

REPRESENTAÇÕES PROTOTÍPICAS E O EXPERIMENTO DE PENSAMENTO DAS TERRAS GÊMEAS

PROTOTYPICAL REPRESENTATIONS AND THE TWIN EARTH THOUGHT EXPERIMENT

Samuel de Castro Bellini-Leite¹

Resumo: O experimento de pensamento das Terras Gêmeas é usado como um dos fortes argumentos a favor da tese de que significados não estão na cabeça, cérebro ou mente. Neste artigo este argumento será considerado, e será criticado a partir de uma análise de exemplos nos quais a referência de um termo no mundo não possui essência fixa e as definições são imprecisas. Em contraste, será argumentado que o significado está ligado a representações internas prototípicas individuais.

Palavras-chave: Terras Gêmeas, Representações prototípicas e Significado.

Abstract: The Twin Earth thought experiment is used as a strong argument for the thesis that meanings are not in the head, brain or mind. I will consider the argument, and criticize it based on analyses of examples on which the reference of a term in the world does not have a fixed essence and its definitions are imprecise. In contrast, I will argue that meanings are related to individual internal prototypical representations.

Keywords: Twin Earth, Prototypical Representations and Meaning.

* * *

Introdução

O internalismo e o externalismo são duas doutrinas frequentemente discutidas na filosofia, cujas teses se estendem sobre campos diversos da epistemologia. Grayling (2007) define teses internalistas como aquelas que assumem que uma crença não pode ser justificada para um sujeito S a não ser que este tenha acesso àquilo que garante a justificação da mesma; já o externalismo entende que S não precisa ter acesso cognitivo à justificação de sua crença em P. Este trabalho foca a discussão sobre o internalismo ou externalismo do significado. Logo, o ponto principal da discussão será se o significado de um termo é dado pelo sujeito cognitivo ou se é determinado pelo objeto do mundo ao qual este termo se refere. Por exemplo, para os internalistas o significado da proposição “o céu está nublado” é criado por capacidades internas, como por exemplo, uma representação mental ou a fenomenologia subjetiva, enquanto que para os externalistas o céu em si seria responsável pelo significado desta proposição. Mas mais especificamente, o objetivo do texto é apenas de tecer uma crítica a um argumento a

favor do externalismo do significado, o experimento das Terras Gêmeas elaborado por Putnam (1973). Para chegar a esta meta, teorias da ciência cognitiva sobre a formação de conceitos e categorias são descritas na primeira seção e utilizadas na última, juntas com algumas considerações filosóficas sobre essências.

Conceitos e Categorias

Conceitos, para a ciência cognitiva (Murphy, 2002), são usados para dividir o mundo em grupos que podem ser tratados como equivalentes por algum motivo. Estes podem se referir a objetos, relações, eventos ou ideias, sendo usados para raciocínio, comunicação e inferências sobre o futuro. Além disso, fornecem informações críticas para nossa interação com objetos e nossa participação em eventos.

Alguns pesquisadores distinguem entre conceitos de informação e conceitos de categorias (MARKMAN, 2005). O primeiro se refere a uma informação apenas, e de acordo com Sternberg (2008), o segundo organiza ou aponta aspectos de equivalência entre outros conceitos. Markman (2005) afirma que o estudo de conceitos abrange representações inseridas em uma categoria, referindo a informações que descrevem uma categoria específica como “goiaba”, e entre categorias que se referem a relações entre diversos tipos de categorias como, por exemplo, “goiaba”, “fruta”, “verde” “vermelha”.

De acordo com Murphy e Lassaline (1997), parte do raciocínio humano advém da habilidade de pensar sobre a mesma coisa por ângulos diversos, assim nos permitindo o acesso a diferentes tipos de conhecimento sobre o objeto (pensar um cachorro como mamífero nos leva a associações biológicas sobre o mesmo, mas pensá-lo como amigo nos leva a outras concepções). Essa flexibilidade de conceitos gera complexidades, pois cada categoria diferente deve ser distinguida e armazenada na memória, sendo que a mais apropriada deve ser escolhida no momento correto.

