

# LINGUAGEM, COMPORTAMENTO E MENTE NO MITO DE JONES DE WILFRID SELLARS

## LANGUAGE, BEHAVIOR AND MIND IN WILFRID SELLARS'S MYTH OF JONES

*Marcelo Masson Maroldi*<sup>1</sup>

**Resumo:** Nos anos de 1950, Wilfrid Sellars procurou apresentar os erros da concepção clássica de “mente”, sugerindo em seu lugar uma abordagem centrada na análise da linguagem pública. Em sua proposta, a linguagem não é o veículo do pensamento, estados mentais não são absolutamente privados e o acesso imediato, por introspecção, é enganador. Essa formulação teve profundas implicações nas pesquisas sobre linguagem e mente da filosofia analítica. Este artigo analisa e discute a reconceitualização dos episódios mentais a partir do “mito de Jones”, ficção sugerida por Sellars para expor algumas de suas ideias sobre pensamento, linguagem e comportamento.

**Palavras-chaves:** Sellars. Mente. Linguagem. Comportamento. Ryle. Mito de Jones.

**Abstract:** In the 1950s, Wilfrid Sellars sought to present the errors of the classical conception of “mind”, suggesting instead an approach that is based on the analysis of public language. In this proposal, the language is not the vehicle of thought, mental states are not absolutely private, and immediate access by introspection is misleading. That account has had important implications for the language and mind research in analytic philosophy. This paper discusses Sellars's reconceptualization of mental episodes in the “myth of Jones”, a fiction created to expose the Sellars's ideas about the thinking, language and behavior.

**Keywords:** Sellars. Mind. Language. Behavior. Ryle. Myth of Jones.

\* \* \*

### Introdução

Ainda que varie em diversos aspectos, a visão geral sobre “mente” presente em Descartes, nos empiristas britânicos e em Kant inclui o princípio de que o sujeito tem acesso inquestionável e não mediado aos seus estados mentais, isto é, possui estados internos e privados dos quais é imediatamente consciente, por introspecção. Embora as dificuldades com este princípio fossem debatidas desde muito antes, a crítica atingiu seu ápice há pouco mais de um século, tornando-se mais elaborada e produzindo um clima

---

<sup>1</sup> Doutorando em Filosofia pela Universidade de São Paulo (USP). Bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). E-mail: marcelomaroldi@yahoo.com.br.

antimentalista diversificado que persistiu pelo menos até 1960. Inicialmente, ainda nas primeiras décadas do século XX, o método introspectivo caiu em descrédito, sendo substituído por abordagens que valorizavam mais a observação externa e a experimentação – foi a fase de maior prestígio das análises behavioristas. Mais tarde, nas décadas de 1940 e 1950, Gilbert Ryle (*The Concept of Mind*, em 1949), Wittgenstein (*Philosophical Investigations*, em 1953) e Sellars (*Empiricism and the Philosophy of Mind*, em 1956) apresentaram suas propostas para a análise dos episódios internos. Simultaneamente, mas por vias distintas, desconstruem o conceito clássico de “mente” promovendo uma análise baseada no funcionamento da linguagem pública. Como comenta Rorty (1979), estas três obras foram fundamentais para a mudança na caracterização da filosofia analítica, exercendo grande impacto junto à comunidade filosófica e contribuindo para um período de estímulo às abordagens anticartesianas.

Embora as *Philosophical Investigations* de Wittgenstein seja certamente a mais lida e comentada das três obras, o livro de Ryle obteve sucesso expressivo na época de sua publicação. Um de seus objetivos era explicar o discurso sobre o mental a partir da observação do comportamento público, recorrendo a um vocabulário que não contivesse os termos da linguagem psicológica. Já o livro de Sellars, apesar da importância que recentemente lhe tem sido atribuída, teve pouca influência direta naquele momento da discussão. Apesar de concordar com a inadequação e a crítica ao modelo cartesiano que *The Concept of Mind* apresentava, Sellars considerava que Ryle ignorava aspectos importantes do mentalismo. Segundo *Empiricism and the Philosophy of Mind*, ficamos com uma sensação desconfortável se tentamos explicar "pensamentos", "intenções", "desejos" etc. unicamente através de comportamentos e disposições comportamentais. O erro de Ryle seria o de ignorar a realidade interior, pois, segundo Sellars, episódios mentais existem, não são "erros categoriais"<sup>2</sup>. Ainda assim, a análise intersubjetiva do uso do vocabulário psicológico presente em *The Concept of Mind* indicava um caminho promissor e abria espaço para alternativas que não considerassem as experiências internas como experiências imediatas. Segundo Sellars, o acesso aos estados internos não precisa ser construído por meio de um modelo "perceptual ou quase perceptual", sendo possível, neste caso,

---

<sup>2</sup> "Unlike Ryle, I believe that meaningful statements are the expression of inner episodes, namely thoughts" (SELLARS e CHISHOLM, 1957).

desenvolver a ideia de que nós temos episódios internos que não são experiências imediatas, mas aos quais temos acesso privilegiado, embora não um acesso "invariável ou infalível".

