda ordem automaticamente. É isso que Burge chama do caráter de autoverificação do autoconhecimento. Assim, se eu estiver pensando em H<sub>2</sub>O (primeira ordem), saberei (segunda ordem) que estou pensando em H<sub>2</sub>O sem ter que verificar o ambiente no qual me encontro.

Para Burge, o pensamento de segunda ordem sobre um pensamento de primeira tem sempre um caráter reflexivo. Ao atribuímos uma crença a nós mesmos, como fazemos no

pensamento de segunda ordem, que “creio que a água é líquida”, o pensamento de segunda ordem opera como uma meta-representação que atribui à representação de primeira ordem (“a água é líquida”) o mesmo conteúdo que a representação de primeira ordem carrega (ou seja, que a água é líquida). Assim, não podemos nos enganar que acreditamos que ‘a água é líquida’, pois o conteúdo daquilo que acreditamos é retirado do próprio pensamento de primeira ordem, qualquer que seja o conteúdo que tenha. Se realmente acredito que ‘a água é líquida’ (crença de primeira ordem), então, tem que ser verdade que ‘creio que ‘a água é líquida’ (crença de segunda ordem), pois a crença de segunda ordem tem seu conteúdo atribuído pela crença de primeira ordem, não importa qual seja o conteúdo. Portanto, não podemos errar ao identificar o conteúdo do pensamento de primeira ordem. Então, mesmo se consideramos que o externalismo de conteúdo é verdadeiro, não podemos estar equivocados em acreditar que creio que ‘a água é líquida’, pois aquilo em que creio que acredito tem seu conteúdo ‘emprestado’ daquilo que acredito<sup>62</sup>. Chamamos a essa forma de conhecer nossos pensamentos de privilégio epistêmico da primeira pessoa. Nessa medida, a autoridade da primeira pessoa não é ameaçada pelo externalismo do conteúdo.

### **3.5 O Externalismo Perceptivo de Davidson**

Na coletânea de artigos publicados em *Subjective, Intersubjective, Objective* (2001), Donald Davidson formula, de forma mais explícita, sua concepção externalista do mental que, de certa forma, permeia toda sua obra. Nesses textos, defende uma concepção do mental que é, em sua origem, externalista e compatível com o privilégio epistêmico da primeira pessoa<sup>63</sup>. Ele mostra que, mesmo admitindo que o conteúdo proposicional de nosso pensamento é determinado por fatores externos, é possível ter um conhecimento de nossas atitudes proposicionais ordinárias que não seja fundamentado na evidência externa.

Na sua formulação original, os propositores do externalismo, ao sustentar a idéia de que nossos pensamentos e os significados do mundo são fixados por fatores externos que ignoramos, foram levados a considerar um erro a idéia cartesiana de que a única coisa que

---

<sup>62</sup> Esse argumento me foi esclarecido por Roberto Pereira em comunicação pessoal.

<sup>63</sup> A defesa que Davidson faz da autoridade da primeira pessoa não significa a manutenção das teses fortes cartesianas do privilégio epistêmico, como a infalibilidade e a incorrigibilidade. Davidson inclui a possibilidade do auto-engano, ou seja, a possibilidade do agente possuir pensamentos que ele mesmo desconhece. Davidson reconhece a assimetria entre primeira e terceira pessoa, mas discorda da interpretação inflacionada do princípio da autoridade da primeira pessoa, formulado pelo ‘postulado da transparência do conteúdo proposicional’, e propõe uma interpretação deflacionada que preserve o conteúdo intuitivo do princípio de discriminação.

podemos ter certeza é o conteúdo de nossa própria mente. Isso levou à conclusão de que a posição externalista não seria compatível com a autoridade da primeira pessoa. Haveria um conflito entre o fato de que os aspectos externos determinam os conteúdos do pensamento e o fato de que conhecemos o que pensamos de uma forma que os outros não podem conhecer. O problema surge se pensarmos que admitir a relevância dos fatores externos na identificação desses estados significa concluir que não podemos conhecer completamente nossos pensamentos. Mas a resposta a essa questão depende de como concebemos a contribuição dos fatores externos na identificação do mental.

Para Davidson, o externalismo como formulado por Putnam-Burge oferece dois obstáculos para compatibilizar essas teses. Primeiro, não reconhece que o conteúdo mental proposicional possa ser idêntico ao estado cerebral, e, com isso, desacredita qualquer teoria da identidade físico-mental, até mesmo a teoria da identidade *token-token* proposta por seu monismo anômalo. Segundo, opera com o que Davidson chama de metáfora dos ‘objetos diante da mente’. Começamos pelos argumentos a favor da identidade entre os eventos físicos e mentais.

No debate acerca da individuação de estados psicológicos, Davidson (2001 [1987]) se posiciona de forma diferente do externalismo de Putnam-Burge. Ele defende a tese de que os estados mentais ordinários satisfazem às duas condições: são estados internos idênticos aos estados corporais, e, ao mesmo tempo, são não-individualistas, pois são identificados em parte por suas relações causais com eventos e objetos externos. Essa tese não ameaça a autoridade da primeira pessoa que pode, sem contradições, ser aplicada aos estados mentais ordinários assim definidos.

Segundo a tese de Putnam, os conteúdos das atitudes proposicionais são em parte determinados por fatores externos ao agente e desconhecidos pelo agente, e, portanto, o indivíduo não tem autoridade sobre esses conteúdos. Deste modo, somente uma terceira pessoa está numa posição privilegiada para saber qual é o conteúdo do pensamento do indivíduo<sup>64</sup>. Em sua crônica das Terras Gêmeas, Putnam imagina duas pessoas exatamente no

---

<sup>64</sup> Esse problema é similar ao problema do ceticismo que considera que poderíamos estar enganados quanto ao mundo exterior e à possibilidade do conhecimento objetivo do mundo. O ceticismo dos sentidos é evitado ao supor que é o mundo que determina o conteúdo dos pensamentos sobre o mundo. Mas o ceticismo não foi derrotado, apenas foi deslocado para o conhecimento de nossas próprias mentes. A dúvida cética pode ser assim formulada: nossas crenças sobre o mundo externo estão dirigidas para o mundo, mas não conhecemos o que acreditamos. A conclusão cética é que a autoridade da primeira pessoa é muito mais limitada do que supúnhamos,

mesmo estado físico e no mesmo *estado psicológico estreito*, mas com histórias e condições absolutamente diferentes de aprendizado do uso dos significados dos termos. Seus *estados psicológicos estreitos* são idênticos, porém os falantes querem dizer coisas diferentes com a mesma palavra. Mesmo que as palavras que pronunciam signifiquem coisas diferentes e que tenham crenças diferentes, ambos dizem a verdade quando pronunciam a palavra ‘água’.

Contudo, se os significados de suas palavras e as crenças expressas pelo uso dessas palavras são parcialmente determinados por fatores externos que os agentes ignoram, então, suas crenças e significados não podem ser *estados estreitos* no sentido que Putnam dá ao conceito. Os *estados psicológicos estreitos* são concebidos como os únicos estados internos no sentido cartesiano do termo, e, portanto, admitidos serem totalmente conhecidos do indivíduo. Como para Putnam uma diferença no sentido não implica uma diferença nos *estados psicológicos estreitos*, isso o leva a concluir que nenhum dos falantes sabe o que quer dizer ou pensa, pois o que determina o significado e a extensão é desconhecido do falante.

No entanto, existe um outro modo pelo qual os fatores externos podem entrar na determinação dos conteúdos da fala e do pensamento. Segundo Burge (1988 [1979]), sempre que uma pessoa estiver enganada, confusa ou mal informada sobre o significado público de uma palavra, ela está enganada, confusa ou mal informada sobre qualquer uma de suas crenças expressas pelo uso dessa palavra. A ‘compreensão completa’ do conteúdo seria a condição para acreditar no conteúdo. Como a ‘compreensão parcial’ é comum ou mesmo normal no caso de um vasto número de expressões em nossos vocabulários, frequentemente devemos estar errados sobre o que acreditamos. Para Burge, o conteúdo do pensamento de uma pessoa não é fixado pelo que se passa com ela ou pelo que é acessível a ela por introspecção. Aquilo que pensamos e significamos (queremos dizer) é determinado pelos hábitos lingüísticos de nossa comunidade.

Para Davidson, a forma como Burge propõe identificar as atitudes proposicionais compromete a autoridade da primeira pessoa, e é incompatível com o tipo de conhecimento que temos das nossas próprias mentes. Davidson acredita que há um modo de concebermos a determinação social dos estados mentais sem abrir mão da noção de autoconhecimento. Ele concorda com a tese externalista de que as atitudes proposicionais (o que eu penso e quero

---

pois nossas crenças sobre o mundo são muitas vezes verdadeiras, mas podemos facilmente estar errados sobre o que pensamos. A dúvida cética é que se não podemos ter crenças indubitáveis, então devemos questionar se o mundo é mesmo como acreditamos ser.

dizer com meus pronunciamentos) não são fixadas somente pelo que se passa com o indivíduo. Mas não concorda com a idéia de Putnam-Burge de que os fatores externos e sociais controlam os conteúdos da mente da pessoa. Ele rejeita a tese de Burge de que estaríamos limitados a dar às palavras da pessoa o significado que a mesma palavra possui na sua comunidade lingüística, e que o uso lingüístico social ditaria o que o falante quer dizer com suas palavras e fixaria seus estados psicológicos estreitos. Para Davidson:

“as atribuições ordinárias de significados e atitudes repousam em afirmações vagas e amplas sobre o que é ou não partilhado (lingüisticamente e de outras formas) por aquele que faz a atribuição, a pessoa a quem a atribuição é feita, e a audiência pretendida daquele que faz a atribuição.” (2001 [1987]: 28)

Dado o caráter holístico das crenças, as crenças não são átomos independentes, pois seus conteúdos necessariamente dependem dos conteúdos de outras crenças. Portanto, não pode haver uma regra rígida para a atribuição de uma única crença.

Davidson acredita que existe um outro sentido no qual os fatores sociais controlam o que o falante pode significar com suas palavras. Ele concorda que o significado de nossas palavras é fixado em parte pelas circunstâncias nas quais aprendemos e usamos as palavras. A questão é como estabelecer a conexão básica entre palavras e coisas. Davidson concorda com Putnam que essa conexão é estabelecida por interações causais entre pessoas e aspectos do mundo, pois nossas “disposições para reagir diferencialmente a objetos e eventos assim estabelecidas são centrais para a correta interpretação dos pensamentos e falas de uma pessoa” (2001 [1987]: 29).

Conforme o externalismo semântico de Putnam que se aplica às palavras de tipos naturais, se aprendo a palavra ‘água’ enquanto experimento H<sub>2</sub>O, a palavra se refere a substâncias com a mesma microestrutura da água. A referência da palavra é fixada pela história natural de sua aquisição. Conseqüentemente, aquilo que causa o uso da palavra determina o que ela significa. Davidson não acredita que a microestrutura comum a dois fenômenos seja necessariamente a similaridade relevante para determinar a mesma referência da palavra ‘água’. Em sua opinião, é um equívoco pensar que a dependência dos significados de fatores extramentais se limita a expressões de tipos naturais e indexicais como Putnam faz crer. É uma característica de toda linguagem e de todo pensamento que as ligações entre mente e mundo se dêem a partir de conexões causais estabelecidas. O fato dos significados serem dependentes do contexto é, na verdade, um fenômeno indissociável do caráter social da

linguagem. Pois, de outro modo, não teríamos como descobrir o que os outros pensam ou o que querem dizer com suas palavras. Como afirma Davidson: “todo pensamento e linguagem devem ter um fundamento em tais conexões históricas, e estas conexões restringem a interpretação da fala e dos pensamentos.” (2001 [1987]: 29). Davidson aceita a afirmação de Putnam-Burge de que:

“o conteúdo intencional das atitudes proposicionais ordinárias ... não pode ser explicado em termos de estados ou processos físicos, fenomenais, funcional-causais, computacionais ou sintáticos, que são especificados de forma não intencional e definidos puramente a partir do indivíduo em isolamento de seu ambiente físico e social.” (Burge *apud* Davidson 2001 [1987]: 30)

Ele concorda com Putnam que duas pessoas idênticas em todos os seus aspectos físicos relevantes podem querer dizer coisas diferentes com suas palavras e ter atitudes proposicionais diferentes. Essas diferenças se devem a diversidades ambientais que os dois agentes podem desconhecer.

Mostra, no entanto, que Putnam se engana ao supor que os dois agentes não sabem o que pensam ou querem dizer no sentido autoritativo da primeira pessoa. Como vimos, Putnam distingue os estados intencionais ordinários dos *estados psicológicos estreitos*, os quais não pressupõem a existência de qualquer indivíduo outro que não o sujeito de tais estados. Os *estados estreitos* são os estados que teriam a propriedade cartesiana de serem conhecidos de uma forma especial pela pessoa que os tem. Ele defende duas assunções: (1) se um pensamento é identificado pela relação a alguma coisa de fora da cabeça, ele não está completamente na cabeça; (2) se um pensamento não está completamente na cabeça, não pode ser apreendido pela mente no modo requerido pela autoridade da primeira pessoa. Nesse sentido, pode-se dizer que se duas pessoas têm a mesma estrutura cerebral e são molécula por molécula idênticas (mesmos pensamentos, *sense data* e disposições), seus *estados psicológicos estreitos* devem ser os mesmos. Assim, os conteúdos estreitos de duas pessoas são idênticos quando seus estados físicos não podem ser distinguidos.

Burge (1988) endossa essa premissa de Putnam. Para ele, as teorias materialistas da identidade cunharam a imagem de que o conteúdo de um evento mental poderia variar enquanto o evento permaneceria fixo. Ou seja, haveria um mesmo evento mental com conteúdos diferentes. Ele acredita que as teorias da identidade físico-mental foram tornadas implausíveis pelos experimentos de pensamento não-individualista das Terras Gêmeas que

mostram que se dois eventos mentais têm conteúdos diferentes, então eles são eventos diferentes.

Davidson não questiona a definição de *estado estreito*, o que ele questiona é a afirmação de que se um pensamento é identificado pela relação com objetos que estão fora da cabeça, então, ele não está completamente na cabeça, pois isso leva à conclusão de que as atitudes proposicionais ordinárias não estariam na cabeça, e, portanto, a autoridade da primeira pessoa não se aplicaria a elas. O fato das atitudes proposicionais serem identificadas, em parte, por elementos externos ao indivíduo não significa que não sejam estados do indivíduo. Uma apreciação dos fatores externos que entram em nossos modos comuns de identificar estados mentais não desacredita uma teoria da identidade físico-mental<sup>65</sup>.

Na tentativa de evitar os problemas do argumento das Terras Gêmeas, Davidson cria o seu próprio *doppelgänger*: o homem do pântano (*the swampman*). Ele imagina uma réplica física de Donald Davidson que surge de uma árvore morta num pântano atingida por um raio. Sua réplica se move e se comporta como se fosse o próprio Davidson, assumindo os mesmos hábitos e atitudes. A diferença é que a réplica não pode reconhecer nem lembrar nenhum elemento da vida de Davidson, pois nunca conheceu nenhuma dessas coisas, embora se comporte como se tivesse conhecido. O homem do pântano não pode querer significar com as palavras as mesmas coisas que Davidson significa, uma vez que as palavras que usa não foram aprendidas no mesmo contexto que poderia ter dado o significado correto. Não podemos afirmar que ele quer significar qualquer coisa com os sons que pronuncia ou que tem qualquer pensamento. A impossibilidade de ter pensamentos não se deve ao fato de ser um objeto criado acidentalmente ou artificialmente. Ele precisa de tempo para adquirir uma história causal na qual faça sentido para as frases que está falando.