Assim, Sellars se dedicou a compreender como podem existir episódios que combinem as noções de acesso privilegiado e intersubjetividade. Como estes conceitos podem ocorrer simultaneamente, sem contradição, ou sem sucumbir à tentação ryliana? Sua sugestão é uma forma modificada da tese de que os pensamentos são episódios linguísticos, mas sem sustentar que eles são experiências imediatas ou comportamento verbal. Ora, como pensamentos podem ser episódios internos se não são experiências imediatas? E como são linguísticos se não são performances linguísticas públicas? Estas dificuldades são discutidas no "mito de Jones", uma breve e difícil ficção de três estágios, apresentada na parte final de *Empiricism and the Philosophy of Mind*. O mito oferece uma explicação para o surgimento dos conceitos mentais em uma sociedade ancestral hipotética, do tipo sugerido pela proposta de Ryle. Sellars acreditava que um cenário artificial, mas logicamente possível, para o aparecimento do discurso mentalista poderia sugerir que adotamos um legado que não é a única maneira de tratarmos do vocabulário mental. Segundo Loux (*The Mind-Body Problem*, In: DELANEY et al., 1977, p. 107), a teoria dos estados mentais do tipo almejado por Sellars precisa:

- (1) Reconhecer a existência dos atos mentais.
- (2) Atribuir intencionalidade aos atos mentais.
- (3) Explicar o acesso privilegiado.
- (4) Explicar o conhecimento dos atos mentais dos outros (problema das outras mentes).
- (5) Não se comprometer com uma ontologia dualista.

Considerando estas ideias introdutórias, este artigo faz uma apresentação comentada do "mito de Jones" e, ao final, avalia se as cinco exigências mencionados acima são satisfatoriamente abordadas pela concepção que o mito oferece.

## **O mito de Jones: a comunidade em seu estágio inicial**

Sellars inicia sua "ficção científica antropológica" pedindo aos leitores que considerem uma comunidade caracterizada com o quadro behaviorista descrito em *The Concept of Mind*. Nesta obra, Ryle procura combater o que ele chama de "o mito do fantasma na máquina", a ideia de que nossas práticas racionais tenham como origem (fundamento) uma mente ou alma imaterial – o que, segundo ele, seria provocado pelo mau uso da linguagem. Em vista disso, Ryle procura mostrar que tudo o que podemos dizer de nossos estados internos poderia ser plenamente analisado em termos linguístico-comportamentais. Ao mostrar que tudo aquilo que estaria "dentro" de nós poderia ser reduzido a comportamentos e disposições, ficaria provado que não há qualquer fantasma na máquina. Como resultado, poderíamos refinar nossa linguagem para que ela não se referisse mais a algo interno e privado.

Segundo Ryle, atribuir a alguém um estado mental equivale a atribuir um comportamento ou uma disposição comportamental, isto é, o vocabulário de estados e eventos mentais pode ser eliminado por uma redução a um vocabulário mais básico. Portanto, é necessário mostrar que os termos mentais são equivalentes a construções lógicas sobre enunciados comportamentais (por isso "behaviorismo lógico"). A função da construção lógica, então, é traduzir os termos mentais em sentenças condicionais que expressem (semanticamente) comportamentos ou disposições comportamentais, em uma linguagem de fatos e eventos públicos. Por isso, devemos supor que a comunidade ryliana, no estágio inicial do mito, possui um vocabulário restrito a objetos públicos espaço-temporais. Todas as expressões racionais ou inteligíveis de seus membros são descritas a partir de seus comportamentos observáveis, em especial, de seus episódios linguísticos. Pensar que-P é analisável por dizer "P" (ou uma sentença semanticamente equivalente) ou ter uma propensão a dizer "P". Assim, afirmar que "João tem *medo* de baratas" seria equivalente à condicional, digamos, "Se aparecer uma barata nesta sala então João irá tremer"<sup>3</sup>, que não contém nenhum termo psicológico. O fundamental é notar que João não tem um conceito ou uma palavra que se refira a seu estado interno (como a nossa palavra

---

<sup>3</sup> E todas as demais possibilidades interligadas por um operador do tipo "ou" ("Se uma barata aparecer na sala então João vai sair correndo" ou "[...] então João vai chorar" ou "[...] então João vai desmaiar" etc.).

"medo"), mas traduziria sua condição na ocasião da aparição de uma barata com uma enunciação pública, declarando, por exemplo: "Minhas mãos tremem pois estou vendo uma barata".

Sellars considera este modelo de Ryle correto, de uma maneira geral, mas limitado e incapaz de explicar estados e eventos interiores que, na sua visão, realmente existem<sup>4</sup>. Um dos objetivos do mito, portanto, é mostrar um cenário no qual a linguagem dos rylianos poderia vir a ser enriquecida de modo a incluir as declarações relativas a estados internos, sem envolver os compromissos cartesianos com um fantasma na máquina. Ou seja, Sellars procura mostrar como a comunidade ryliana poderia passar de sua linguagem fisicalista e behaviorista a outra linguagem que inclua o vocabulário mental, mas sem os prejuízos do mentalismo tradicional. Ora, o que precisa ocorrer com a linguagem destes "animais falantes" para eles se reconhecerem como seres dotados de "pensamentos", "crenças", "sensações" etc.?

O primeiro passo é verificar se a linguagem ryliana pode vir a admitir um discurso *semântico*, "os recursos necessários para fazer enunciados semânticos característicos como '<Rot> significa vermelho' e '<Der Mond ist run> é verdadeiro se e somente se a lua for redonda'"(SELLARS, 1963, p. 179)<sup>5</sup>. Segundo Sellars, não há problema em supor que isso possa ocorrer. A comunidade ryliana poderia passar por um estágio do desenvolvimento de sua linguagem em que seus membros criam e aprendem a utilizar os termos semânticos. Isto é, eles podem aprender a dizer que sua linguagem *significa* isso ou aquilo, que sua linguagem diz que isso ou aquilo é *verdadeiro* etc. Paralelamente, eles podem começar a utilizar operadores lógicos (negação, quantificação, etc.) misturados a esse novo vocabulário, aumentando sua capacidade expressiva e tornando-se aptos a certos movimentos sobre a linguagem – por exemplo, fazer inferências com enunciados. Ou seja, os rylianos fazem evoluir sua linguagem de modo a formularem frases como 'Solteiro significa homem não casado' e 'O céu é azul é verdadeiro' mas, também, 'Se há nuvens negras no céu então cairá água' ou 'Se eu tocar André com um martelo então ele dirá <ai>'.  

---

<sup>4</sup> É importante notar que a linguagem dos rylianos é suficiente (completa) para o modo de vida de seus membros, ela apenas não inclui certos usos que Sellars considera importantes para que uma linguagem cumpra seu papel cognitivo máximo. Mas os behavioristas rylianos não são menos racionais apenas por não terem em sua linguagem um vocabulário mental e psicológico.

<sup>5</sup> "[...] the resources necessary for making such characteristically semantical statements as '<Rot> means red', and '<Der Mond ist run> is true if and only if the moon is round'" (SELLARS, 1963, p. 179).

E esse progresso da linguagem ryliana não requer nenhuma alteração no paradigma vigente da comunidade (que se mantém, portanto, behaviorista e fisicalista).

Até aqui, então, os rylianos possuem um vocabulário que contém referências a fatos públicos, objetos materiais, comportamentos públicos e termos semânticos usados em comportamentos linguísticos. Eles ainda não têm palavras e conceitos relacionados a pensamentos, sensações, sentimentos, crenças, introspecção ou sonhos. O estágio seguinte da ficção é o enriquecimento desta linguagem com a adição de um discurso *teórico*, o que ocorre quando os membros da comunidade começam a elaborar hipóteses para explicar as alterações em fatos observáveis, por exemplo, a mudança no estado físico da água que fica muito tempo sobre o fogo. Ora, isso é plausível porque o fato de que a linguagem seja pública não significa que ela esteja permanentemente restrita a objetos públicos. Sua publicidade não precisa excluir ou negar as noções e os termos teóricos. Assim, a comunidade pode desenvolver um vocabulário e um modo de uso deste vocabulário que distingue o que é observável daquilo que é teórico, mas pode ser utilizado para explicar as situações observáveis (públicas). Desse modo, a comunidade passa a se acostumar com frases como 'A água esquentada porque as moléculas se agitam' ou 'A aceleração da gravidade na Terra é  $9,8\text{m/s}^2$ ', embora nenhum de seus membros tenha observado moléculas ou a aceleração da gravidade.

Portanto, se à linguagem ryliana original já havia sido adicionado o discurso semântico, podemos assumir que também lhe foi inserida uma camada teórica. Estamos prontos para o passo decisivo, que ocorre quando surge Jones.

### **O aparecimento de Jones**

Nesse momento da narrativa, a comunidade neoryliana possui uma linguagem usada para descrever objetos públicos, enriquecida com termos semânticos e teóricos, mas ainda restrita ao vocabulário não teórico de uma psicologia essencialmente behaviorista. Tal linguagem não inclui palavras para descrever estados psicológicos. Então, Sellars pede para imaginarmos que na comunidade apareça um estrangeiro chamado Jones, um pensador adepto de um tipo de behaviorismo que permite a inclusão de termos teóricos relacionados ao vocabulário observacional, mas sem implicar qualquer ontologia (a distinção entre

termos teóricos e objetos observacionais é epistemológica). No início, Jones não tem dificuldade em aceitar que as pessoas da comunidade estejam agindo de forma racional quando falam, mas ele acha curioso que elas estejam se comportando deste mesmo modo quando agem em silêncio<sup>6</sup>. Ou seja, se os casos de comportamento inteligente que acompanhavam as manifestações verbais pareciam triviais, Jones se inquietou com aqueles casos em que isso não ocorria, quando as pessoas pareciam somente falar consigo mesmas. Essas atitudes exigiam uma investigação suplementar.

Jones observou que, em alguns casos, ocorria um processo de aprendizado que produzia sempre os mesmos comportamentos, enquanto que, em outros momentos, parecia ocorrer um ato de deliberação interna, sem qualquer evidência observável, modificando os comportamentos esperados. Esse processo "interior" parecia ser a causa dos indivíduos virem a se comportar distintamente. Por exemplo, se a propensão a responder ao surgimento de um pássaro no céu deveria equivaler a falar "Há um pássaro no céu", Jones nota que essa propensão poderia ser substituída por outra enunciação qualquer, por um comportamento (como fuga ou choro) ou simplesmente seguia-se um período de silêncio. É nesse instante que Jones começa a revolucionar a vida dos membros da comunidade ao propor uma teoria para explicar o que ocorre em tais casos, uma teoria para explicar o comportamento das pessoas. Ele desenvolve a tese de que os episódios de comportamento público, incluindo o comportamento verbal, são apenas o ponto culminante de um processo iniciado por episódios internos do sujeito. Assim, Jones sugere considerar a atividade mental como uma espécie de enunciação interna, uma linguagem do pensamento cujo desfecho é a fala pública, sua expressão natural. A partir disso, ele passa a postular entidades teóricas ("pensamento", "crença", "dor", "medo" etc.) e desenvolve sua teoria como uma explicação causal. Segundo a nova explicação, o comportamento de João de gritar e tremer ao ver uma barata é o resultado de um estado interno de João, o "medo". Ou seja, as entidades criadas passam a ser utilizadas na explicação dos comportamentos, permitindo afirmar que "João está com *medo* da barata" se João gritar ou tremer ao avistar uma barata. Portanto, o comportamento manifesto já não é mais tudo o que resta para ser