O experimento do homem do pântano confirma a premissa de Burge de que se dois eventos mentais têm diferentes conteúdos eles são seguramente eventos diferentes. O argumento da crônica das Terras Gêmeas mostra que pessoas idênticas em todos os aspectos físicos relevantes podem diferir no que elas significam ou pensam. Mas existiria alguma coisa diferente nessas pessoas mesmo no mundo físico: suas histórias causais são diferentes e eles são objetos físicos únicos. Portanto, o simples fato de que os estados e eventos mentais

---

<sup>65</sup> Davidson contra-argumenta que os estados e eventos individuais não pressupõem *conceitualmente* qualquer coisa neles mesmos; porém, algumas de suas *descrições* podem pressupor.

ordinários sejam individuados em termos de suas relações com o mundo exterior não serve para desacreditar as teorias da identidade físico-mental como tais. O externalismo de certos estados e eventos mentais pode ser usado para desacreditar as teorias da identidade *type-type*, mas serve para sustentar as teorias da identidade *token-token*. Davidson conclui que Burge está errado ao pensar ter demonstrado que as teorias da identidade são implausíveis, pois é possível imaginar que as pessoas podem ser idênticas em todos os aspectos físicos relevantes enquanto diferem psicologicamente. Está é a posição do monismo anômalo de Davidson.

Vejam agora o obstáculo criado pela metáfora dos ‘objetos diante da mente’. Para Davidson, os externalistas que atacam o suposto caráter subjetivo dos estados psicológicos ordinários, muitas vezes, empregam a mesma metáfora: a idéia de que os objetos mentais – percepções, *sense data*, representações, atitudes proposicionais – são entidades que a mente pode ‘apreender’, ‘ter diante de si’ ou ‘estar consciente de’. Ou seja, permanecem presos ao ‘mito do subjetivo’, ao proporem que as atitudes proposicionais sejam identificadas por objetos (como proposições ou representações) ‘diante da mente’ e ao incorporarem objetos fora da mente do agente. Essa forma de subjetivismo do mental não permite saber onde localizar os objetos da mente: se estão na mente ou são vistos através da mente. A metáfora de ‘objetos diante da mente’ baseia-se na idéia de que os objetos mentais são ‘dados’ à mente, e que a atividade de pensar requer essas entidades mentais.

A descoberta de que os fatores externos participam da individuação dos estados mentais perturba essa imagem. A solução proposta por Putnam, e seguida por outros, foi criar a distinção entre os verdadeiros estados internos, *os estados estreitos*, que são ‘objetos diante da mente’, e acerca dos quais a mente mantém sua autoridade; e os estados ordinários (crenças, desejos, intenções e significados) que são comprometidos com suas conexões com o mundo público e social. Davidson recusa essa solução. Pois, como argumenta, se ter um pensamento fosse ter um objeto diante da mente e a identidade do objeto determinasse o que é o pensamento, então, seria sempre possível estar equivocado acerca do que se está pensando, pois seria impossível conhecer todos os aspectos de um objeto. Nesse sentido, seria óbvio imaginar que as pessoas ignoram os aspectos infinitos de todo objeto externo e desconhecem partes do mundo que participam da determinação de suas crenças.

Contudo, Davidson rejeita essa imagem. Para ele, ‘ter crenças’ não corresponde a ter objetos mentais, pois:

“os objetos aos quais relacionamos as pessoas na tentativa de descrever suas atitudes não necessitam ser, em nenhum sentido, objetos psicológicos, objetos a serem apreendidos, conhecidos pela pessoa cujas atitudes são descritas. (...) Ter uma atitude não é ter o objeto diante da mente.” (2001 [1987]: 37)

Davidson acredita que se nos livrarmos dessa metáfora, a autoridade da primeira pessoa deixa de ser ameaçada pelo externalismo do mental.

Em “Epistemology Externalized” (2001 [1990]), Davidson propõe o que ele denominou de *externalismo perceptivo*, uma forma de externalismo que se distingue dos argumentos de Putnam-Burge em relação à forma como os eventos mentais são individuados.

Burge argumenta a favor de duas formas de externalismo: o externalismo social e o que Davidson (2001 [1990]) chama de externalismo perceptivo. O *externalismo social* baseia-se na idéia de que os significados das palavras de uma pessoa e os conteúdos do seu pensamento dependem em parte das práticas lingüísticas da sua comunidade até mesmo nos casos em que o indivíduo se equivoca a respeito de que práticas lingüísticas são relevantes. O *externalismo perceptivo* enfatiza os modos pelos quais os conteúdos dos pronunciamentos e pensamentos dependem da história causal do indivíduo, em particular em sua conexão com a percepção.

Davidson não está interessado nos argumentos em favor do externalismo social, na forma como Burge o formula, pois considera equivocado supor que compreendemos e interpretamos a fala e o pensamento de um agente nos termos do que os outros significariam pelas mesmas palavras. Acredita que compreendemos melhor um agente quando o interpretamos do modo como ele deseja ser interpretado. A rejeição ao externalismo social de Burge não significa que Davidson ignora o papel dos fatores sociais no externalismo do mental. Porém, ele introduz o fator social de um modo que o conecta diretamente com o externalismo perceptivo, ao localizar o papel da sociedade nonexo causal que inclui a interação entre pessoas e o resto da natureza. Ambos estão de acordo que o que causa certos estados mentais é relevante para determinar o conteúdo desses estados mentais.

Davidson está interessado no *externalismo perceptivo* de Burge. Segundo Burge (2001 [1988]), temos um conhecimento perceptivo do mundo sem que tenhamos que nos certificar de forma independente que não estamos enganados de um modo ou de outro. No conhecimento perceptivo, o conteúdo de nossos pensamentos é parcialmente determinado por aquilo que causa o pensamento. Os conteúdos dos pronunciamentos e pensamentos dependem da história causal do indivíduo, em particular em sua conexão com a percepção. Para Burge, é

a interação com a água, por exemplo, que fornece as condições que tornam possível pensar em alguma coisa como a água, com as propriedades da água. A tese importante para o externalismo é sustentar o papel que a história das minhas relações com a água ocupa em tornar falso meu pensamento de que eu estou vendo água apesar do pensamento sobre a água. O falante que pensa que isto é água, provavelmente está certo, pois aprendeu o uso da palavra ‘água’ num ambiente aquoso, e o erro é sempre determinado a partir de um background de interação normal.

Essa forma de conceber o externalismo é corroborada por Davidson ao afirmar que os conteúdos de nosso aprendizado precoce e das frases básicas devem ser determinados pelo que está no mundo e pelo que nos leva a mantê-los como verdadeiros. Para Davidson, “é aqui que os laços entre a linguagem e o mundo são estabelecidos e os constrangimentos centrais de significado são fixados; e dada a estreita conexão entre pensamento e linguagem, comentários análogos vão para os conteúdos das atitudes”. (2001 [1990]: 200)

Não obstante, aceitar o externalismo perceptivo de Burge, Davidson não concorda com a resposta de Burge a dúvida cética sobre a existência do mundo e do autoconhecimento. Para Davidson, o externalismo não impõe nenhuma restrição ao privilégio do autoconhecimento. Pois se alguma coisa está sistematicamente causando certas experiências ou respostas verbais, é exatamente sobre isso que os pensamentos e pronunciamentos se referem. Conseqüentemente, isso exclui o erro sistemático. Se aceitamos o externalismo perceptivo, então, sabemos que não podemos estar sistematicamente enganados sobre a existência das coisas do mundo exterior.

Davidson fornece um exemplo esclarecedor desse tipo de relação causal implicada na formulação de um pensamento na percepção de um objeto físico. Ele afirma:

“um exemplo é o modo como o fato de um certo estado mental estar sendo tipicamente causado por ver vacas nos permite pensar ‘Existe uma vaca’ mesmo quando nenhuma vaca está presente. Mas aqui aparece um problema. O que determina o conteúdo de tais pensamentos básicos (e o que queremos dizer pelas palavras que usamos para expressá-los) é o que tem tipicamente causado pensamentos similares. Mas o que *tem* tipicamente causado esses pensamentos? Existem muitas escolhas, por exemplo, eventos que ocorreram antes de todas as vacas, ou eventos que estejam espacialmente mais próximos daquele que pensa do que qualquer vaca.” (2001 [1990]: 201)

Para Davidson, o modo como interagimos perceptivamente com o mundo exterior é determinante para a forma como classificamos o mundo. Somos projetados evolutivamente de

forma a agruparmos e classificarmos o mundo com aprendizado mínimo, a partir das similaridades das respostas que obtemos ou das causas de nossas experiências. Essas similaridades se tornam relevantes, pois consideramos simples e natural agruparmos essas respostas. Se não fosse assim não teríamos razões para afirmar que os outros estão respondendo aos mesmos objetos e eventos aos quais estamos respondendo também. Nas palavras de Davidson:

“a identificação dos objetos de pensamento repousa, então, numa base social. Sem uma criatura para observar outra, a triangulação que localiza os objetos relevantes num espaço público não poderia ocorrer. Não quero dizer com isso que uma criatura observando outra fornece a cada criatura o conceito de objetividade; a presença de duas ou mais criaturas interagindo uma com a outra e com um ambiente em comum é, na melhor das hipóteses, uma condição necessária para tal conceito. Somente a comunicação pode fornecer o conceito, pois ter o conceito de objetividade, os conceitos de objetos e eventos que ocupam um mundo partilhado, de objetos e eventos cujas propriedades e existência são independentes de nosso pensamento, exigem que estejamos conscientes do fato que partilhamos pensamentos e um mundo com outros”.  
(2001 [1990]: 202)

É por isso que não podemos resolver o problema dos conteúdos mentais do ponto de vista de uma única criatura. Pois se não partilhássemos um mundo de objetos e pensamentos com outros, não haveria critérios para selecionarmos uma causa e não outra como sendo a causa que fixa o conteúdo mental.

Como vimos, a posição externalista coloca um problema para a autoridade epistêmica da primeira pessoa. Se eu adoto a tese que os conteúdos mentais são determinados pelo exterior como posso admitir que tenho um conhecimento privilegiado de meus estados mentais? Como compatibilizar a suposição externalista de que os conteúdos das atitudes proposicionais de uma pessoa são parcialmente determinados por fatores ignorados pela pessoa com a autoridade da primeira pessoa? É a compatibilidade entre a posição externalista e a perspectiva da primeira pessoa que Davidson tenta resolver.

A tese da autoridade da primeira pessoa reconhece o fato que cada pessoa geralmente sabe o que está na sua própria cabeça de um modo que não pode saber o que está na cabeça de outra pessoa. A autoridade da primeira pessoa tem para Davidson duas características: produz um conhecimento que não está baseado na inferência ou na evidência e revela uma assimetria entre como conhecemos nossas mentes e como conhecemos a mente dos outros. A parte perdida do elo explicativo pode ser preenchida quando reconhecemos o modo como a

interação com outras pessoas determina, em parte, os conteúdos dos estados mentais. O conhecimento da própria mente é pessoal, mas o que individua aquele estado é, ao mesmo tempo, o que o torna acessível a outros, pois **o estado é individuado pelo jogo causal de três elementos: aquele que pensa; os outros com os quais ele se comunica; o mundo objetivo que eles conhecem e partilham.**

Como Davidson sugere (2001 [1987]), o insight de Wittgenstein – que o apelo à evidência e à observação são aspectos essenciais de nosso uso de certos predicados mentais – é uma resposta ao problema de como conhecemos outras mentes, mas não resolve a dúvida cética de como afirmações da primeira pessoa sem o suporte da evidência e da observação podem ter precedência sobre afirmações baseadas na evidência. A dúvida cética baseia-se no fato de que se uma crença é formada sem o suporte da evidência, a mesma crença não pode ser impugnada pelo questionamento da verdade ou relevância da evidência. Contudo, Davidson insiste que não podemos concluir que crenças sem suportes evidenciais são mais confiáveis que as crenças baseadas na observação. O ceticismo coloca o problema nos seguintes termos: sei o que estou buscando ao atribuir pensamentos aos outros, mas por que deveria pensar que são pensamentos que estou atribuindo a mim mesmo ao aplicar os mesmos predicados a mim mesmo? Como a evidência que uso no caso dos outros é pública, não há razão para atribuir pensamentos a mim mesmo do mesmo modo que atribuo aos outros, ou seja, não há razão para tratar meus pensamentos da mesma forma que trato os dos outros. Há uma assimetria na forma como conheço meus pensamentos e conheço a mente dos outros que resta a ser explicada. A possibilidade de que eu me engane sobre os meus próprios pensamentos não pode negar a presunção de que a pessoa sabe em que acredita: que “je puisse me tromper parfois” não significa que “je me trompe toujours”. Em geral, o fato de que eu acredito que tenho uma crença é suficiente para justificar a crença. Na maioria das vezes, eu sei em que estou pensando sem recorrer à evidência ou à observação. Isso faz com que o conhecimento da primeira pessoa, o autoconhecimento, mesmo que não seja nem infalível nem incorrigível, tenha uma autoridade sobre o conhecimento da segunda e da terceira pessoa. Davidson não apenas reconhece esse fato como também tenta fornecer uma explicação para o mesmo.

O externalismo de Davidson é uma revisão da imagem equivocada proposta por Descartes que analisamos na primeira sessão deste capítulo. Na tradição de pensamento cartesiana, a autoridade da primeira pessoa foi vista por muitos filósofos como a fundação

para o resto do conhecimento do mundo externo e de outras mentes O conhecimento é concebido como estando fundado essencialmente na experiência da primeira pessoa. A concepção do conhecimento como emergindo de um ponto de vista subjetivo *a priori* resultou na idéia de que poderíamos formar conceitos privados sobre o mundo. O pilar da contra-argumentação de Davidson reside na idéia de que, não importa se nos referimos a um conhecimento do exterior ou do interior, todo conhecimento proposicional requer a posse do conceito de verdade objetiva, e este tipo de conhecimento só é acessível a criaturas em comunicação com outras. Portanto, o autoconhecimento só é possível quando partilhamos um mundo de objetos num tempo e espaço compartilhado. O conhecimento do conteúdo de nossas próprias mentes não pode fornecer o fundamento para as crenças partilhadas, pois é logicamente independente de nossas crenças sobre o mundo exterior. Isso quer dizer que o conhecimento do conteúdo de nossas mentes não pode ser a base para o resto de nosso conhecimento. Como afirma Davidson,

“por causa do caráter holístico da crença empírica, então, é impossível que todas as nossas crenças acerca do mundo sejam falsas. (...) O que isso mostra é que não podemos nutrir crenças particulares falsas (ou verdadeiras) acerca dos objetos individuais ao menos que tenhamos muitas crenças verdadeiras sobre a natureza de tais objetos” (2001 [1990]: 195).

Não é possível que a maior parte de nossas crenças sobre o mundo sejam falsas. Pois os conceitos não são formados em primeiro lugar e depois descobrimos a que se aplicam; ao contrário, é a aplicação do conceito, seu uso na comunicação, que determina o conteúdo do conceito. *A condição que normalmente causa uma crença determina as condições na quais ela é verdadeira.* Nesse sentido, os conceitos são individuados por múltiplos critérios empíricos. Mesmo que possamos nos enganar eventualmente sobre o conteúdo dessas experiências, não nos enganamos sempre, na maior parte das vezes, estamos certos.

O externalismo perceptivo está baseado na triangulação entre o *conhecimento de outras mentes*, o *conhecimento do mundo externo* e o *conhecimento de nossa própria mente*. Esse conhecimento não ocorre de forma progressiva, ao contrário, emerge de forma holística. Davidson oferece **uma imagem intersubjetiva triádica da mente**. Ao definir conhecimento como, de início, intersubjetivo, inverte a lógica solipsista cartesiana e empirista que pretende construir uma visão do mundo a partir do indivíduo e seu ponto de vista subjetivo perceptivo, entendido como evento isolado e irreproduzível.