---

<sup>6</sup> Como observam deVries e Triplett (2000), a capacidade de agir silenciosamente deveria ser muito pouco usual nessa comunidade imaginada. Algumas atividades mentais, sem a contraparte pública, nem sempre foram facilmente entendidas. Nas suas confissões (Agostinho. *Confissões*. Trad. de Maria Luiza Jardim Amarante. São Paulo: Paulus, 1984), por exemplo, Agostinho narra que ler em silêncio era bastante incomum em sua época.

utilizado na explicação da conduta dos membros da comunidade, pois a teoria adiciona um nível de explicação anterior a ele. É importante frisar, porém, que as entidades criadas não tiveram motivação empírica, mas são puramente teóricas, introduzidas metodologicamente pela teoria de Jones.

Após elaborar essa explicação, Jones começa a ensinar aos membros da comunidade como aplicar a teoria, de modo que todos possam justificar as ações uns dos outros segundo o novo vocabulário e as implicações teóricas que seu uso estabelece. Assim, a teoria vai naturalmente se incorporando à vida das pessoas, de modo que sua utilização gera menos estranheza a cada dia, a ponto de se tornar plenamente aceita e passar a integrar o quadro conceitual da comunidade. Nesse sentido, podemos dizer que surge uma psicologia cotidiana e científica.

Mas como Jones chegou a elaborar sua teoria? Em primeiro lugar, ele percebeu que o significado de um termo da linguagem pública servia para classificar as expressões funcionalmente, em termos de seus papéis na prática linguística. Então, ele transfere essa mesma ideia para as novas entidades que concebeu, dando-lhes um tratamento igualmente funcional<sup>7</sup>. Isto é, um estado do tipo que a teoria trata, como o "medo", é uma noção classificatória, um conceito para descrever um estado do sujeito que desempenha um papel em uma situação específica. O significado de um estado mental depende da função que ele cumprir na vida do agente. A camada semântica da linguagem (que já havia sido introduzida na linguagem da comunidade previamente à chegada de Jones) é crucial por permitir transpor seus termos para o novo vocabulário, teórico (que a comunidade também já estava habituada a usar antes de Jones aparecer), para as novas entidades "mentais" recém-criadas<sup>8</sup>. Ou seja, do mesmo modo que as pessoas falavam que uma declaração qualquer "R" significava algo, que era sobre algo, que se referia a algo ou que era verdadeira, passam a falar que um pensamento "S" significa algo, que é sobre algo, que se refere a algo ou que é verdadeiro. Portanto, o que Sellars está sugerindo, através do mito, é uma teoria psicológica funcional que mantém o que ele entende ser a característica

---

<sup>7</sup> Ou seja, Jones é adepto de uma teoria funcionalista, tanto do conteúdo linguístico como do conteúdo mental. De fato, a concepção de pensamentos como papéis funcionais desenvolvida por Sellars foi precursora do movimento filosófico contemporâneo conhecido na filosofia da mente como funcionalismo, embora praticamente nenhum crédito lhe tenha sido atribuído por isso. Cf. MAROLDI, M. M. *O funcionalismo de Sellars: uma pesquisa histórica*. Ciências & Cognição 14 (3), 24-38, 2009.

<sup>8</sup> A intencionalidade primária pertence à linguagem, e apenas de modo derivado aos estados mentais.



fundamental de um estado mental, a sua intencionalidade (referir-se ou dirigir-se a algo), que é construída sobre o modelo semântico da linguagem pública. Os novos rylianos fazem referência à intencionalidade de seus estados mentais usando as categorias semânticas de sua própria linguagem. Portanto, fica claro por que Sellars inicia mostrando que os rylianos podem desenvolver um discurso semântico, como uma etapa intermediária obrigatória para o desenvolvimento do vocabulário psicológico. Se os rylianos podem vir a ter um vocabulário semântico, está aberta a possibilidade de introduzir o vocabulário mental, como intencional e teórico.

A ideia de que o vocabulário semântico é mais básico (anterior) que o vocabulário psicológico permite conceber a intencionalidade dos episódios mentais considerando que o processo de atribuição de significação aos estados mentais depende da linguagem pública. Todavia, apesar da independência e aparente primazia cronológica da linguagem, a significação dos episódios internos e dos termos da linguagem pública vão se desenvolvendo conjuntamente, isto é, a habilidade de dar sentido a pensamentos, sentimentos e outros estados internos vai sendo adquirida simultaneamente à aquisição da fala pública. Assim, embora o conhecimento dos episódios linguísticos seja primeiro na "ordem do conhecimento", ele não tem prioridade ontológica sobre o mental. Não é a linguagem que cria estados mentais, a teoria de Jones é uma explicação epistemológica. (Não é que os rylianos não tinham estados mentais antes de a teoria ter sido criada, mas que sua intencionalidade, seu papel na vida das pessoas, foi posterior à teoria de Jones e sua análise da atividade linguística).

Vale notar, finalmente, que Sellars não está afirmando que em algum momento da história humana um indivíduo inventou tudo isso, reuniu as pessoas e explicou sua teoria, que todos passaram a usar. Ele está sugerindo que o vocabulário mentalista *poderia* vir a ser introduzido na linguagem de uma comunidade desta maneira, que os seres humanos poderiam ter adotado um esquema conceitual progressivo para descrever suas experiências internas e que, incorporado à linguagem ordinária, teria seu uso tão assimilado a ponto de ser totalmente justificado<sup>9</sup>. Também é preciso esclarecer que Jones não construiu uma

---

<sup>9</sup> Assim, o mito de Jones não pretende servir como uma explicação histórica. Contudo: "But is my myth really a myth? Or does the reader not recognize Jones as Man himself in the middle of his journey from the grunts and groans of the cave to the subtle and polydimensional discourse of the drawing room, the laboratory, and the study, the language of Henry and William James, of Einstein and of the philosophers who, in their efforts

teoria completa, imune a erros e que não exige revisão. Explicar estados internos como ele fez é estender a linguagem para abranger um novo nível de explicação para o comportamento dos indivíduos, uma explicação que pode ser alterada, se necessário. (Isso seria perfeitamente possível, por exemplo, se os avanços da ciência fizesse Jones reformular os usos de certos termos que ele criou).

### **Acesso privilegiado**

Como vimos, o fato de que os eventos mentais tenham sido introduzidos como "internos" não significou que eles tenham sido introduzidos como o resultado de experiências imediatas dos sujeitos, pelo contrário. Os membros da comunidade não atribuem significado a seus estados internos simplesmente por experienciá-los, mas após dominarem uma técnica de inferi-los segundo os papéis que estes desempenham em suas vidas, seguindo a explicação de Jones. Por isso, o mito não se encerra após a formulação e a disseminação da teoria, tal como encontra neste momento da narração. Sellars mostrou como podemos entender os conceitos mentais a partir da dimensão intersubjetiva. Nada foi dito, no entanto, a respeito do acesso privilegiado, que precisa explicar o conhecimento direto dos estados mentais, aquilo que não depende de inferências sobre evidências comportamentais. Em outras palavras: como compreender a perspectiva mental em primeira pessoa? Isto é, Jones desenvolveu uma teoria que explica o comportamento dos membros da comunidade como a expressão de episódios internos e ensinou a comunidade a avaliar e descrever os comportamentos alheios usando a teoria. Mas como passar do conhecimento baseado na observação externa para o autoconhecimento? Segundo Sellars, é preciso um "pequeno passo" para que a linguagem dos neorylianos comece a ser utilizada na autodescrição de eventos internos.

---

to break out of discourse to an *arché* beyond discourse, have provided the most curious dimension of all" (SELLARS, 1963, p. 195-6). Rebecca Kukla (Myth, Memory and Misrecognition in Sellars's 'Empiricism and the Philosophy of Mind'. *Philosophical Studies* 101 (2-3):161-211, 2000) faz importantes observações sobre o papel do mitológico na obra de Sellars. Ela escreve: "[...] the philosophical function of these myths is irreducibly distinct from the function of literal or pseudoliteral descriptions of how things are [...] Sellars' mythical explanations in EPM [Empiricism and the Philosophy of Mind] employ a very specific and rhetorically complex methodology, and likewise that we will not be in a position to critically assess the paper's arguments unless we give careful attention to its overall textual structure and to the nature of the mythical explanations it employs" (KUKLA, 2000, p. 162-3).

Até este instante, devemos pensar que Paulo, após aprender a usar a teoria de Jones, poderia afirmar, por análise comportamental, que "João está com medo da barata". Do mesmo modo, o próprio João, usando a mesma evidência comportamental e aplicando a mesma teoria, seria capaz de declarar: "Eu estou com medo da barata". Entretanto, num momento posterior à assimilação da teoria, João pode vir a ser treinado a dar relatos confiáveis dos seus próprios estados mentais sem ter de observar o seu comportamento manifesto. Ou seja, se no princípio ele aprende a inferir a existência de uma entidade a partir de uma evidência, Jones o ensina a considerar seus estados mentais sem a necessidade de tal evidência. João se torna capaz de identificar seu "medo", de atribuir a este seu estado não apenas uma palavra, mas todo um contexto intencional em que ele se insere, sem ter de observar o próprio comportamento (desenvolvendo a habilidade de identificar tal estado não mais como a contraparte comportamental). E isso é possível porque Jones treina João a reconhecer seus estados psicológicos reforçando, positiva e negativamente, as enunciações que João faz, diante das evidências externas disponíveis a Jones. Uma vez treinado, a evidência pode ser desprezada, e aprende-se a identificar o próprio estado.