A matriz do pensamento intersubjetivo é o conhecimento da terceira pessoa, daquilo que está na mente do outro. Para desenvolvermos um conceito de objetividade, uma imagem dos objetos no espaço e no tempo públicos, devemos saber o que o outro pensa e para tal devemos estar em comunicação com o outro. É a partir dessa interação que desenvolvemos um conhecimento compartilhado do mundo externo. Há uma simultaneidade entre o conhecimento e as trocas comunicativas, pois não existem idéias antes de compartilharmos uma visão de mundo. O pensamento, entendido como a capacidade proposicional de formação de juízos ou de percepções proposicionais, no sentido de ‘perceber que’ e ‘sentir que’, não é possível senão no contexto de uma língua, de uma comunicação intersubjetiva. Assim definido, o conhecimento é uma capacidade interdependente, pois não é possível atribuir um conteúdo proposicional a criaturas que não tem capacidade de palavra. Portanto, a emergência do conhecimento compreende sempre, pelo menos, duas pessoas e uma série de eventos compartilhados. **A terceira face da tríade intersubjetiva é autoconhecimento, o conhecimento do que ocorre dentro de nós mesmos.**

## CAPÍTULO III

### AS FONTES EXTERNAS DO EU

#### 1. O Lugar Em Que Vivemos

Na tradição cartesiana, o eu foi identificado à consciência reflexiva, privada acessível apenas à reflexão em primeira pessoa. Habitamo-nos a concebê-lo como um ponto de vista coerente e unificado, uma perspectiva constante e estável. O eu seria a agência racional que não somente guiaria o comportamento humano, mas seria responsável pelo arranjo das informações cognitivas e perceptivas. As teorias internalistas do mental deram origem ao eu da autoconsciência como o *olho interno* que examina as representações privilegiadas da mente (Rorty 1994), e funciona como um princípio transcendental de unidade, um ‘meio combinador’ aonde se ancora a identidade pessoal. A metáfora do eu como um intermediário mental dominou boa parte das ciências sociais e psicológicas, e pode ser encontrada em certas formulações sobre o eu em psicanálise (ver Costa 1992, Bezerra 1994, Ferreira 2000).

O eu cartesiano é uma autarquia isolada e empobrecida em suas relações sociais, concebido como um índice transcendental sem conteúdo e sem história que acompanha a vida mental. Essa acepção não leva em consideração a **dimensão da alteridade**. Conforme vimos, a autoconsciência e o autoconhecimento se desenvolvem na presença da alteridade, seja na relação organismo-ambiente ou eu-outro. Para compreendermos a nós mesmos como uma referencia é preciso mais do que um índice semântico, um princípio formal, é necessária a experiência que inclui a experiência dos outros, dos objetos e do mundo físico e cultural.

O internalismo resultou em teses controvertidas acerca da gênese do eu, que têm sido revistas e criticadas pela psicologia e filosofia contemporânea, em particular pelas ciências cognitivas. Essas abordagens se opõem ao solipsismo e ao intrinsecalismo inerente ao consciencialismo que advoga a existência de um ‘quartel general’ ao qual se atribui responsabilidade por atos e comportamentos. De forma geral, a crítica se concentra em três idéias: a noção de self como agência central, identificação da mente à consciência, o solipsismo do mental.

Em oposição às abordagens internalistas que descrevem os conteúdos mentais a partir de relatos introspectivos da primeira pessoa, a concepção da mente externa (Land 2001)

amplia as fronteiras do eu para além da autoconsciência da criatura. O eu não é constituído apenas por aquilo que pode estar disponível à consciência, mas por todo campo de experiência externalizado, ecológico e não-consciente, por todos os instrumentos disponíveis no ambiente cultural, de palavras a objetos, que o organismo usa na construção ativa de suas fronteiras existenciais.

O que chamamos de mente e que vem sendo incluído na rubrica de ‘mundo interno’, ‘razão’, ‘realidade psíquica’ tem sido redescrito pelo naturalismo pragmático como algo que não se restringe às oposições conceituais de interno-externo, mental-físico, subjetivo-objetivo, organismo-ambiente. A naturalização da mente abandona as oposições essencialistas clássicas que concebiam os estados mentais como uma característica intrínseca dos humanos (Rorty 2000). Nossa única particularidade é a capacidade de estabelecer com outros objetos um conjunto de relações que nenhum outro objeto estabelece com o que quer que seja. Nossas propriedades, como consciência, linguagem e percepção, não são independentes das relações que mantemos com as outras coisas do mundo. Ao abandonarmos a distinção entre intrínseco e extrínseco, e entre realidade e aparência, nos livramos de preocupações sobre a existência de barreiras entre nós e o mundo.

A compatibilização do anti-essencialismo com as teses evolutivas darwinianas tornou possível ver todo o comportamento humano como sendo um contínuo com o comportamento animal, inscrevendo a humanidade numa temporalidade e contingência nunca antes pensada. A teoria evolucionista foi associada à noção de que é a linguagem, e não a consciência ou a mente ou a razão, o acento diferencial de nossa espécie, criando a imagem de continuidade do comportamento humano e da própria mente humana com o resto da natureza. Ao absorvermos a história evolutiva de Darwin, tornamo-nos pouco propensos a acreditar em versões transcendentais e atemporais de nós mesmos. Substituímos gradualmente a tentativa de nos vermos do lado de fora do tempo e da história, pela construção de um futuro melhor para nós mesmos. A metáfora evolutiva darwiniana nos levou a substituir a auto-imagem baseada na transcendência ou na metafísica pela auto-imagem como criaturas em evolução, produto de combinações biológicas aleatórias, originando um ser singular e contingente, com capacidade para ação criativa e responsabilidade moral.

Na ontologia darwiniana, o eu emerge do processo natural do desenvolvimento, cuja transformação contínua dá origem a um organismo capaz de fazer distinções entre um antes e

um depois, um dentro e um fora, com uma interação complexa com o ambiente. A permeabilidade da fronteira entre o interior e o exterior, como afirma Lewontin, é uma característica universal dos organismos, “que não se adaptam a nichos ecológicos pré-existentes, mas não cessam de fabricar e redefinir seu ambiente. (...) Esta co-evolução é quase sempre *topologicamente* contínua”. (Lewontin 2003 : 66) Esse processo de desenvolvimento é um processo interacionista concebido tanto da perspectiva filogenética da espécie quanto ontogenética. O processo filogenético é entendido como a *construção* permanente de uma rede complexa de interações entre o organismo e o mundo exterior (Lewontin 2003). Não pode haver organismo sem ambiente e não pode haver ambiente sem organismo. É claro que existe um mundo físico exterior ao organismo, que são as condições físicas a partir das quais os ambientes se constituem. O ambiente não preexiste ao organismo, ele é a penumbra produzida pelo organismo no meio exterior com o qual interage. Como sistemas abertos, exibem uma permeabilidade da fronteira entre o exterior e o interior que propicia fenômenos de internalização, assimilação, projeção, exteriorização e expulsão.

Em resposta aos modelos internalistas que privilegiam o indivíduo, os modelos externalistas valorizam a interação organismo-ambiente e buscam acentuar o aspecto interativo da gênese do eu. Nessa acepção, o eu é fruto de suas interações com a alteridade e sua origem está inteiramente associada ao modo como percebemos os outros e os objetos do ambiente. Elas podem ser divididas em dois eixos de análise: as *abordagens duais* e as *abordagens triádicas*. Ambas tomam a relação organismo-ambiente como ponto genético da constituição do eu, mas se diferenciam ao conceber a lógica da interação.

As *abordagens duais* incluem o ecologismo e o funcionalismo e se caracterizam pela ênfase na interação recíproca entre eu-ambiente ou eu-outro. A abordagem ecológica investiga o processo ontogenético de constituição do eu a partir dos aspectos corporais como um processo que se inicia desde o nascimento e, portanto, antes do aprendizado da linguagem e da aquisição de conceitos de eu. Descreve o eu a partir do desenvolvimento do organismo, dando ênfase à habilidade corpórea de selecionar e obter informação do ambiente. Sua contribuição é a investigação dos processos que constituem o *self agente*, concebido como uma projeção cognitiva que começa a fabricar sua mente por meio de sua ação corporal, antes mesmo da aquisição da linguagem. As abordagens funcionalistas focalizam o processo filogenético e concebem o eu como fruto do processo de aquisição da capacidade de narrativa. O

funcionalismo de Dennett imagina o eu como efeito da introjeção das histórias passadas de interação do organismo com o seu ambiente que se organizam em narrativas que constituem o *self narrativo*. O eu é a projeção de narrativas autobiográficas. Nessa acepção, o eu é *resultado* do mundo interpretado.

Pretendo analisar as *abordagens duais do eu*, exemplificadas pelo *self ecológico* de Neisser e o *self narrativo* de Dennett. Ambas as démarches propõem uma gênese externa do eu e concebem o eu como um modo de cognição, como um modo de obter informação sobre o mundo e nós mesmos. Mas, parecem insuficientes para explicar a gênese da autoconsciência e do autoconhecimento, ou seja, para explicar como o eu aprende a se referir a ele mesmo e aprende a conhecer seus próprios estados. Uma como outra, apesar de levar em consideração as interações com o ambiente e ter a externalidade como ponto zero do desenvolvimento, são incapazes de elucidar o modo pelo qual nos tornamos seres autoconscientes, isto é seres que podem auto-referir-se e meta-representar a auto-referência.

O funcionalismo e o ecologismo descrevem duas dimensões importantes da constituição do eu: a *existência corporal* e a *narrativa introspectiva*. Dito de outra forma, o eu que se autopercebe e o eu que se auto-representa a partir de sua ação no mundo. Contudo, apesar de serem aspectos centrais na constituição da identidade pessoal, estes processos não nos permitem entender como se produz o sentimento de continuidade do eu, a experiência de ser um self integrado que persiste no tempo. Ora, o eu integrado em seus diversos aspectos e expandido no tempo é o referente da autoconsciência da criatura. Assim, a emergência da autoconsciência está diretamente associada à capacidade de se ver como um eu que persiste no tempo.

A insuficiência das abordagens duais reside, a meu ver, no modo como interpretam a interação com a alteridade. Para ambos, a gênese do eu se explica a partir de uma lógica dual eu-outro ou eu-mundo. No entanto, as abordagens triádicas incluem o terceiro termo.

Como vimos, o externalismo perceptivo de Davidson apresenta uma gênese triádica do autoconhecimento. Para Davidson, a aquisição de conceitos, a aquisição do conceito de eu e a referência a si mesmo, são obtidas na relação com o intérprete. É necessária a mediação da alteridade, de um Outro que interpreta o Mundo, para que se forme a **gramática da interioridade**. Embora não aborde diretamente a questão do self, Davidson descreve o surgimento da mente e da capacidade de pensar. Podemos, então, empregar sua lógica ternária

para explicar como a referência do eu é dada por sua história interativa, que consiste na tríade: eu-mundo-intérprete.

Winnicott, na mesma direção, descreve o processo de gênese do eu de um ponto de vista que contempla uma lógica similar a de Davidson, a lógica dos fenômenos e dos objetos do espaço transicional. A experiência do sentimento de continuidade do eu, de que o eu persiste no tempo, está diretamente relacionada à interação no espaço transicional eu-objeto-outro. Para Winnicott, a gênese do eu nem está apenas no corpo e suas interações, nem apenas na narrativa autobiográfica. Para ele, a dimensão imaginativa é fundamental para a constituição da referência do sentimento de continuidade, e esta dimensão só se desenvolve e manifesta plenamente no espaço transicional. O *self expandido* no tempo está diretamente relacionado à aquisição da gramática da interioridade que é fornecida pela relação com um outro que interpreta o mundo para o eu.

Uma das preocupações centrais de Winnicott é a tentativa de compreender o que ele chama de paradoxo do ser humano, que são os processos de diferenciação entre eu e não-eu. A resposta ao paradoxo insolúvel winnicottiano é a criação de um espaço híbrido, a área transicional, para a qual concorrem tanto o mundo interno quanto o mundo exterior, formando uma continuidade entre os dois pólos da existência (Winnicott, 1975). A metáfora transicional winnicottiana situa a identidade pessoal na construção das fronteiras intersubjetivas.

## 2. A Gênese Dual do Eu

Onde começa e onde termina o self<sup>66</sup>? Quais são seus limites topográficos e qual é sua extensão temporal? São os contornos corporais que o definem? Ou são as narrativas existenciais que desenham suas fronteiras?

Existem duas formas em filosofia e psicologia de responder a essas questões. Uma delas é conceber o self como a unidade corporal mínima capaz de comportamento intencional,

---

<sup>66</sup> Para fins heurísticos, empregarei os termos “self” e “eu” como equivalentes. Como esclarece Costa, o termo “self é a forma substantivada do pronome reflexivo “se”. O self, o “si”, tem, aproximadamente, o mesmo referente do termo “eu”, qual seja, a identidade subjetiva. Falar de “identidade”, contudo, implica aludir à idéia de mesmidade. O quê, porem permanece idêntico no sujeito, se sabemos que jamais somos os mesmos ao longo da história pessoal? As teorias, deste aspecto revezam-se. Ora as expressões físicas do corpo, ora as mentais, são tidas como o invariável identitário. No primeiro caso, as particularidades anatomofisiológicas ou genéticas aparecem como aquilo que dá o toque singular a vida mental de cada um; no segundo, os eventos psíquicos como as fantasias; o inconsciente; os desejos, o equipamento cognitivo; as formas de obter prazer etc, são tomados como o elemento constante da subjetividade.” Costa, Jurandir Freire, *O self e a interioridade*, inédito, s/d.

cujos limites são os mesmos da pele corporal. O self mínimo (*minimal self*), como chama Gallagher (2000), é o ponto de origem pré-reflexivo para ação, o pensamento e a experiência existencial, do qual derivam tanto o sentido de ser um agente (*sense of agency*) quanto o sentido de posse do eu (*ownership*) para ação. O self mínimo é desprovido de extensão temporal e limitado ao que é acessível à consciência presente e imediata, ou seja, à consciência dos estados mentais e corporais. É o eu da perspectiva fraca da primeira pessoa de Baker. Nessa acepção, os limites do self são dados pela ação corporal intencional, na qual o sentido da agência e o sentido da posse coincidem e são indistinguíveis. O self mínimo engloba tanto o ‘self ecológico’ quanto o ‘self interpessoal’ de Neisser (1988). Nos dois casos, o eu é definido como o indivíduo ativo em seu ambiente imediato tanto na relação ecológica com os objetos quanto na relação interpessoal. Segundo Strawson (1997), este eu é uma presença mental que é ponto de partida da experiência em primeira pessoa, mas ainda não é uma pessoa. Chamarei esta entidade de **self agente**<sup>67</sup>.

A outra concepção é o self narrativo (*narrative self*), que requer continuidade ao longo do tempo e está diretamente relacionado à memória e à identidade pessoal. O eu narrativo é estendido no tempo e inclui as memórias passadas e intenções futuras, englobando os aspectos do ‘self expandido’ e do ‘self conceptual’ de Neisser (1988). Este eu emerge das histórias contadas, que estabilizam uma versão sobre nós mesmos que transcende o momento presente. Nessa acepção, o eu é uma ficção narrativa (Dennett, 1992), na qual quem somos nós depende das histórias que contamos e que os outros contam sobre nós. A construção da narrativa egóica se inicia na infância precoce com o surgimento do self conceptual, e continua durante toda vida na interação com a comunidade lingüística a que pertencemos. É o eu da perspectiva forte da primeira pessoa de Baker, que está diretamente vinculado à aquisição do conceito de ‘eu’. Chamarei esta entidade de **self narrativo**.

As duas acepções respondem a questões sobre como nos diferenciamos e nos individualizamos espacialmente do conjunto de outros objetos do mundo e como construímos a memória autobiográfica com a qual nos identificamos. Esses aspectos são geralmente abordados separadamente como pertencentes a eixos de análise de diferentes aspectos da identidade pessoal (ver A. Rorty 1976). Alguns autores buscam compreender como esses

---

<sup>67</sup> O self agente, como afirma Costa (inédito), é “uma instância originalmente ativa, isto é, capaz de escolher e preferir instrumentos, meios e finalidades para realizar ações que a favoreçam”.

aspectos interagem na constituição do eu. A clássica divisão de James – do ‘eu empírico’ e o ‘eu sujeito’ – e a conhecida classificação em cinco aspectos do eu de Neisser (1988) – self ecológico, self estendido, self interpessoal, self privado, self conceptual – são tentativas de apreender esses aspectos constitutivos da identidade. Ao propor dois padrões de resposta, não busco afirmar que existam apenas duas concepções de eu. As concepções sobre o eu são abundantes e não consensuais. Em “The self” (1997), Strawson, por exemplo, enumera 21 concepções de self. Proponho, ao contrário, que a partir dessas duas perspectivas correntes na literatura, investigarmos o processo de desenvolvimento emocional do self do ponto de vista do externalismo e dos distintos níveis de interação que o constituem.