Então, quando Tom, observando Dick, tem a evidência comportamental que sustenta o uso da frase (na linguagem da teoria) 'Dick está pensando <p>' (ou 'Dick está pensando que p'), Dick, usando a mesma evidência comportamental, pode dizer, na linguagem da teoria, 'Eu estou pensando <p>' (ou 'Eu estou pensando que p'). E agora se revela – precisaria? – que Dick pode ser treinado a dar autodescrições razoavelmente confiáveis, usando a linguagem da teoria, **sem ter de observar seu comportamento público**. Jonas consegue isso, de forma aproximada, aplaudindo enunciações de Dick de 'Eu estou pensando que p' quando a evidência comportamental apoia fortemente o enunciado teórico 'Dick está pensando que p'; e desaprovando enunciações de 'Eu estou pensando que p', **quando a evidência não apoia essa declaração teórica**. Nossos ancestrais começam a falar do acesso privilegiado que cada um de nós tem a seus próprios pensamentos. *O que começou como uma linguagem com um uso puramente teórico ganhou um papel de relato* (SELLARS, 1963, p. 189, grifo nosso).<sup>10</sup>

---

<sup>10</sup> "Thus, when Tom, watching Dick, has behavioral evidence which warrants the use of the sentence (in the language of the theory) 'Dick is thinking <p>' (or 'Dick is thinking that p'), Dick, using the same behavioral evidence, can say, in the language of the theory, 'I am thinking <p>' (or 'I am thinking that p'). And it now turns out – need it have? – that Dick can be trained to give reasonably reliable self-descriptions, using the language of the theory, without having to observe his overt behavior. Jones brings this about, roughly by applauding utterances by Dick of 'I am thinking that p' when the behavioral evidence strongly supports the theoretical statement 'Dick is thinking that p'; and by frowning on utterances of 'I am thinking that p', when

Portanto, a habilidade de relatar as experiências internas na ausência de evidências públicas e sem depender de inferências permite um sentido em que podemos dizer que o conhecimento das experiências internas é "direto". A linguagem passa a ser usada para descrever as próprias disposições comportamentais:

Ela [Sally] é capaz de fazê-lo enquanto permanece em silêncio na ocasião, pois ela também foi treinada para 'segurar' seus pensamentos-em-voz-alta em circunstâncias apropriadas. Com o tempo [...] a ryliana Sally estará apta, se chamada, a dar relatos-em-voz-alta altamente elaborados e unicamente confiáveis (embora, de maneira nenhuma, infalíveis) de sua história recente de pensar-em-voz-alta e suas propensões para tal. Ninguém pode ganhar dela, ela consegue de olhos fechados. Ela confessa seus próprios pensamentos-em-voz-alta diretamente ou não inferencialmente, isto é, sem qualquer inferência nestes casos de evidências comportamentais 'externas'. Agora ela tem *acesso privilegiado* [...] aos seus próprios pensamentos-em-voz-alta.<sup>11</sup> (O'SHEA, 2007, p. 99).<sup>12</sup>

E a introspecção, entendida como a capacidade de identificar o próprio estado interno através de uma "olhada para dentro" é, segundo Sellars, uma capacidade aprendida: o acesso individual aos estados mentais surge após treinamento. Dessa maneira, a evidência comportamental torna-se dispensável na autoaplicação da teoria. É assim que nasce a noção de "acesso privilegiado", ainda compreendido como o que apenas o indivíduo pode identificar em si, mas que não é tratado como uma capacidade inata ou imediata – o acesso privilegiado não é "dado" e só ocorre após o domínio da linguagem pública. Nota-se, assim, a intenção de Sellars em mostrar como pode ser possível compreender a noção de "episódio

---

the evidence does not support this theoretical statement. Our ancestors begin to speak of the privileged access each of us has to his own thoughts. *What began as a language with a purely theoretical use has gained a reporting role.*" (SELLARS, 1963, p. 189).

<sup>11</sup> Sobre isso, Brandom (IN: SELLARS, 1997, p. 176) afirma: "one can develop a conditioned reflex in someone [...] to report noninferentially what heretofore could only be inferred". Para uma discussão completa veja O'SHEA (2007, pp. 97-105) e também de Vries e Triplett (2000). O'SHEA menciona o quão objetável pode ser tal noção de acesso privilegiado e de Vries e Triplett a julgam difícil de ser compreendida, criticável e altamente misteriosa.

<sup>12</sup> "She [Sally] is able to do so while remaining silent on this occasion, for she has also been trained to 'hold back' her thinkings-out-load in appropriate circumstances. By the time [...] Rylean Sally will be able, if called upon, to give highly elaborate and uniquely reliable (but by no means infallible) reports-out-loud on her own recent personal history of thinking-out-loud and her propensities to such. No one can beat her at it; she can do it with her eyes shut. She avows her own thoughts-out-load directly or non-inferentially, that is without any inference in these cases from 'outer' behavioral evidence. She now has a *privileged access* [...] to her own thoughts-out-load" (O'SHEA, 2007, p. 99)

interno" – aquilo que está sob a pele, como diria um behaviorista – como um conceito intersubjetivo e, principalmente, que o fato de alguém relatar seus episódios internos na ausência de evidências públicas pressupõe e é totalmente conforme a essa intersubjetividade. Em outras palavras, o fato da linguagem psicológica ser aprendida em contextos públicos pode ser conciliado com a exigência de privacidade, ainda que não se trate de uma privacidade "absoluta". Se o mito de Jones reconhece que episódios internos podem ser relatados na ausência de comportamentos públicos, também sugere que a evidência comportamental deve estar embutida na própria lógica do conceito, "assim como o fato de que o comportamento observável dos gases é evidência para episódios moleculares está incorporado na própria lógica da fala sobre moléculas" (SELLARS, 1963, p. 195)<sup>13</sup>. Em outras palavras, o que Sellars mostra é como os rylianos poderiam adquirir conceitos relativos a estados mentais como internos e privados, combinados a noção de acesso privilegiado, em uma reformulação que exclui as desvantagens do mentalismo clássico e do mito do dado (rejeitada na primeira parte de *Empirismo e Filosofia da Mente*).<sup>14</sup>