Os estudos sobre o desenvolvimento psicológico do bebê (Gibson 2000; Rochat 2001; Neisser 1988) têm permitido reagrupar os aspectos do eu e redescrevê-los como processos mentais da constituição da identidade da pessoa. O eu não é uma parte da pessoa ou de seu cérebro, ou de seu corpo; o eu é a pessoa como um todo considerado de um determinado ponto de vista. Na démarche ecológica, tanto o ‘self agente’ como o ‘self narrativo’ são aspectos do eu que fazem parte do contínuo de experiências em complexidade crescente nos quatro primeiros anos de vida. Se, por um lado, a constituição do self agente pode ser vista como uma constante antropológica universal do desenvolvimento humano, por outro lado, as teias que constituem o self narrativo são totalmente determinadas pela autobiografia do indivíduo situado em dado contexto cultural.

Tanto Dennett quanto Neisser buscam dar conta da existência externalizada do eu, cuja centralidade, seja espacial ou ficcional, é continuamente recriada nos espaços interativos. O eu é a posição corporal resultante da ação do organismo no campo experiencial, e é a memória dessa vida existencial que servirá de dado primeiro para a construção de narrativas. Nesse sentido, para investigarmos o que é o eu e suas delimitações é necessário considerar a experiência existencial do *self agente* e tomar o campo existencial como agente causal e produtor de realidades discursivas e narrativas egóicas.

## **2.1 O Self Agente de Neisser**

A partir da tese de Gibson (1979) de que toda percepção envolve uma co-percepção do ambiente e do eu, Ulric Neisser (1988) distingue cinco formas em que o eu ao longo do desenvolvimento obtém informações que o especificam em seus diferentes aspectos. As

informações que especificam o eu estabelecem tipos de cognição que diferem em suas origens e histórias, mas que estão localizados no mesmo espaço e engajados nas mesmas atividades.

O que chamo de **self agente** inclui dois aspectos da auto-percepção descritos por Neisser (1988): o *self ecológico* e o *self interpessoal*. O *self ecológico* é o eu percebido em relação ao ambiente físico. Segundo Gibson, há um fluxo contínuo de informação óptica que ancora o sistema visual e seu ambiente imediato. A criança localiza o self no ponto de observação especificado pelo campo do fluxo óptico<sup>68</sup>.

É a partir da ação do agente que se estabelece um sentido de eu. O self ecológico não coincide com o corpo biológico. Qualquer coisa que se movimenta com o corpo ou qualquer objeto que se movimenta junto com o ponto de observação são percebidos como parte do eu. O que percebemos somos nós mesmos situados no ambiente e agindo em relação ao ambiente. Podemos modificar o ambiente imediatamente perceptível e os objetos cujos movimentos e alterações podemos controlar são experimentados como fazendo parte do eu. O eu e o ambiente são especificados pela informação que existe objetivamente e que nos permite perceber não apenas a localização do self ecológico, mas também a natureza de sua interação com o ambiente.

O *self interpessoal* é o self percebido por meio de sinais da relação emocional e da comunicação que são específicos de nossa espécie. É o eu engajado na interação imediata e não-reflexiva com outra pessoa. Assim como o self ecológico, parte de sua informação é obtida pelo sistema cinestésico que engloba estruturas permanentes no tempo. Contudo, essa informação só é obtida quando o bebê está engajado em uma interação interpessoal. Nas interações face-a-face, cada participante responde ao outro de forma imediata e coerente, na ação e na emoção. As respostas comunicativas características dos congêneres, como as expressões, gestos e vocalizações, especificam o self interpessoal. O aspecto interpessoal é produzido por meio das informações fornecidas por estas formas precoces de comunicação, pela orientação e pelo fluxo dos gestos expressivos do outro.

Desde o nascimento, o bebê se identifica com outros humanos e imita suas ações. Ele nasce com uma informação estrutural peculiar que o capacita a reconhecer que os membros de

---

<sup>68</sup> A autopercepção depende da duração da ação e de diferentes modalidades perceptivas. “Posso ver e sentir o que faço: as estruturas ópticas e cinestésicas que são dadas e que especificam as consequências do movimento são exatamente sincrônicas, e ambas coincidem com a atividade eferente por meio da qual a própria ação é produzida.” (Neisser 1988: 39)

sua própria espécie são como ele e a diferenciá-los dos objetos inanimados. O recém-nascido é capaz de diferenciar o eu do não-eu e exibe um sistema inato que possibilita desenvolver a imitação precoce. Esse sistema inato exibe um tipo de *intencionalidade corporal* que predispõe o self agente para a ação interativa. Como mostram Gallagher & Meltzoff (1996), o sistema é composto pela propriocepção, o esquema corporal e a imagem corporal, subsistemas que interagem entre si e são coordenados no contexto do comportamento intencional.

A *propriocepção* é um sistema de informação da posição dos membros e da postura corporal gerado pelos proprioceptores fisiológicos localizados em toda superfície corporal. Há uma consciência corporal dada pela propriocepção somática que é não-perceptual, pois ainda não se tornou objeto da percepção (Gallagher 2003). A consciência proprioceptiva é atentivamente recessiva e fornece uma consciência não-reflexiva. Mesmo na percepção tátil, a propriocepção somática contribui como parte recessiva da ação, de um corpo que não reflete sobre si próprio, pois intenciona o mundo e o corpo do outro. O *esquema corporal* é um sistema de habilidades e hábitos motores que permite o movimento, a manutenção da postura, e funciona sem a necessidade do monitoramento perceptivo. É composto de um conjunto de performances e processos não conscientes que controlam a postura e o movimento. E, embora participe e sustente a atividade intencional, opera abaixo da intencionalidade auto-referencial do agente. A *imagem corporal* é um sistema de percepções, atitudes e crenças, nas quais o objeto intencional é o corpo próprio. No self agente, há somente uma experiência perceptual do corpo próprio e de sua imagem corporal.

A propriocepção, o esquema corporal e os elementos perceptivos da imagem corporal são inatos e estão envolvidos na imitação invisível. A imitação invisível é a imitação que o bebê faz dos movimentos do outro usando partes do seu corpo que lhe são invisíveis, como os movimentos faciais. A imitação é um processo de tradução intermodal entre o sentido interno e o sentido externo que torna possível a experiência de conexão entre o eu e o outro. No processo de imitação, o bebê aprende a coordenar o esquema corporal com o movimento e os gestos visíveis que percebe do outro. Ele se engaja em atividades motoras e gestuais no esforço de se dirigir ao outro. Nessa experiência de conexão com o outro, o self agente experimenta a similaridade, analogia e sincronia entre seu corpo e o corpo do outro. O bebê não imita um gesto isolado, ele imita uma ação contextualizada. No entanto, o mecanismo de

imitação não demanda um sentido de continuidade do eu para que ele seja iniciado. A imitação é uma ação momentânea sem continuidade no tempo.

Nos primórdios, a imitação emerge da consciência perceptiva, ou seja, da experiência fenomenal, porque não há ainda autoconsciência, no sentido da perspectiva forte da primeira pessoa. O bebê pode ser dito consciente do objeto que imita, mas não tem consciência da experiência de imitar. Ele tem consciência da experiência, mas não tem consciência de que está tendo a experiência de imitar. Portanto, não tem ainda poderes discriminativos para distinguir a experiência, apesar de poder agrupar e classificar experiências de acordo com algum senso perceptivo e reagir a diferentes tipos de objetos e sons diferenciados. A experiência da imitação é uma experiência preste a ser digitalizada, que será futuramente conceitualizada e tornada complexa, vindo a ocupar lugar central nos processos de identificação. No entanto, não há introspecção na imitação. Não há autoconsciência de que se está imitando, pois não há posse do conceito de imitar. Estar consciente da qualidade da experiência é tornar-se consciente de que se está tendo uma experiência de um certo tipo, e, em relação ao ato de imitar, alguém não pode estar consciente sem compreender o que significa a ação de imitar.

O *self agente* exibe um sentido primitivo de eu que é inerente ao fluxo de experiência e que pode ser encontrado em qualquer estrutura subjetiva por meio da experiência de ser um ponto de vista. Contudo, é limitado ao que é acessível à consciência imediata e presente e envolve o sentido de ser um agente (*sense of agency*) e de posse do eu (*self-ownership*). Na ação motora, temos a consciência de que é o meu corpo que se move e de que sou eu a fonte da ação. Por exemplo, na apreensão de objetos, o sentido da agência e o sentido da posse são indiscerníveis. O autor da ação é a pessoa que está causalmente envolvida na produção da ação.

O sentido de ser responsável pela ação motora pode ser explicado em termos da consciência perceptiva ecológica construída pelo movimento e pela percepção. Pode-se dizer que existe um mecanismo muito mais precoce e corporificado do sentido do eu que emerge quando nos tornamos uma perspectiva. Há um acesso não-conceptual ao *self agente* cujo conteúdo não depende do uso da primeira pessoa. O desenvolvimento da autoconsciência é um processo gradual que implica a habilidade pré-existente de se diferenciar dos outros e de diferenciar pessoas de objetos. Todavia, não há ainda um agente meta-representacional capaz

de autoconsciência. De acordo com o princípio da imunidade, o acesso auto-referido do pronome da primeira pessoa só é possível ao ‘eu-mentalês’ que é capaz de comunicação lingüística e de usar o pronome da primeira pessoa nos discursos indiretos. Esse eu é construído a partir das trocas sociais e é uma aquisição que emerge da interação triádica entre o eu, o outro e o mundo compartilhado.

## **2.2 O Self Narrativo de Dennett**

Dennett (1996) criou uma concepção naturalista da mente. A mente é menos definida pelo material orgânico que a compõe, e mais definida pelo que é capaz de realizar. O que faz uma coisa ser uma mente é o que ela pode fazer, seu papel funcional. E o que a mente faz? A mente coleta, transforma e processa informações. Mas essa não seria uma capacidade única dos humanos. Possui mente qualquer organismo ou mecanismo no qual a atribuição de atitude intencional é bem sucedida do ponto de vista preditivo. A mente intencional não é intrínseca; é derivada da Mãe Natureza, do processo de pensamento sem pensador resultante da evolução por seleção natural que é capaz de autocontrole, consciência de si, comportamento intencional, e de funções mais complexas como atividade reflexiva, ação criativa e autodeterminação, ou seja, o tipo de mente exibida pelos humanos.

Para Dennett, não há um lugar específico do corpo humano que abrigue os atributos mentais. O corpo não é a sede da mente. A mente também não é o cérebro e não está situada num lugar específico do corpo. O processo evolutivo espalhou a mente por todo organismo, incorporando informações no corpo. As informações aproveitadas pelo cérebro não estão representadas numa parte central da mente. Não há um agente racional ideal que seria o responsável pela tomada de decisão e agiria como um self possuidor de um ponto de vista unitário e coerente do mundo.

Os organismos humanos são um *design* complexo projetado pelo processo de interação ambiental com capacidade para o que chama de controle de segunda ordem. São capazes de reavaliação e reflexão de seus próprios atos, capazes também de ações inteligentes e criativas. Mas, segundo Dennett, vários organismos, como os mamíferos, podem ser descritos como sistemas intencionais de ordem superior. Por que, então, não são considerados como uma pessoa? Qual seria a diferença? A diferença não estaria na configuração cerebral, nem no tamanho do cérebro, pois muitos mamíferos têm um cérebro com capacidade de estocar

informação tão grande quanto o nosso. É a habilidade de fala, a capacidade de comunicação, que permite a ampliação de um hábito já existente de “‘espargir’ (off-loading) o máximo possível suas tarefas cognitivas no nosso ambiente, fazendo literalmente uma espécie de extrusão de suas mentes no mundo que as cerca” (Land 2001: 230).

Desenvolvemos, pela herança cultural, o hábito de espalhar nossas mentes para o espaço ambiental em que vivemos e onde se encontram nossas ferramentas. Nossa mente resulta da interação corpo-ambiente. Várias espécies utilizam marcas do ambiente como parte dos processos cognitivos sem utilizar nenhum tipo de pensamento reflexivo. Os humanos utilizam coisas do ambiente para pensar. Ao longo dos anos, vão depositando marcos que funcionam como “gatilhos para seus hábitos”. Parte considerável de suas mentes está espalhada pelo habitat, numa prática de extrusão ativa. Dennett (1996) chama isso de “marcas de checagem”, que são tanto externas no mundo, quanto internas no cérebro.

Porém, os humanos têm uma mente adicional capaz de manipular símbolos. Desenvolvemos a capacidade de re-representação dos eventos e das formas de representação disponíveis que foram infinitamente ampliadas pela aquisição da habilidade de fala. Usamos uma série de próteses mentais que são exteriorizações de índices presentes em nosso cérebro. As ferramentas mais importantes obtidas da cultura são as palavras: para se comunicar, falar consigo próprio e criar o hábito do autocomentário. O autocomentário – que primeiro é um monólogo externo para depois ser internalizado – só se tornou possível depois de adquirirmos a capacidade de extrusão do mental. O monólogo interno foi uma aquisição posterior à capacidade de extrusão do mental, à aquisição da linguagem falada e ao uso das palavras como um hábito de re-representar para si mesmo atividades, como marcas de checagem e etiquetas internas.

O ‘si mesmo’ ou self resulta da atividade autonarrativa contínua desde a infância até o fim da existência. O desenvolvimento individual da pessoa é o aspecto mais relevante na aquisição de um self, tal como o processo de aprendizado e amadurecimento de um bebê que aprende uma estratégia de autocontrole<sup>69</sup>. O self dennettiano é “uma rede de crenças que não

---

<sup>69</sup> Ao mesmo tempo em que enfatiza o aprendizado, Dennett condena a idéia de que nascemos totalmente sem recursos. Há uma importância causal da história filogenética passada que equipou o organismo. O self autocriado não pode ser considerado totalmente responsável por si próprio, há sempre algum evento causal que está fora de seu controle e que tem efeitos causais sobre o self. Para atribuímos vontade livre e responsabilidade moral à pessoa, é necessário supormos a existência de um ‘elbow room’, um espaço de manobra, onde razões, crenças e desejos, autocriação e autocontrole possam ter eficácia causal.

para de se tecer, capaz de se autodescrever e produzir novos valores e regras de conduta” (Land 2001). O self é o centro de gravidade da narrativa (Dennett 1992), não devendo ser visto como um lugar real ou uma substância.

Dennett (1991) mostra que, na história evolutiva, a distinção entre o self e o outro, com a emergência de entidades individualizadas começa com a autopreservação: um ser que resiste à destruição e que transmite isso para sua geração. Para se autopreservar, a existência de limites e fronteiras são importantes. A distinção entre tudo que está dentro da fronteira fechada e tudo que está no mundo externo é um princípio biológico fundamental de distinção do self para o mundo. Esse senso mínimo de self não deve ser compreendido como um centro ou uma substância interna que faria uma supervisão central. O self mínimo é uma entidade abstrata que tende a distinguir, controlar e preservar porções do mundo, uma organização que cria e mantém fronteiras permeáveis e flexíveis. As delimitações do self são circunstanciais.

O que nos distingue é a forma como usamos as palavras. Nossa narrativa funciona como autoproteção, autocontrole e autodefinição. Contamos histórias sobre quem somos e controlamos as histórias que os outros contam sobre quem somos. O efeito de nossa narrativa é tentar posicionar um centro de gravidade da narrativa. O self não é uma entidade interna que vai se desenvolvendo e se torna uma agência central. É uma abstração que é efeito de nossa habilidade de contarmos histórias sobre nós mesmos, narrativas construídas em função da interação com o mundo externo. Inconscientemente, criamos um ou mais selves ideais fictícios e, então, há uma preponderância daquele que melhor oferece suporte para ocupar o papel de ‘senhor da mente’.

O que chamamos de mundo interno faz parte do conhecimento portátil que temos que ter sobre nós mesmos. A borda entre interno e externo não é delimitada pela pele, mas pela aquisição de um ‘saber como lidar’ com uma porção do mundo que é o ‘si mesmo da criatura’.

Falar de ‘mundo interno’ e ‘mundo externo’ só faz sentido depois que emerge da história evolutiva criaturas com um conhecimento sobre si mesmas. O critério deflacionado de diferenciação entre interno e externo proposto é que os componentes do ambiente interno são portáteis e os componentes do ambiente externo mudam de forma independente de nossa vontade. Em suma, nossa mente é o conjunto de todos os circuitos que garantem uma série de competências cognitivas em relação ao ambiente. Não há nenhuma descontinuidade entre o corpo, a mente, o ambiente, tudo se conhece automaticamente. A mente é a expressão de uma

função do organismo com capacidade de auto-regulação, auto-reprodução e auto-satisfação, e surge como uma resposta evolutiva da interação do organismo com o meio.

Essa concepção externalizada da mente se afasta das proposições que reduzem os processos mentais à consciência. Seria muito difícil imaginar a consciência estendida para fora da mente. Podemos imaginar processos cognitivos que estão para além da fronteira da consciência, pois nem todo processo cognitivo é um processo consciente: processos lingüísticos, aquisição de habilidades, retenção de memória. Existem processos externos mesmo quando há uma experiência interna, de primeira pessoa. Podemos perfeitamente imaginar e, até mesmo, experimentar, momentos em que temos uma experiência consciente em primeira pessoa a partir de uma exteriorização da mente.

Um dos problemas da mente externa dennettiana é supor uma relação desacoplada entre a mente e o ambiente, ou seja, como se os acoplamentos e desacoplamentos fossem circunstanciais. A mente expandida, proposta por Clark & Chalmers (1998), é extremamente dependente de certos arranjos com o ambiente. Os estudos sobre as patologias que afetam a memória tornam cada vez mais evidente o processo de acoplamento entre a mente e seu ambiente. Um dos maiores sofrimentos da perda da memória é a sensação de não familiaridade, ou seja, perdemos a capacidade de nos reconhecermos no ambiente que criamos, pois nos tornamos dependentes das informações do ambiente para estruturar nossas vidas.

Ao propor uma posição externalista ativa, Clark & Chalmers (1998) almejam aperfeiçoar o externalismo passivo<sup>70</sup> de Putnam e Burge. Segundo a abordagem externalista, como vimos, os conteúdos das crenças são causados historicamente por fatores externos. Se podemos dizer que estão certos em relação à externalidade do conteúdo, não está claro que esses aspectos externalistas ocupam um papel causal ou explicativo na geração da ação. No externalismo passivo, o ambiente imediato é considerado irrelevante, apenas o ambiente histórico é levado em consideração. Se a démarche externalista desfaz a intrinsecidade presente às abordagens internalistas do mental, e, conseqüentemente, torna inútil a redução do

---

<sup>70</sup> Clark e Chalmers chamam de externalismo passivo o externalismo que considera que os fatores externos (históricos e distais) formariam uma longa cadeia causal responsável pelas diferenças entre crenças. Ou seja, haveria uma externalidade causal do conteúdo das crenças determinada pela história causal de aquisição da crença. Contudo, os aspectos externos do **presente** seriam considerados não-relevantes, passivos, inertes causalmente, pois são por natureza distal e não ocupam papel causal em dirigir o processo cognitivo no **aqui e agora**. Parece, então, que estariam de acordo com o externalismo de Davidson que valoriza os aspectos proximais da interação perceptiva do organismo-ambiente.

mental ao cérebro, por outro lado, ela não explica o funcionamento da mente no tempo presente. Para o **externalismo ativo** (Clark & Chalmers 1998), os aspectos externos são ativos e ocupam um papel central no **aqui e agora**, pois são acoplados ao corpo humano e têm um impacto direto no organismo e no seu ambiente. As partes relevantes do processo interativo (organismo e ambiente) estão conectadas, e não são apenas as pontas de uma longa cadeia causal. Tanto os aspectos internos quanto os aspectos externos ocupam um papel causal relevante. O fato de vivermos espalhados num espaço híbrido de conteúdos internos e externos, onde a fronteira do eu está para além da consciência não é em si um paradoxo, mas sim uma das características mais interessantes de nossa mundanização: somos criaturas no mundo.

Isso altera a forma como concebemos os processos cognitivos. Pois a ação deixa de ser vista apenas como uma projeção do pensamento; ela é parte do pensamento. Usamos tanto nossa memória quanto nossos objetos como fontes externas ambientais. Nossa cognição é um processo contínuo com o ambiente. Nesse sentido, os objetos ambientais não são apenas algo que portamos e que nos auxilia, uma espécie de prótese amplificadora de nossa potência cognitiva, como o afirma Dennett. Os objetos que portamos são aqueles com os quais realmente formamos sistemas acoplados confiantes, pois, como afirma Clark & Chalmers, para que um sistema de acoplagem funcione como o centro da cognição é importante que de fato seja um sistema confiável. Os sistemas confiáveis configuram um conjunto de objetos que fazem parte do pacote básico de fontes cognitivas. Assim como o cérebro se desenvolve de um modo a complementar as estruturas ambientais externas, a mente também se desenvolve de um modo a complementar as estrutura externas. Nesse sentido, a mente estendida é um processo de cognitivo no qual o ambiente ocupa um papel central. Um dos instrumentos mais eficazes pelo qual os processos cognitivos se estenderam pelo mundo é a linguagem. O cérebro trata o envelope lingüístico que nos circunda desde o nascimento como estruturas confiáveis. Palavras e símbolos externos fazem parte da parafernália cognitiva que ajuda a constituir o pensamento humano.

A constituição do self narrativo só é possível porque somos *language-users*, ou seja, usamos a linguagem não apenas como uma extensão cognitiva, mas também como extensão de nossa identidade. Muito daquilo com o qual nos identificamos só pode ser formulado com palavras e está extremamente relacionado com as palavras que escutamos, ou seja, com aquilo

que é dito de nós mesmos pelo outro. Assim como o superego freudiano, que configura o complexo sistema de identificação do eu, que é um tipo de decantação das narrativas que escutamos durante o processo de desenvolvimento.

O *self narrativo* é um centro de gravidade, cuja única propriedade é sua posição espaço-temporal na rede de narrativas. As narrativas do eu são construídas pelo *self narrativo*, que é um intérprete que fornece seu próprio relato de suas atividades no mundo. Somos todos confabuladores, romancistas autobiográficos, contando e recontando a nós mesmos histórias de nossas próprias vidas, com pouca atenção para a veracidade das histórias.

Mesmo que haja margem para criatividade autobiográfica, construímos nossas narrativas de vida a partir de nossas experiências existenciais. Contudo Dennett não se preocupa em descrever ou explicar como as narrativas de vida se conectam com a experiência do *self agente*. Veremos a seguir como Winnicott oferece uma explicação das origens do self cuja continuidade da existência servirá de matéria autobiográfica.

### **3. Triangulação Eu-Outro-Objeto**

#### **3.1 O Self Expandido de Winnicott**

Herdeiro da tradição ferencziana que enfatiza o ambiente, Winnicott (1990, 2001) apresenta uma abordagem interacionista do desenvolvimento emocional como um processo de diferenciação do eu. A *experiência de interação* é o ponto de partida para emergência do self individualizado, numa relação contínua de oposição e cooperação com o ambiente.

Desde o nascimento, o bebê interage com seu nicho desenvolvimental (Reed 1995). Não existe, nas palavras de Winnicott (1993 [1952]), ‘o bebê’; só existe ‘o bebê e sua mãe’. O centro de gravidade do ser não se coloca no ser, mas na unidade mãe-bebê, na “estrutura de interação *de-dois-corpos*”. O que há, no início, é um par corporal, pois um bebê não pode existir sozinho, ele é essencialmente parte de uma relação. Na unidade *de-dois-corpos*, não há ainda o sentimento de habitar um corpo que tornará o bebê uma pessoa separada da mãe. Não há ainda um cerne localizado no corpo com uma membrana delimitadora, embora a mãe e o bebê possuam vivências corporais inteiramente diferentes e separadas, mesmo que interdependentes. Há uma co-evolução do par mãe-bebê, que funciona como um subsistema de troca constante entre os processos internos ao bebê e o mundo exterior.

A realidade exterior atinge o organismo por meio do filtro fabricado pelo próprio organismo, que no caso da espécie humana é o espaço transicional – um continuum entre o subjetivo e o objetivo. O ambiente humano, para Winnicott (1975), é o espaço transicional. É por meio do espaço transicional, que o bebê conhece a realidade externa. Isso não significa dizer que o mundo não tem um papel causal sobre o indivíduo. Ao contrário, a realidade exterior produz alterações sobre o indivíduo. E cada indivíduo reagirá a essa realidade de acordo com seu repertório de ações possíveis. A realidade exterior é aquilo que existe independente e fora do ambiente humano, aquilo que é não familiar e que se impõe ao organismo.

A constituição do self deriva do processo precoce de adaptação e ajustamento do ambiente ao bebê, oferecendo condições ideais para que o indivíduo possa criar o sentido de continuidade da existência<sup>71</sup>. A realidade externa é o espaço no qual se dão os fenômenos de externalização ativa do eu. Sem a experiência de externalização do self não há construção do sentimento de continuidade. O self será a unidade que garantirá a experiência de continuidade do ser. O self emerge da experiência corporal, não apenas da forma como representamos o corpo no psiquismo, mas, principalmente, como o corpo age no mundo. É por meio da *ação corporal* que se desenvolve o sentido de realidade exterior e o sentimento de continuidade da existência. Esse postulado naturalista, como sublinha Costa (2004), vai dar ao corpo “um papel proeminente na organização subjetiva”.

A experiência de ser uma entidade self que unifica as experiências do corpo físico<sup>72</sup> (ver Costa 2004) não é inata, é uma aquisição do processo de personalização. No ponto zero, mais precoce, o que há é a *natureza humana* com capacidade de sentir, se emocionar, perceber, pensar e agir. O corpo físico, tal logo nascemos, exhibe, quase como um automatismo da espécie, um funcionamento discriminativo e auto-regulador, como o *self agente* que exhibe a consciência perceptiva (*awareness*), que dota o corpo da capacidade de agir de forma seletiva

---

<sup>71</sup> Winnicott não busca as regularidades do processo do desenvolvimento humano. Está interessado na experiência como ruptura com o ambiente e no restabelecimento da conexão com o ambiente. Utiliza os casos em que a experiência de continuidade do eu fracassa, para mostrar como se desenvolve o sentimento de continuidade do eu. Propõe conceituações que ofereçam inteligibilidade às situações em que a experiência criativa fracassou, como nos casos de conduta anti-social e de desintegração do eu.

<sup>72</sup> Adoto a noção de *corpo físico* como proposta por Costa (2004) que inclui tanto a “soma das trocas metabólicas com o ambiente que visam à auto-regulação dos organismos individuais e à reprodução da espécie” quanto a dimensão fenomenológica da “intencionalidade corporal”. A solução de Costa resolve tanto os problemas do dualismo cartesiano ou do eliminativismo, que exclui a intencionalidade da matéria corpórea, como evita as imprecisões conceituais da fenomenologia. (Ver Costa 2004).

e discriminada. A essa estrutura mínima rudimentar, Winnicott dá o nome de psique-soma, uma “totalidade individual” expressiva. (ver Costa 2004)

Do ponto de vista emocional, Winnicott (1999 [1948]) descreve o desenvolvimento a partir de 3 aspectos: 1) a integração da personalidade; 2) o sentimento de habitar um corpo; 3) o contato com a realidade externa.

#### 1) A integração da personalidade

Em relação à integração da personalidade, o processo de desenvolvimento psíquico ocorre no sentido da não-integração para integração. Na origem, o que há é um **psique-soma** que se encontra num estado de não-integração, e que não é capaz de se distinguir do ambiente<sup>73</sup>. A integração do psique-soma é algo a ser conquistado no desenvolvimento e não um dado *a priori*. Psique-soma é um conceito usado por Winnicott para descrever um organismo com capacidades mentais e somáticas. A psique refere-se à elaboração imaginativa das funções corpóreas, das sensações, da percepção. É a representação mental do que é vivido corporalmente, é a memória das interações corporais. O soma é a atividade física.

É do processo de inter-relação do psique-soma com o ambiente que emerge a mente, que inclui a consciência, os processos cognitivos e a capacidade de simbolização, posteriormente enriquecida pela aquisição da linguagem. O desenvolvimento da mente é variável e depende de um ambiente que favoreça a continuidade da existência e o desenvolvimento do self. Dessa experiência, resultam a constituição do mundo interior com capacidade imaginativa e da mente com capacidade intelectual.

O estado de integração do psique-soma<sup>74</sup> e da mente deriva da aquisição do sentimento de unidade das experiências somáticas, imaginativas e intelectivas que envolvem a corporeidade mas que não têm por referência um lugar específico no corpo. A mente é

---

<sup>73</sup> Winnicott imagina um ponto zero de indiferenciação eu-não-eu a partir do qual o self corporal do bebê irá se desenvolver na direção da diferenciação. Contudo, proponho que essa ‘indiferenciação’ seja concebida menos como uma simbiose mãe-bebê e mais como uma continuidade entre dois pontos heterogêneos que, portanto, comportam um dualismo primordial eu-outro, pois o bebê nasce com mecanismos perceptivos inatos que o diferenciam, desde o primeiro momento, do ambiente. Sobre o *mito da indiferenciação* em psicanálise e a importância da presença dos aspectos interativos e sociais, desde o nascimento, para a emergência do self individualizado em Winnicott ver Pedro Salem (2006).

<sup>74</sup> A expressão ‘psique-soma’, como uma tentativa de superação do dualismo mente-corpo, é problemática, pois mantém o uso de expressões distintas para algo que, do ponto de vista ontológico, é concebido de forma monista. Winnicott parece afirmar que há uma realidade orgânica com propriedades distintas, mas não explica como ocorre a interação entre essas propriedades. Não há uma explicação coerente de como se dá a interação causal entre a função corpórea e a elaboração imaginativa da mesma. A relação causal é pressuposta, mas não é explicitada.

entendida como algo distinto da realidade material cerebral, a despeito da necessidade da integridade do cérebro para o exercício da atividade mental. A mente é a face representacional do cérebro. O desenvolvimento da mente não se restringe a aspectos inerentes ao organismo, pois está diretamente relacionado à interação ambiental. O funcionamento cerebral também se modifica com as interações ambientais. O cérebro inter-relaciona-se com o ambiente desde o início e sofre transformações oriundas da interação ambiental. A estrutura neuronal suporta modificações dentro do limite com o qual o cérebro foi projetado. A mente, ao contrário, não é pressuposta de início. É uma função projetada pelo processo evolutivo para representar o mundo, mas que é realizada apenas na interação do organismo com o meio. A mente winnicottiana existe pela externalização, e, tal como a mente das criaturas gregorianas de Dennett (1996), é capaz de extrusão e de se espargir pelo mundo.

## 2) O sentimento de habitar um corpo

O self corporal emerge da interação com a alteridade e está relacionado ao contexto de vida, portanto, seu contorno final é sempre provisório e relativo ao espaço externo no qual ele vive. A definição do que somos parece ser uma matéria oscilante com limites plásticos. O corpo ora é tratado como uma parte do self, ora é tratado como simplesmente self. A centralidade do self nem é dada pelos contornos do corpo, nem pela localização cerebral. O self é uma posição de equilíbrio na experiência existencial, *é um momento situacional do corpo*, resultante do movimento corporal no campo experiencial ambiental compartilhado<sup>75</sup>. Não é prévio à atividade. O eu que age nunca aparece antes da ação corpórea, constitui-se no exercício da ação e porta consigo a memória de suas interações.

A força vital originária do self é a expressão da motilidade corporal que, ao encontrar oposição ou obstáculo no meio externo, se subjetiva e se expressa de forma criativa ou agressiva. Ao nascer, todo bebê traz em si “um impulso inato poderoso e primário de auto-realização como a base da agressão”<sup>76</sup>. A agressividade original winnicottiana é sinônimo de motilidade. A força vital torna-se agressiva quando o bebê encontra um objeto externo, que não é o objeto da satisfação, e oferece resistência à sua ação. Para Winnicott, agressividade é evidência de vida e está na base do gesto espontâneo da criatividade primária.

---

<sup>75</sup> Sobre a noção de corpo humano como conceito situacional ver Patočka (1998).