Mas isso não é tudo. Antes de desaparecer da comunidade, Jones acaba se equivocando, confundindo-se em relação à teoria que ele mesmo criou. Acostumado a seu uso, ele se esquece de que as entidades representadas pelo vocabulário psicológico foram criadas e introduzidas como entidades *teóricas* e, assim, "confunde seu próprio enriquecimento criativo da estrutura do conhecimento empírico com a análise do conhecimento tal como foi" (SELLARS, 1963, p. 195)<sup>15</sup>. Isto é, Jones esquece que o uso das entidades teóricas aplicadas ao discurso mental envolveu uma dimensão puramente epistemológica, e não ontológica. Como consequência, ele começa a crer que as entidades postuladas realmente existem e, desse modo, elas passam a ser usadas como fundamento do saber – como parte do esforço e da tendência humana de querer ir além do discurso e procurar um fundamento (por vezes metafísico) para as suas explicações. Em outras

---

<sup>13</sup> "[...] just as the fact that the observable behavior of gases is evidence for molecular episodes is built into the very logic of molecule talk" (SELLARS, 1963, p. 195)

<sup>14</sup> Jones, após de ter desenvolvido uma teoria para uma classe específica de episódios mentais ("pensamentos", num sentido amplo), a estende para abarcar outra classe de episódios internos, as "impressões" – o que ele faz analogamente aos "pensamentos". Por acreditar que o tópico das "impressões" é extenso e bastante problemático, exigindo um trabalho totalmente dedicado a ele, o tema não será tratado aqui.

<sup>15</sup> "[...] confuses his own creative enrichment of the framework of empirical knowledge, with an analysis of knowledge as it was" (SELLARS, 1963, p. 195)

palavras, acostumado à sua teoria, Jones passa a considerar as entidades teóricas como entidades existentes. É o ponto de partida para o mentalismo clássico.

### **Considerações finais**

Será que o "mito de Jones" atende às características que Sellars desejava contemplar em sua filosofia da mente, apresentadas sucintamente no "mito de Jones"? Como lembra Loux, são elas:

- (1) Reconhecer a existência dos atos mentais.
- (2) Atribuir intencionalidade aos atos mentais.
- (3) Explicar o acesso privilegiado.
- (4) Explicar o conhecimento dos atos mentais dos outros (problema das outras mentes).
- (5) Não se comprometer com uma ontologia dualista.

O ponto central da argumentação do "mito de Jones" visa explicar que o vocabulário mental pode ser introduzido como meramente teórico, mostrando sua relação com o comportamento público. Jones parece ser bem sucedido ao postular a existência de entidades inobserváveis para explicar o comportamento observável. Segundo Sellars, ter boas razões para adotar uma teoria é ter boas razões para crer que as entidades que ela postula realmente existam<sup>16</sup>, o que satisfaz (1).

A estrutura teórica psicológica que Jones criou envolveu um modelo mais familiar, a fala pública. Então, Sellars pôde tomar os "pensamentos" como um tipo de discurso interno e sua intencionalidade como uma extensão das propriedades semânticas do discurso linguístico. Assim, estados mentais são sobre (ou de) algo como uma apropriação e extensão do uso do vocabulário semântico ordinário. Isso atende (2).

Sellars acredita que a teoria de Jones não é válida (aplicável) apenas como uma conclusão de inferências, mas que também é possível empregá-la em relatos autodescritíveis, por introspecção. Por funcionar em um contexto não inferencial, Sellars

---

<sup>16</sup> Cf. *The Language of Theories* (In: SELLARS, 1963).

acredita que sua teoria atende (3). Nesse caso, enquanto uma pessoa pode relatar seus estados internos não inferencialmente, os demais só podem relatar os estados internos dessa pessoa a partir de evidências públicas.

Os seres humanos de fato atribuem estados mentais para si e para outras pessoas e usam um vocabulário psicológico para descrever, entender e predizer seu comportamento e o dos outros. Estados mentais, entendidos como entidades teóricas, podem ser satisfatoriamente usados para explicar o comportamento e os estados individuais das pessoas, sobretudo de terceiros. Ou seja, a teoria de Jones pode ser utilizada para explicar a nossa capacidade de termos uma "teoria da mente" (ou "psicologia popular"). Isso satisfaz (4).<sup>17</sup>

Por fim, Sellars quer evitar uma ontologia dualista do tipo cartesiana. Ora, o significado dos atos mentais é uma questão de atribuição de papéis. Portanto, na teoria de Jones, os estados mentais não são descritos ontologicamente, mas funcionalmente. Um ato mental intencional é uma entidade teórica, e nada é dito sobre sua natureza – deve-se aceitar a existência de estados mentais (compromisso ontológico), mas caracterizá-los a partir de suas funções e não por sua ontologia. Desse modo, não se deve concluir que sejam entidades físicas ou imateriais, por exemplo. O mito de Jones, ao desvincular do dualismo as noções de intencionalidade e de acesso privilegiado, mostra como acomodar estes dois conceitos em uma teoria que não se compromete com qualquer ontologia particular. Portanto, a teoria de Sellars é sobre a natureza dos conceitos mentais e ontologicamente neutra. Isso contempla (5)<sup>18</sup>.

No desfecho da saga de Jones, a linguagem dos rylianos foi enriquecida com termos semânticos e teóricos e, por último, com a adoção do próprio vocabulário psicológico. É importante notar que ela realmente é uma nova linguagem, ampliadora da capacidade de expressão anterior, e que permite explicar e sustentar afirmações com base em características "internas" e "privadas". As pessoas aprendem o uso dessa teoria e verificam sua superioridade epistemológica em relação ao velho modelo ryliano, passando a adotá-la. Mas por que a explicação de Jones é mais adequada que a antiga explicação? Em primeiro

---

<sup>17</sup> O mito de Jones sugeriu uma versão preliminar do que atualmente é conhecido como "teoria da teoria" (uma hipótese para a nossa "teoria da mente").

<sup>18</sup> Assim, o mito de Jones permite inúmeros desenvolvimentos, tanto materialistas quanto metafísicos. Ela não precisa ser tomada com uma forma cartesiana e não cria problemas ao realista científico, por exemplo, podendo, inclusive, ser combinada à teoria da identidade ou ao eliminativismo.

lugar, ela não precisa ser formulada como um behaviorismo lógico. Sellars acredita que o behaviorismo lógico é correto em defender a dimensão intersubjetiva do significado, mas erra ao dar uma visão limitada da conduta racional e ignorar a complexidade da realidade interior e sua importância para a compreensão do comportamento manifesto. Além disso, Sellars acredita que o behaviorismo lógico acaba caindo em uma circularidade lógica injustificável. Um segundo ponto (também apontado por O'SHEA, 2007) é que a teoria de Jones oferece uma explicação melhor de certas evidências comportamentais básicas, por exemplo, que podemos pensar mais rápido do que podemos falar. Essa velocidade de mudar de pensamento nem sempre acompanha a mudança da expressão verbal, de modo que as disposições nem sempre se refletem em verbalizações. Isso sugere que há episódios internos que não são propensões à fala pública. A terceira observação diz respeito ao fato de que, uma vez que a teoria tenha sido concebida segundo um modelo de entidades inobserváveis, ela pode explicar com sucesso comportamentos e disposições a partir de regularidades observadas. Ou seja, se a teoria de Jones é uma boa teoria, as disposições e os comportamentos estarão contemplados pela teoria, através de uma relação causal. As entidades teorizadas por Jones cumprem com precisão um papel explicativo, mesmo que sejam identificadas somente com o auxílio da linguagem pública. Assim, trata-se de um modelo teórico adequado à predição e à explicação do comportamento. Outra vantagem é que a explicação de Jones permite uma abordagem convincente para o problema de outras mentes. Vimos que a teoria de Jones explica a capacidade humana de atribuir estados mentais a si e aos demais membros da comunidade, ou seja, uma psicologia popular. Sellars sugere que o conhecimento dos termos mentais é aprendido a partir da interação das pessoas com a sua comunidade, revertendo a noção de que nosso conhecimento surge de "dentro" para "fora". É possível iniciar com o conhecimento dos objetos públicos e evoluir para um vocabulário psicológico eficiente. Finalmente, o grande mérito da teoria de Sellars é iluminar alguns conceitos da filosofia da mente através de uma hipótese plausível. Se o mito de Jones é uma hipótese possível, não é preciso aceitar cartesianismo ou inatismo na explicação dos conceitos referentes à realidade mental, mas uma melhor explicação pode vir a ser desenvolvida de maneira adequada à prática científica. O resultado final do "mito de Jones" é bem capturado pela declaração de Rorty (In: SELLARS, 1997) de que Sellars teria mostrado como ser wittgensteiniano sem ser ryliano – isto é, como podemos aceitar as



objeções ao mental feitas por Wittgenstein (principalmente o argumento da linguagem privada) sem ter que desconsiderar a existência dos estados mentais, como Ryle.

## **Referências**

- DELANEY, C. F. et al. *The Synoptic Vision. Essays on the philosophy of Wilfrid Sellars*. Indiana: University of Notre Dame Press, 1977.
- DEVRIES, W. A.; TRIPLETT, T. *Knowledge, Mind, and the Given: reading Wilfrid Sellars's "Empiricism and the Philosophy of Mind" including the complete text of Sellars's essay*. Indianapolis/Cambridge: Hackett Publishing, 2000.
- O'SHEA, J. R. *Wilfrid Sellars: naturalism with a normative turn*. Malden: Polity, 2007.
- RYLE, G. *The Concept of Mind*. London: Hutchinson, 1975.
- RORTY, R. *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press, 1979
- SELLARS, W. *Science, Perception and reality*. London: Routledge & Kegan Paul, 1963.
- \_\_\_\_\_. *Empiricism and the Philosophy of Mind*. Cambridge, Mass.: Harvard University Press, 1997.
- \_\_\_\_\_.; CHISHOLM, R. M. Intentionality and the Mental: Chisholm-Sellars Correspondence on Intentionality. In: FEIGL, H.; SCRIVEN, M.; MAXWELL, G. (Eds.). *Minnesota Studies in the Philosophy of Science*, 2, p. 521-539, 1957. Disponível em <http://www.ditext.com/sellars/sccor-f.html>.