<sup>76</sup> Cada bebê tem um potencial erótico instintivo biológico, que é mais ou menos o mesmo. Por outro lado, o componente agressivo é extremamente variável, porque depende da oposição dada pelo ambiente. Há uma imprevisibilidade na constituição subjetiva, pois não é possível saber como o ambiente irá se comportar. (Phillips 1988)

A pulsão criativa winnicottiana é o esforço hiperorgânico corporal como um centro autônomo gerador do movimento. Consiste numa ação e num movimento em direção a algo que existe fora do corpo. O que caracteriza a subjetividade corporal é o esforço. O esforço aparece na alteração dos estados corporais de atividade e passividade, relaxamento e excitação. O esforço do movimento faz fronteira a algo que resiste a ele. O bebê experimenta no contato seu esforço face ao corpo da mãe que lhe resiste. A espacialidade do corpo vivenciada no sentimento do esforço é diferente da espacialidade objetiva percebida pelos sentidos.

O bebê nasce com a capacidade para o movimento pulsional em busca do corpo materno, que Bowlby (1990) chama de impulso para o apego. Em seu movimento pulsional em direção ao mundo, o corpo age de modo a tornar o mundo familiar. O organismo humano nasce com o que Merleau-Ponty (1999) chama de intencionalidade motora – um atributo prévio, uma capacidade *a priori* da espécie. A *intencionalidade motora* é a capacidade de discriminação ativa, que revela um invariante fenomenológico da corporeidade. O corpo é afetado não apenas como um corpo que reage aos estímulos do ambiente, mas como um corpo que se dirige ao mundo. O estado de prontidão proprioceptiva para a ação e percepção é um equipamento para experimentar o ambiente. A corporeidade apresenta uma disponibilidade imediata para percepção, mas, como afirma Merleau-Ponty (1999), é preciso que o ambiente esteja lá para que o eu-corporal possa percebê-lo. É preciso que algo exterior se ofereça como obstáculo e que se preste à determinabilidade.

O self corporal inclui a dimensão mental representacional que é elaboração imaginativa das funções corpóreas, e a dimensão mental não-conceptual que é a própria atividade corporal. E pode ser entendido como uma entidade que comporta três dimensões: *corpórea*, *internalizada* e *externalizada*. 1) A dimensão mental não-conceptual do corpo inclui as propriedades proprioceptivas e exteroceptivas e o esquema corporal, que resultam em bem-estar, equilíbrio, conforto, prontidão para agir, esquemas motores e confiança nos hábitos corpóreos. 2) A dimensão mental conceptual da relação do corpo com os objetos e as pessoas é a interiorização das vivências relacionais, de hábitos culturais, com a criação de padrões imagéticos, como a imagem corporal. 3) A dimensão mental da relação do corpo com os objetos que é o modo como o self se exterioriza no ambiente por meio da ação corporal.

### 3) Contato com a realidade externa

Do ponto de vista da interação organismo-ambiente, o recém-nascido apresenta um potencial para se discriminar do ambiente. O ambiente facilitador origina-se a partir da identificação natural da mãe ao bebê. A construção do ambiente emocional depende de como a mãe se comporta, da sua história de vida e da memória das relações. O par mãe-bebê é construído por meio das interações corporais, que começam desde o útero materno e se estende pela forma como a mãe segura o bebê, o toca e acaricia, no manuseio da rotina dos cuidados corporais – o *handling*. A mãe cria um entorno adequado às dimensões, textura e temperatura do corpo do bebê, em constante mudança nos 10 primeiros meses de vida em função da variação de tamanho. O corpo do bebê se *desenvolve* no contato com o corpo do outro cuidador, por meio da comunicação sensorial. O toque participa dos processos de instauração do interior e do exterior do corpo pelo contato com a linguagem tátil do corpo materno. Ser tocado, como mostra Andrieu (2004), não é um ato passivo, a pele é um órgão receptivo que ativamente percebe as informações tangíveis e as incorpora como sensações internas. A totalidade do recém-nascido se expressa na pele – uma teia que cobre o corpo todo e que funciona como delimitador do espaço corporal, como uma interface fisiológica e psíquica dinâmica entre sujeito e mundo, como afirma Anzieu (1974)<sup>77</sup>.

A mãe estabelece uma identificação empática com o bebê, de modo a se modificar em função das alterações vivenciadas pelo bebê, num processo de co-transformação mãe-bebê. Tanto a mãe altera constantemente o meio do lactente para satisfazê-lo, quanto as capacidades de mobilidade, percepção e integração do corpo subjetivo do lactente também o transformam. O corpo da mãe funciona como um processo vinculado ao ser do lactente e as alterações de seu ser. Ela imprime no corpo do lactente a marca das primeiras experiências incorporadas como hábitos proprioceptivos. Os movimentos que a mãe repete na rotina de cuidado criam padrões de experiências, um estilo corporal. No primeiro momento, a mãe é identificável como um *processo* de transformações interna e externa acumulativas. Na concepção de Bollas (1987)<sup>78</sup>, o corpo materno é o *objeto transformacional*, que “tanto mantém a vida do bebê

---

<sup>77</sup> A experiência da taticidade não é esquecida ou reprimida quando a criança acede à linguagem, ela é incorporada como memória corporal. O sistema simbólico não é uma mediação entre o corpo e os sentimentos, mas pode se dizer que a aquisição da linguagem, hábitos culturais e práticas transformam a experiência do corpo fenomenológico. O aprendizado da gramática da exteriorização das sensações táteis é o aprendizado do modo de manifestação de uma prática social. O sentir só tem significação na relação do eu com o mundo.

<sup>78</sup> A noção de objeto transformacional é uma ampliação da matriz teórica winnicottiana. O modelo proposto por Bollas é de uma interação entre um sujeito ativo e um objeto ativo. “O objeto transformacional é identificado experimentalmente pelo lactente com processos que alteram a experiência do self. É uma identificação que surge

quanto transmite ao lactente, por meio de seu idioma próprio de maternagem, uma estética do ser que se torna uma característica do *self* do lactente.” (Bollas 1987: 27) O *self*, como diz Bollas (1987), é uma delimitação das fronteiras do eu, é o fator constitutivo do conhecido não pensado, um conhecimento mais existencial, em oposição ao conhecimento conceptual.

A mãe ambiente é aquela que apresenta o mundo no lugar mesmo em que o bebê está pronto para criá-lo, ou seja, apresenta o mundo ao bebê na medida em que este vai criando um repertório de experiências em que possa criar prontidões possíveis para o objeto, de modo que no primeiro momento cria-se a ilusão de que foi o bebê que criou o objeto. Se o desenvolvimento é satisfatório, a criança cria, fora de seu desejo, a mãe que está pronta para ser encontrada. Como afirma Winnicott, é esperado e até desejado que a mãe comece gradualmente a falhar em sua adaptação, favorecendo o contato do bebê, em pequenas doses suportáveis, com um mundo que lhe resiste, e oferece obstáculos. É assim que o bebê começa a experimentar a desilusão e a descobrir a mãe como um objeto fora de seu controle mágico.

O estágio de independência relativa do bebê é alcançado pela via do uso de seu primeiro objeto não-eu: *o objeto transicional*. O objeto transicional é aquele que introduz a dimensão da alteridade e que resiste à subjetivação. Ele tem uma dupla função: apresenta-se como um obstáculo que resiste à onipotência criativa e deixa-se alterar pelo *self*. O ‘eu’ constitui-se a partir da experiência de interação no espaço transicional e ao se deparar com aquilo que é ‘não-eu’. O ‘não-eu’ é aquilo que resiste e faz oposição ao ‘eu’. Uma posse pode vir a ser reconhecida como parte do ‘eu’ em algum momento da experiência. Todavia, a diferença entre o eu e suas posses é fundamental. A possibilidade de possuir algo é posterior a sua definição de *self*. O ‘eu sou’ precede ao ‘eu tenho’. Para Winnicott, o sentimento de posse só é possível no momento em que surge o *self expandido*.

O *self expandido* surge de uma primeira experiência externalizada e não de uma vivência prévia interna e privada. Ao agir e encontrar uma resistência, o *self* se volta para si. Depois disso, ele age, ou melhor, tem um ato criativo, e, assim, constitui, do ponto de vista funcional, a realidade externa e modifica a interna. Ao agir no mundo, o *self* se exterioriza, expressa-se do lado de fora e há uma modificação do corpo subjetivo. Quanto mais a criança modifica o objeto mais ela tem a experiência do sentido de continuidade do *self expandido*. O

---

de um relacionar simbiótico, onde o primeiro objeto é ‘conhecido’, não tanto por o colocarem em uma representação objetual, mas como uma experiência recorrente do ser – um conhecimento mais existencial, em oposição ao representativo” (Bollas 1987: 28).

processo do desenvolvimento é a crescente capacidade para tolerar o contínuo processo de ilusão-desilusão-re-ilusão. É assumido que a tarefa de aceitação da realidade nunca é completada, que nenhum ser humano está livre do esforço de relacionar realidade interna e externa. O fenômeno transicional faz com que a criança suporte as transições iniciais, de estar fundida com a mãe para ser algo separado dela, que são fundamentais para a constituição da subjetividade e objetividade.

A dimensão do reconhecimento é constitutiva da experiência da alteridade. É necessário um outro que exerça as funções de **holding, handling e apresentação do objeto**, para que haja o processo de constituição do self. A presença de um outro real – que é mãe suficientemente boa winnicottiana – instaura o espaço transicional para que emergja um sentido de self. Por isso, o mecanismo de constituição do self ocorre sempre na dimensão da externalidade e resulta da internalização de algo que foi primeiro experimentado no espaço transicional.

A fase transicional inaugura a experiência da alteridade O uso do objeto transicional é o primeiro ato criativo do lactente em relação a essas capacidades. Esse momento perceptivo registra a estrutura do conhecido não-pensado. A constituição do self corpóreo obedece menos a um modelo reacional de afetação, ou seja, um self que surge em reação a um objeto que o intenciona, como o desejo dos pais. Ao contrário, a constituição do self corpóreo se dá pela memória das sensações psicossomáticas de fusão relativas ao objeto transformacional. É a memória do processo ontogenético do eu não-conceptual.

Winnicott imagina, portanto, que o desenvolvimento emocional se dirige à conquista da experiência tridimensional. Realidade interna, externa, e transicional são perspectivas espaciais construídas pela posição situacional do self, e que separam experiências que se interpenetram, se sobrepõem, e se diferenciam, conforme a interação entre o self e o mundo. Na tópica tridimensional, há uma redefinição das fronteiras entre interno e externo. O que chamamos de “exterior” são os atos sensório-motores intencionalmente dirigidos ao objeto; o que chamamos de “interior” são as representações imagéticas e ideicas memorizadas. Exterior é o presente da atividade mental; interior, o passado da ação efetivada. O psiquismo tridimensional de Winnicott é um mapa metapsicológico do *self expandido* que se constitui pela ação corpórea, ao se exteriorizar no ambiente.

Tanto Winnicott quanto os interacionistas se interessaram pela propensão humana a se inclinar em direção ao ambiente, a buscar suportes ambientais, sejam cognitivos, emocionais ou culturais. Filogeneticamente, nossa espécie se desenvolveu criando e amplificando os instrumentos de suportes a cognição humana, e uma das mais complexas criações foi a habilidade de usar símbolos lingüísticos. O processo evolutivo nos tornou *language-users* (Ramberg 2002). Contudo, o processo de desenvolvimento do bebê mostra que antes de aprendermos a manipular sinais lingüísticos aprendemos a manipular objetos. Somos *objets-users*. A capacidade de usar símbolos lingüísticos deriva, em certa medida, da capacidade precoce do uso do objeto. Nesse sentido, podemos dizer que o self agente se constitui no processo de interação com os objetos, animados ou inanimados, presentes no espaço transicional do bebê.

No primeiro momento, o corpo do bebê é um corpo para mãe, investido em sua direção. Ao nascer, ou até mesmo antes do nascimento, o bebê e a mãe não existem separadamente. Definir a relação como de dependência não significa postular um estado inicial de indiferenciação. O bebê não é capaz de ter autoconsciência da diferença. É capaz de ter consciência de seus estados e de se discriminar do ambiente, mas não exhibe ainda uma autoconsciência como uma entidade separada e persistente no tempo.

A experiência inicial com a alteridade não é traumática, no sentido freudiano, pois ela é incluída na experiência subjetiva do lactente. Esse momento de ilusão é um momento no qual o bebê não reconhece a existência de nada externo a pseudo-díade. Esse estado é ‘puramente subjetivo’, uma área de ilusão, cujo correlato emocional é a experiência de onipotência. Contudo, a indiferenciação suposta não é corporal. São dois corpos subjetivos em contato e comunicação. O que o bebê ainda não é capaz de experimentar é a autoconsciência reflexiva do *self expandido*.

A partir do momento que o bebê se torna capaz de objetificar uma das partes, então, ela começa a existir. Ao adquirir a capacidade de tomar consciência das experiências corporais, o *self agente* será objetificado, tornando-se objeto do pensamento do próprio self. A constituição do *self expandido* é a constituição de um ponto de vista, por meio do qual percebemos, sentimos, observamos e refletimos sobre as experiências distintas e independentes do ser. É a constituição do *self expandido* que permite a objetificação e conceitualização do eu.

Para Winnicott, a aquisição da noção de objeto-objetivo, quer dizer, do objeto cuja realidade é não apenas percebida, mas concebida como exterior ao controle subjetivo do bebê, está relacionada à capacidade do bebê de usar objetos. A primeira posse do eu é o *objeto transicional* e está diretamente relacionada ao aprendizado de manipular objetos a partir de um corpo que se situa em relação ao objeto percebido. Ao manusear objetos, somos capazes de ações epistêmicas, quer dizer, são ações que alteram o mundo assim como ajudam a aumentar os processos cognitivos. Ao usar o mundo como parte do processo cognitivo experienciamos o mundo como parte de nós mesmos, Clark (1996) chama isso crédito epistêmico, pois incorporamos uma parte do mundo ao self, que passa a funcionar em sintonia com o self. A relação mãe-bebê é um sistema acoplado (*coupled system*) (Clark 1996) no qual todos os componentes do sistema exercem um papel causal ativo e controlam o comportamento.

Essa relação inaugural precoce mãe-bebê vai se expandir como uma experiência estética inaugural para a forma como o bebê usa os objetos. Clark mostra como esse tipo de processo é experienciado pelo sujeito como uma parte dele mesmo. Perder essa parte do sistema é perder uma parte de seu próprio corpo. Pois o corpo constitui um sistema acoplado com o ambiente. Nesse sentido, não é apenas a mente que é acoplada ao ambiente, mas também o corpo. Acoplamos nosso corpo ao ambiente pela ação corporal, por meio da propriocepção, dos esquemas corporais e perceptivos, e da imagem corporal.

Em seu processo de exteriorização em busca de suporte ambiental, o *self agente* desenvolve um complexo sistema de acoplamento com as ferramentas lingüísticas. A construção de narrativas, escritas ou faladas, é um dos usos mais complexos da linguagem. Por exemplo, a escrita – palavras desenhadas sobre o papel – permite objetificar nossos pensamentos e organizá-los ao **externalizá-los**. A externalização nas palavras (escrita ou oral), nas ações ou representações é a raiz do progresso interior, pois permite a expansão da memória. Portanto, a interioridade é causada pelas práticas de externalização do eu. Ao contrário do que afirma Dennett, ao dizer que o *self narrativo* das ficções autobiográficas não precisa estar ancorado em selves reais, considero que não somente a capacidade de construir autonarrativas está diretamente relacionada à experiência existencial do *self agente* situado em relação aos objetos e as pessoas, como afirmo também que é das experiências do self expandido no tempo – capaz de experimentar o sentimento de continuidade do eu, como diz Winnicott – que se constitui o *self narrativo*.

O *self expandido* de Winnicott é uma estrutura mental que evolui dessa dialética interativa entre o interno e o externo. O sujeito, aquilo que Dennett (1995) chama de personalidade (Land 2001), chega muito depois, quando o self já se constituiu a partir da negociação entre o vivido corporal e os cuidados ambiente-somáticos. Entre as capacidades transformacionais, a aquisição da linguagem materna é a mais significativa, pois resulta na alteração da percepção do mundo e do eu e introduz o sentido da personalidade, ou seja, a capacidade ter a perspectiva da primeira pessoa, de fazer inferências de segunda ordem e construir autonarrativas. Esse processo resulta na constituição da identidade pessoal como uma complexa rede de meta-representações sobre o eu.

A transição entre a experiência perceptiva *do self agente* para a experiência de conceitualização do *self narrativo* e a emergência da autoconsciência depende da constituição do *self expandido*, ou seja, do sentimento de continuidade do eu. O *self expandido* emerge da relação triádica eu-objeto-outro. É da relação do eu com o outro que apresenta o objeto que emerge a experiência de continuidade que servirá de âncora da referência do eu. Somente um ser que experimenta o sentimento de continuidade parece ser capaz de aprender a usar os conceitos necessários à autoconsciência e a se auto-referir de ‘dentro’. É a partir da aquisição da autoconsciência e da capacidade de se auto-referir que podemos contar histórias sobre nós mesmos.

A novidade do processo proposto por Winnicott reside no fato de que **a capacidade de se objetificar, de se meta-representar e de construir a referência sobre si só é possível na relação triangular bebê-mãe-objeto**. A mãe ambiente é o intérprete que apresenta o mundo dos objetos físicos e culturais ao bebê. É ela que apresenta também os objetos lingüísticos e os conceitos. Portanto, o conceito de eu, fundamental para a emergência da autoconsciência, é adquirido nessa relação triangular. É na relação triangular que a criança aprende a se auto-referir, a conhecer seus estados mentais e corporais e a construir a gramática da interioridade. A complexa rede de crenças e desejos que configuram a identidade pessoal depende, portanto, do ambiente social no qual a criança está inserida, e depende daquilo que as pessoas da comunidade pensam e acreditam. A **gramática da interioridade** é estruturada por normas corporificadas por meio de interações comunicativas constituídas por complexos padrões de interação causal com os outros e com o mundo, aquilo que nos torna *language-users* de uma comunidade de pessoas.

## CONSIDERAÇÕES FINAIS

O *Enigma de Kaspar Hauser* de Werner Herzog relata a história de Kaspar Hauser, uma enigmática pessoa que viveu até os 16 anos sem contato humano. O mundo de Kaspar Hauser se restringia a alguns poucos objetos e um ambiente precário e monótono, com poucos estímulos e poucas alterações. Até os 12 anos, vivia num pequeno quarto escuro com uma cama de palha e um cavalo de madeira e alimentos que eram deixados no local. Teve alguns contatos com um único homem que o ensinou a repetir uma mesma frase, com o qual nunca estabeleceu uma interação, no sentido real do termo. Ao sair de seu isolamento, Kaspar Hauser encontra um mundo que lhe é completamente desconhecido. Ele se dirige ao mundo com um olhar de perplexidade e espanto que denuncia sua não familiaridade com o ambiente mutável e de horizontes nunca antes percebidos. Não só desconhece objetos e paisagens, mas também o movimento, a perspectiva, o pensamento, a fala. Seu corpo não sabe se movimentar no espaço de objetos animados e inanimados. Sua mente não sabe operar com signos e significados. Desconhece também o contato com outro ser humano e se mostra inábil para conversações e trocas comunicativas.

Os anos de privação da interação humana comprometeram consideravelmente o desenvolvimento mental e emocional de Kaspar Hauser. Apesar de ter nascido com o sistema perceptivo que o permite se diferenciar do mundo, discriminar o ambiente e se auto-perceber, Kaspar Hauser não desenvolveu a consciência de si nem a capacidade de se auto-representar – funções adquiridas na interação com os congêneres. Não exhibe consciência de si nem capacidade de autoconhecimento. Quando é encontrado pela comunidade que o acolhe, Kaspar Hauser se assemelha a um ser autista, em isolamento, incapaz não só de se comunicar com os outros, mas também de exibir capacidade de conhecimento mínimo do mundo que o circunda e de seus próprios estados mentais e corporais.

Três incapacidades de Kaspar Hauser revelam a importância crucial da interação triádica com pessoas com mentes na emergência da autoconsciência e do autoconhecimento: incapacidade de imitar, incapacidade de se colocar em perspectiva, e incapacidade de exibir meta-pensamentos. Como nunca havia sido exposto às interações face-a-face, típicas dos jogos imitativos que os bebês são incitados a participar no contato com a mãe-ambiente, Kaspar

Hauser nunca aprendeu a usar a mímica facial nas exteriorizações de suas emoções. Também nunca aprendeu as vocalizações nem os gestos que fazem parte de toda interação humana. Mantinha sempre uma expressão de espanto revelando uma rigidez da mímica facial e desinteresse pela busca do contato atento e afetivo do outro. Sua face não expressava senão um sorriso social que vemos em um bebê de dois meses. Outra incapacidade de Kaspar Hauser é a ausência de perspectiva: ausência da experiência de movimento, de interagir com o objeto a partir de diferentes pontos de vista. Kaspar Hauser não teve a experiência de ser um ponto de vista subjetivo de onde emerge a capacidade de compartilhar com o outro – seja o intérprete de Davidson ou a mãe-ambiente de Winnicott – a atenção a um objeto físico ou lingüístico. Está também comprometida a capacidade de ter pensamentos sobre si mesmo, sobre o outro e sobre o mundo, e de compartilhar com os outros uma gramática de interpretação do um mundo em comum. No vocabulário winnicottiano, podemos dizer que em Kaspar Hauser três esferas necessárias ao estabelecimento do sentimento de continuidade do eu estão em desequilíbrio: (1) a ausência do sentimento de habitar um corpo dado pelo encontro com objetos e outros corpos humanos que imprimem a marca de uma experiência estética inaugural; (2) a ausência da capacidade de agir criativo que engendra a experiência de ser uma perspectiva entre outras e de ser capaz de modificar o objeto transicional; (3) a ausência da capacidade imaginativa dada pelo aprendizado da língua materna que introduz o bebê nos jogos interpretativos que permitem interpretar seu comportamento e o comportamento do outro.

Contudo, ao ser introduzido num espaço transicional de trocas comunicativas, vemos uma alteração significativa de Kaspar Hauser que, pouco a pouco, vai adquirindo consciência de si mesmo e de sua história. A história de Kaspar Hauser corrobora as teses que insistem que aquilo que consideramos como mais característico da personalidade só emerge da interação entre pessoas e um mundo compartilhado. A experiência existencial de ser um corpo fenomenológico externalizado em interação, apesar de ser condição necessária não é condição suficiente para a emergência do sentimento de continuidade do eu e, conseqüentemente, da autoconsciência da criatura. A capacidade de se referir a si mesmo emerge da interação com um outro que nos interpreta o mundo e nos apresenta a gramática da interioridade. Uma das cenas mais comoventes do filme é quando Kaspar Hauser toma consciência da significação de sua existência para o grupo social do qual faz parte. Em uma festa em que se torna alvo de

olhares curiosos, Kaspar Hauser experimenta a vergonha frente à autoconsciência da posição social que ocupa. Seu sentimento de constrangimento frente à compreensão de seu lugar social, como objeto de olhares que denunciam sua excentricidade e a curiosidade por sua anormalidade, revela que *a autoconsciência de nossa identidade social só é possível na interação entre o eu, o outro e o grupo*. Dito de outro modo, a interação triádica se altera significativamente quando os objetos que o intérprete apresenta e significa não são simplesmente um objeto físico ou um objeto lingüístico, mas um grupo social. Portanto, é apenas diante do grupo social que Kaspar Hauser pode tomar conhecimento de sua existência social, de sua identidade em seu grupo social. É só no contato com um outro que interpreta o mundo social para nós que somos capazes de construir significados sobre nossa existência pessoal e social.

Esta tese teve como objetivo investigar as origens externas da interioridade. A partir da metapsicologia de Winnicott, indicamos quais os processos psicológicos engendrados pela interação triádica que explicam a emergência do sentimento de continuidade do eu. Deixo para investigações futuras a proposição de uma metapsicologia da gênese externalizada da interioridade. Nesse primeiro momento, podemos apenas apontar o ponto genético inicial da tríade relacional eu, outro, objeto que emerge no espaço transicional.

## REFERÊNCIAS BIBLIOGRÁFICAS

- ABBAGNANO, N. (1988) *Dicionário de filosofia*. São Paulo: Martins Fontes.
- ABRAM, Jan (2000) *A linguagem de Winnicott*, Rio de Janeiro: Revinter.
- ANDLER, Daniel (2004) *Introduction aux sciences cognitives*, Paris: Ed. Gallimard.
- ANDRIEU, Bernard. *Le somaphore: naissance du sujet biotechnologique*, Belgique : Ed. Sils Maria, 2003.
- \_\_\_\_\_. *La nouvelle philosophie du corps*, Ed. Erès, Coll. Philosophies, 2002.
- \_\_\_\_\_. Mutologie de l'identité et externalité corporelle. In : *Multitudes*. Paris: Presses Universitaires de France, 2004.
- \_\_\_\_\_ (1999) Les théories matérialistes de l'esprit dans les neurosciences, in MISSA, Jean-Noël, *Matière pensante. Études historiques sur les conceptions matérialistes em philosophie de l'esprit*. Paris, Vrin.
- \_\_\_\_\_ (1998) *La Neurophilosophie*, Paris, P.U.F., coll. " Que sais-je ? " n°3373, 130 p.
- \_\_\_\_\_ (1995) "Éléments pour un matérialisme dynamique", *Revue philosophique de la France et de l'étranger*, PUF, 1995, n. 1, p. 71-82.
- \_\_\_\_\_ (2004) *Être touché. Sur l'haptophobie contemporaine*, Nancy, Les Éditions de la Maison Close.
- ANZIEU, Didier, 1974, *Le Moi-Peau*, Nouvelle Revue Psychanalytique, n.9, p. 195-208.
- AUDI, Robert (ed), (1998) *The Cambridge Dictionary of Philosophy*, Cambridge University Press.
- BAUMAN, Zygmunt. (1998) *O mal-estar na pós-modernidade*. Rio de Janeiro: Jorge Zahar.
- BAKER, Lynne Rudder, (1998) "The first-person perspective: a test for naturalism". In *American Philosophical Quarterly*, oct 1998 p327 (1).
- BERGSON, Henri. (1999) *Matéria e memória*, São Paulo: Martins Fontes.

BERNESTEIN, Richard (1971) *Praxis and action: contemporary philosophy of human activity*, Philadelphia: University of Pennsylvania Press.

BEZERRA Benilton. (1994) Descentramento e sujeito: versões da revolução copernicana de Freud. In: COSTA, Jurandir. *Redescrições da psicanálise: ensaios pragmáticos*. Rio de Janeiro: Relume-Dumará, 1994. P. 121-167

\_\_\_\_\_ (1996) A Paixão da Linguagem. Processos de Subjetivação Contemporâneos: Uma Leitura de Clarice Lispector. Tese (Doutorado). Instituto de Medicina Social. Universidade do Estado do Rio de Janeiro.

BOLLAS, Christopher, (1987) *The shadow of the object. Psychoanalysis of the Unthought Known*, New York: Columbia University Press.

BOUVERESSE, Jacques, (1976) *Le mythe de l'intériorité*, Paris, Munit.

BOWLBY, John, (1990) *Apego e Separação*, vol. 1, Rio de Janeiro: Martins Fontes.

BORRADORI, Giovanna, (2003) *A Filosofia Americana*, São Paulo: Editora Unesp.

BULLER, David J. (1997) Individualism and Evolutionary Psychology (or: In Defense of 'Narrow' Functions). *Philosophy of Science* 64:pp. 74-95.

BROOK, Andrew, (2005) *Externalism and The Varieties of Self-Awareness*, mimeo.

BURGE, Tyler, (1988 [1979]) "Individualism and the Mental", in Ludlow & Martin (ed.) (1988) *Externalism and Self-Knowledge*, Center for the Study of Language and Information, Stanford.

\_\_\_\_\_ (1988) "Individualism and Self-Knowledge", in Ludlow & Martin (ed.) (1988) *Externalism and Self-Knowledge*, Center for the Study of Language and Information, Stanford.

BUTTERWORTH, George, (1995) Self as Object of Consciousness, ROCHAT, Philippe, (1995) *The Self in Infancy: Theory and Research*, Amsterdam: Elsevier.

CASSAM, Quassim, (1998) "Introspection and Bodily Self-Ascription", in BERMÚDEZ, José Luis, MARCEL, Anthony & EILAN, Naomi, *The Body and The Self*, Cambridge: MIT Press.

CAVELL, Marcia. *The psychoanalytic mind: from Freud to philosophy*. United States of America: Harvard University Press, 1993.

CHALMERS, David, (1995) Facing Up to The Problem of Consciousness, *Journal of Consciousness Studies*, 2(3): 200-19, 1995.

CHANGEAUX, Jean-Pierre (1983) *L'homme neuronal*. Paris, Fayard.

CLARK, Andy, 1996, *Being there: putting brain, body, and world together again*, Cambridge, MIT Press.

CLARK, Andy & CHALMERS, David (1998) 'The Extended Mind', *Analysis* 58: 1:1998 p. 7-19

COSTA, Jurandir, (1992) *A Inocência e o Vício. Estudos Sobre o Homoerotismo*, Rio de Janeiro: Relume-Dumará.

\_\_\_\_\_ (1994) "Pragmática e Processo Analítico: Freud, Wittgenstein, Davidson e Rorty", In COSTA, J. *Redescrições da psicanálise: ensaios pragmáticos*. Rio de Janeiro: Relume-Dumará. P. 121-167.

\_\_\_\_\_, (1995) *A Face e o Verso: estudos sobre o homoerotismo II*, São Paulo, Escuta.

\_\_\_\_\_, (1998) *Sem fraude nem favor: estudos sobre o amor romântico*, Rio de Janeiro: Rocco, 1998.

\_\_\_\_\_. (2001) "A questão do sentido na teoria psicanalítica", in *Corpo, afeto, linguagem: a questão do sentido hoje*, BEZERRA JR., B. & PLASTINO, C.A., Rio de Janeiro: Rios Ambiciosos.

\_\_\_\_\_. (2002) Prefácio. In: *Agnes Heller entrevistada por Francisco Ortega*. Rio de Janeiro: EdUERJ, 2002.

\_\_\_\_\_. (2004) *O Vestígio e a Aura. Corpo e Consumismo na Moral do Espetáculo*, Rio de Janeiro: Garamond.

\_\_\_\_\_ *O self e a interioridade*, inédito.

COUTINHO, Ana Maria Ribeiro (1994) Repensando a questão da subjetividade em uma perspectiva pragmática. In: COSTA, Jurandir (1994) *Redescrições da psicanálise: ensaios pragmáticos*. Rio de Janeiro: Relume-Dumará, P. 61-117.

DAVIDSON, Donald (2001 [1984]) "First Person Authority" *Subjective, Intersubjective, Objective*, Clarendon Press, Oxford.

\_\_\_\_\_ (2001 [1988]) "The myth of the subjective", in *Subjective, Intersubjective, Objective*, Clarendon Press, Oxford.

\_\_\_\_\_ (2001 [1987]) "Knowing One's Own Mind", in *Subjective, Intersubjective, Objective*, Clarendon Press, Oxford.

\_\_\_\_\_ (2001 [1990]) "Epistemology Externalized", in *Subjective, Intersubjective, Objective*, Clarendon Press, Oxford.

\_\_\_\_\_ (1982) Paradoxes of irrationality. In: WOLLHEIM, Richard, HOPKINS, James (ed) *Philosophical essays on Freud*. Cambridge: Cambridge University Press, P. 289-305.

\_\_\_\_\_ (1992) *Mente, mundo y acción: claves para una interpretación*. Buenos Aires: Paidós.

\_\_\_\_\_. Thinking Causes. In: HEIL, John, MELE, Alfred (Ed.). *Mental causation*. Oxford: Clarendon Press, 1995. P. 3-17

DENNETT, Daniel. (1981a) "Conditions of Personhood", in *Brainstorms. Philosophical Essays on Mind and Psychology*, The MIT Press.

\_\_\_\_\_ (1981b) "Intentional Systems", in *Brainstorms. Philosophical Essays on Mind and Psychology*, The MIT Press.

\_\_\_\_\_ (1991) The origins of the selves, In: KOLAK, D. & MATIN, R. (eds) *Self & Identity. Contemporary philosophical Issues*. Macmillan.

\_\_\_\_\_ (1992) The self as a center of narrative gravity, in Kessel,F, Cole,P. & Johnson, *Self and consciousness, Multiples perspectives*, Hillsdale, NJ: Erlbaum.

\_\_\_\_\_ (1995) *Elbow room*, Cambridge: MIT Press.

\_\_\_\_\_ (1996) *Kinds of minds: toward an understanding of consciousness*, Nova York: Basic Books,.

\_\_\_\_\_ (2001) The Fantasy of First Person Science, draft.

\_\_\_\_\_ (2003) Who's On First? Heterofenomenology Explained, *Journal of Consciousness Studies*, 10.

\_\_\_\_\_ (2004) *Théorie Évolutionniste de La Liberté*, Paris : Odile Jacob.

DESCOMBES, Vincent, (2004) 'Un dedans derrière ce qui est le dedans', in *L'intériorité, Rue Descartes* n. 43, Paris, PUF, mars 2004.

DESCARTES, René (2000) *Meditações metafísicas*, São Paulo, Martins Fontes.

DEWEY, John. The influence of Darwinism on philosophy (1910), in *The structure of experience*, The philosophy of John Dewey, Chicago: University of Chicago Press, 1981.

\_\_\_\_\_. Nature, life and body-mind, in *The Later Works*, 1925-1953, vol. 1: 1925, Experience and nature, Illinois: Southern Illinois University Press, 1981.

DRETSKE, Fred, (1997) *Naturalizing The Mind*, Cambridge: MIT Press.

DUMONT, Louis, (1966) *Homo Hierarchicus – Le système de caste et ses implications*, Paris, Gallimard.

DUPUY, Jean-Pierre (1988) L'homme machine et les adorateurs du signifiant: Cybernétique et sciences de l'homme. *Le Débat*, Paris, v.49, mars-avril, pp.32-49.

\_\_\_\_\_ (1994) *Aux origines des sciences cognitives*. Paris, Éditions La Découverte.

EHRENBERG, Alain (2000), *La fatigue d'être soi. Dépression et Société*, Paris, Ed. Odile Jacob.

\_\_\_\_\_ (2004), « Le sujet cérébral », in *Esprit*, 11/ 2004, « Les Guerres du Sujet ».

EVNINE, Simon (1991) *Donald Davidson*, California: Stanford University Press.

FARIA, Paulo, (2001) Discriminação e conhecimento de si, in Ulysses Pinheiro, Marco Ruffino e Plínio Junqueira Smith (eds.), *Ontologia, Conhecimento e Linguagem: um Encontro de Filósofos Latino-Americanos* (Rio: Mauad/ FAPERJ, 2001).

\_\_\_\_\_, (2006) *Verdade Como Valor*, no prelo.

FAUSTINO, Silvia (1995) *Wittgenstein: o eu e sua gramática*. São Paulo: Ática.

FERREIRA, Cláudia Passos (1998) Sobre a trajetória da reflexão ética e as implicações do neopragmatismo na clínica terapêutica. *Saúde, sexo e educação*, n. 15, out//dez, 1998.

\_\_\_\_\_ (2000) *A causalidade psíquica em Freud*. Dissertação de Mestrado (Saúde Coletiva). Instituto de Medicina Social, UERJ.

\_\_\_\_\_ (2004) Causalité Psychique: La Machine Sémantique de Freud, in *Concepts, Revue Semestrielle de Philosophie*, n. 9. Belgique : Les Éditions Sils Maria.

\_\_\_\_\_ (2006) Conflit Psychique. In ANDRIEU, Bernard, (org) *Dictionnaire du Corps : en sciences humaines et sociales*. Paris : CNRS, 98-99.

FISSETTE, Denis & POIRIER, Pierre (2000) *Philosophie de L'Esprit : États des Lieux*, Paris : Vrin.

FLANAGAN, Owen, (1991) *The Science of Mind*, MIT Press.

FONAGY, Peter et al. (2005) *Affect regulation, mentalization and the development of the self*. London: Karnac.

GALLAGHER, Shaun (2000) Philosophical conceptions of the self: implications for cognitive science, *Trends in Cognitive Science* 4 (1):14-21.

\_\_\_\_\_ (2003) Bodily-self-awareness and object perception, in *Theoria et Historia Scientiarum: International Journal for Interdisciplinary Studies*, 7 (I) – in press.

GALLAGHER, S. & MARCEL, A.J. (1999) The self in contextualized action, in *Journal Consciousness Studies* 6, 4-30.

GALLAGHER, Shaun & MELTZOFF, Andrew (1996) The Earliest Sense of Self and Others: Merleau-Ponty and Recent Developmental Studies, in *Philosophical Psychology*, 9: 213-236.

GIBSON, Eleanor & PICK, Anne, (2000) *An Ecological Approach to Perceptual Learning and Development*, Oxford: Oxford University Press.

GIBSON, Eleanor, (1995) Are We Automata?, In ROCHAT, Philippe (ed.), *The Self in Infancy: Theory and Research*, Amsterdam: Elsevier.

GIBSON, James (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin; reprinted, (1986). Hillsdale, NJ: Erlbaum.

GLOCK, Hans-Johann, (1997) *Dicionário Wittgenstein*, Rio de Janeiro: Jorge Zahar Editor.

- JACOB, Pierre (2004) *L'Intentionnalité. Problèmes de Philosophie de L'Esprit*, Paris : Odile Jacob.
- JAMES, William, (1992a) *Writings 1878-1899*, edited by Gerald Myers, New York: The Library of America.
- \_\_\_\_\_, (1992b) *Writings 1902-1910*, edited by Gerald Myers, New York: The Library of America.
- JONES, Doug, (1999) *Evolutionary Psychology*, Annu. Rev. Anthropology, 1999, 28:553-75.
- KIM, Jaegwon (1998), *Mind in a physical world: un essay on the mind-body problem and mental causation*. Cambridge: Massachusetts Institute of Technology.
- KRIEGEL, Uriah, (2006) *Self-Consciousness: A Tutorial*. Mimeo.
- \_\_\_\_\_, (2005) "Naturalizing Subjective Character", In *Philosophy and Phenomenological Research*, 71 (2005) 23-57.
- LAND, Marcelo (2001), *A mente externa*, Rio de Janeiro: Garamond.
- LEWONTIN, Richard, (2003). *La triple hélice. Les genes, l'organisme, l'environnement*, Paris, Éditions du Seuil.
- LUDLOW & MARTIN, (ed.) (1988) *Externalism and Self-Knowledge*, Center for the Study of Language and Informatio, Stanford.
- MELTZOFF, Andrew & MOORE, M. Keith, (1995) A theory of the role of imitation in the emergence of self, in ROCHAT, Philippe (ed), (1995) *The Self in Infancy*, Amsterdam : Elsevier.
- MENDONÇA, Wilson, (1997a) "Descartes, Individualismo e Autoconhecimento." In S. Fuks (org.) *Descartes: um legado científico e filosófico*. Rio de Janeiro: Relume Dumará, 1997, pp. 217-228
- \_\_\_\_\_, (1997b) "Quão Individualista Deve Ser a Psicologia Intencional?" Palestra proferida no Centro de Ética e Filosofia da Mente/ IFCS/UFRJ. Mimeo.
- MERLEAU-PONTY, Maurice (1999) *Fenomenologia da Percepção*. São Paulo: Martins Fontes.

NAGEL, Thomas, (1992) "What's Like to Be a Bat?" in *Mortal Questions*, Cambridge University Press.

NEISSER, Ulric (1988) Five kinds of self-knowledge. *Philosophical Psychology* 1, 35-59.

ORTEGA, Francisco. "Da ascese à bio-ascese, ou do corpo submetido à submissão ao corpo", in RAGO, Margareth & NETO, Alfredo (org) *Imagens de Foucault e Deleuze : ressonâncias nietzschianas*, Rio de Janeiro, DP&A, 2002.

Patočka, Jan. (1998). *Body, community, language, world*, Illinois, Open Court Publishing.

PEREIRA, Roberto Horácio de Sá, (2006) *Tugendhat On Self-Consciousness*, mimeo, inédito.

PHILLIPS, Adam (1988) *Winnicott*, Cambridge: Havard University Press.

PUTNAM, Hilary (1995 [1975]) The Meaning of 'Meaning', in *Mind, Language and Reality, Philosophical Papers*, vol.2, Cambridge University Press.

\_\_\_\_\_ (1992) *Razão, Verdade e História*, Lisboa, Publicações Dom Quixote.

\_\_\_\_\_ (1994) "How Old is the Mind", in *Words and Life*, Massachusetts, Havard University Press.

\_\_\_\_\_ (1999) *Threefold cord: mind, body and world*, Columbia University Press.

\_\_\_\_\_ (1999) "Are appearances 'qualia'?" in *Threefold cord: mind, body and world*, Columbia University Press.

\_\_\_\_\_ (2000) *Pragmatism: an open question*, Oxford & Cambridge: Blackwell.

RABINOW, Paul. (2002) *Antropologia da Razão*, Rio de Janeiro: Relume Dumará.

RAMBERG, Bjorn, (2002) Post-ontological philosophy of mind: Rorty versus Davidson, in

BRANDON, R, (2002) *Rorty and his critics*, Blackwell Publishers.

\_\_\_\_\_ (1997). *Naturalizing idealizations: pragmatism and the interpretivist strategy*, inédito, exemplar fotocopiado, p. 3.

REED, Eduard, (1995) Becoming self, in ROCHAT, Philippe (ed), (1995) *The Self in Infancy*, Amsterdam : Elsevier.

ROCHAT, Philippe (ed) (1995) *The Self in Infancy*, Amsterdam : Elsevier.

ROCHAT, Philippe, (2004) *The Infant's World*, New York: Harvard University Press.

- \_\_\_\_\_ (2002) Naissance de La Co-Conscience, In *Intellectica*, 2002/1, 34, 99-123.
- \_\_\_\_\_ (2003) Five Levels of Self-Awareness as They Unfold Early in Life, In *Consciousness and Cognition*, 12 (2003) 717-731.
- ROCHAT, Philippe & STRIANO, Tricia, Perceived self in infancy, in *Infant Behavior and Development*, 23 (2000) 513-530.
- RORTY, Amélie (edt), (1976) *The Identities of Persons*, Berkeley: University of California Press.
- RORTY, Richard, (2000) “Um mundo sem substâncias ou essências”, in *Pragmatismo: a filosofia da criação e da mudança*, MAGRO, C. & PEREIRA, A. (org), Belo Horizonte: Ed. UFMG.
- \_\_\_\_\_, (1988) *A filosofia e o espelho da natureza*. Lisboa: Publicações Dom Quixote.
- \_\_\_\_\_. (1994) “Droits de l’homme, rationalité et sentimentalité”, in *Ambiguïté et limites du postmodernisme*, Librairie Philosophique J. Vrin, Paris, 13-36.
- RUSSO, Jane & POCIANO, Edna, (2003) *O Sujeito da Neurociência - da naturalização do homem ao re-encantamento da natureza*, artigo inédito.
- RYLE, Gilbert, (1984), *The Concept of Mind*, Chicago: The University of Chicago Press.
- SALEM, Pedro, (2006) *A Gramática da Quietude. Um Estudo sobre Hábito e Confiança na Formação da Identidade*. Tese (Doutorado). Universidade do Estado do Rio de Janeiro. Instituto de Medicina Social.
- SALEM, Pedro e COSTA, Jurandir. Sobre a confiança em Balint. In: *Textura: Revista de Psicanálise*. São Paulo: Publicação das reuniões psicanalíticas, 2003.
- SEARLE, John, (2004) *Liberté et Neurobiologie*, Nouveau Collège de Philosophie, Paris : Bernard Grasset.
- SELLARS, Wilfrid, (1997) *Empiricism and Philosophy of Mind*, Harvard University Press.
- SENETT, Richard (2002). *A corrosão do caráter*. Rio de Janeiro: Record.
- SHOEMAKER, Sidney, (1996), *The First-Person Perspective and Others Essays*, Cambridge: Cambridge University Press.

- SOUZA, Octavio. Trauma, defesa e criatividade. Mimeo, s/d.
- STRAWSON, Galen. (1997) 'The self', *J. Consciousness Studies* 4, 405-28.
- TAYLOR, Charles, (1997) *As fontes do self. A construção da identidade moderna*, São Paulo: Edições Loyola.
- TOMASELLO, Michael (1995) "Understanding The Self As Social Agent", ROCHAT, Philippe (ed) (1995) *The Self in Infancy*, Amsterdam : Elsevier.
- VARELA, Francisco, THOMPSON, Evan & ROSCH, Eleanor (1994) *L'inscription corporelle de l'esprit. Sciences cognitives et expérience humaine*, Paris : Seuil.
- VARELA, Francisco & SHEAR, Jonathan, (1999) "First-Person Methodologies : What, Why, How ?" *Journal of Consciousness Studies*, 6, n.2-3, 1999, pp. 1-14.
- VARELA, Francisco (1996), *Invitation aux sciences cognitives*, Paris, Ed. du Seuil.
- \_\_\_\_\_ (1996a) «Neurophenomenology: a methodological remedy to the hard problem», *Journal of Consciousness Studies* 3, pp. 330-350.
- \_\_\_\_\_ (2004) *Quel Savoir pour L'Éthique ? Action, Sagesse et Cognition*, Paris : La Découverte.
- VIDAL, Fernando & ORTEGA, Francisco. (2004) 'O sujeito cerebral. Impacto das neurociências na sociedade contemporânea'. Projeto PROBAL (CAPES & DAAD), Rio de Janeiro.
- VIDAL, Fernando, «Brains, bodies, selves, and science. Anthropologies of identity and the resurrection of the body», *Critical Inquiry*, sous presse.
- \_\_\_\_\_, (2002) L'histoire et les enjeux des neurosciences. L'avènement du "sujet cerebral" au XX siècle. Mimeo s/d
- WILSON, Robert, (2004) *Boundaries of The Mind. The individual in the Fragile Sciences*, Cambridge: Cambridge University Press.
- \_\_\_\_\_ (2005) "Meaning Making and The Mind of The Externalist", in MENARY, Richard (ed), *The Extended Mind*, (2005), Ashgate.
- WILSON, Robert & CLARK, Andy, (2006) How to Situate Cognition: Letting Nature Take Its Course, To Appear In M. Ayded & P. Robbins, *The Cambridge Handbook of Situated Cognition*.

WINNICOTT, Donald. (1993) “Desenvolvimento emocional primitivo” (1945), in *Da pediatria à psicanálise*, Rio de Janeiro: Francisco Alves.

\_\_\_\_\_. (1993) “A mente e sua relação com o psique-soma” (1949), in *Da pediatria à psicanálise*, Rio de Janeiro: Francisco Alves, 1993.

\_\_\_\_\_. (1993) Ansiedade associada à insegurança (1952) in *Da pediatria à psicanálise*, Rio de Janeiro: Francisco Alves.

\_\_\_\_\_. (1993) “Preocupação materna primária” (1956), in *Da pediatria à psicanálise*, Rio de Janeiro: Francisco Alves.

\_\_\_\_\_. (1975) *O brincar e a realidade*, Rio de Janeiro: Imago.

\_\_\_\_\_. (1990) *O ambiente e os processos maturacionais*, Porto Alegre: Artes Médicas.

\_\_\_\_\_. (1993) “Agressão e sua relação com o desenvolvimento emocional” (1950), in *Da pediatria à psicanálise*, Rio de Janeiro: Francisco Alves.

\_\_\_\_\_. (1999) Le premier contact avec la réalité externe (1948) in *L'enfant, la psyché et le corps*, Paris, Éditions Payot & Rivages.

\_\_\_\_\_. (2001) *A família e o desenvolvimento individual*, São Paulo: Martins Fontes.

\_\_\_\_\_. (1990) *Natureza Humana*, Rio de Janeiro: Imago.

ZAHAVI, Dan, (2005) *Subjectivity and selfhood. Investigating the first-person perspective*. Cambridge: MIT Press.