É possível separar primeiramente categorias em quatro tipos: naturais, artificiais, *ad hoc* e nominais. Naturais acontecem naturalmente no mundo, como árvores e mamíferos; artificiais são agrupamentos consensuais dos seres humanos, como cozinha ou machado; *ad hoc* diz respeito a categorias formadas tendo um propósito específico em mente, como por exemplo: “coisas para serem estudadas”. O tipo nominal é uma atribuição arbitrária de um nome a algo que cumpre um conjunto de condições. Por exemplo, viúva ou amante. Como é difícil saber de que forma o significado é obtido por

meio de categorias, principalmente nominais, diversos modelos são propostos (STERNBERG, 2008).

Uma das formas mais básicas de entender como categorias são formadas é a partir de características definitórias. Características necessárias e suficientes para algo pertencer à categoria. Para ser um “x” tal coisa necessariamente tem que ter “y”. Dessa forma, por exemplo, para ser um “marido” é preciso ser um homem e ser casado. Não tendo uma dessas duas características definitórias, o objeto em questão não será um marido, e tendo estas características *necessariamente* será um marido. Dessa forma, a categoria marido é formada por duas características definitórias. No primeiro momento, tal teoria parece abarcar coerentemente as categorias, mas se analisarmos com exemplos, esta teoria não parece apresentar descrições completas das categorias (STERNBERG, 2008). Um exemplo muito problemático para essa teoria é a definição de jogo. Parece ser difícil, se não impossível, encontrar características necessárias e suficientes para definir um jogo. Algumas das características são envolver prazer, ter um desafio a ser cumprido, ter interação entre participantes. Essas são características comuns de um jogo, mas não necessárias ou suficientes. Um jogo não precisa envolver prazer, você pode jogar um jogo entediante. Você pode praticar uma atividade desafiante que não seja um jogo. Também, pode jogar sem interação com alguém, ou até mesmo sem interação com uma máquina (como um jogo de baralho sozinho). A teoria baseada em características definitórias se mostra atrativa, mas de fato não abarca todas as possibilidades das categorias.

Uma teoria mais completa e mais resistente a tentativas de falseamentos é a teoria dos protótipos. Esta diz que as características são formadas com base em um modelo (prototípico) da categoria. De acordo com Murphy (2002), este protótipo não é um elemento dentro da categoria, ou um exemplar, mas uma representação idealizada que abarca algumas das principais características da categoria. Sternberg (2008) afirma que o protótipo é sempre o item no qual os futuros modelos se baseiam, sendo este o melhor para definir a classe na qual a categoria se baseia.

O protótipo é caracterizado e descrito pelas categorias típicas. A categoria típica não é como uma característica definitória que precisa estar sempre presente, simplesmente está presente na maioria das vezes, assim, quanto mais características típicas o conceito tiver, mais ele se encaixa em uma categoria, quanto menos tiver, menos se encaixa. Isso implicou na diferenciação entre conceitos clássicos e conceitos vagos (*fuzzy*). Os clássicos são facilmente explicados por características definitórias, os

vagos dependem mais de um protótipo criado por associações de categorias típicas (STERNBERG, 2008).

Esta representação idealizada segue a forma de um sumário, listando características típicas e atribuindo forças a estas. Por exemplo, para um objeto ser uma arma é mais relevante que tenha a característica de danificar do que a característica de ser feita de metal. Logo, a característica de danificar tem mais peso na categoria do que a de ser feita de metal. O grau de tipicidade de uma característica também pode ser atribuído por esta ter propriedades específicas da categoria e não ter propriedades de outras categorias semelhantes (Murphy, 2002).

De acordo com Murphy (2002), a teoria dos protótipos explica as falhas da visão clássica de conceitos que seguem características definitórias. Primeiramente, nenhuma característica específica é necessária para a categorização de um item. Desde que um item tenha, por exemplo, muitas características de um cachorro podemos classificá-lo como tal, sem que haja requisitos fixos para as características preencherem.

Pela teoria de protótipos é possível entender porque alguns casos são de fronteira, nos quais não há acordo de classificação entre os falantes. Ora, se um item possui diversas características de duas categorias, não fica fácil decidir em qual este se encaixa. Dependendo da ocasião, algumas características podem ser focadas e outras esquecidas, e até mesmo isto pode influenciar na classificação. Ainda, a teoria dos protótipos está de acordo com evidências de que elementos típicos de uma categoria são mais facilmente classificados do que elementos atípicos (MURPHY, 2002).

Hampton (1982) mostra evidências de que categorias naturais não seguem leis de transitividade. Nesta pesquisa, indivíduos afirmaram que proposições de categorias eram verdadeiras, ao mesmo tempo em que concordavam com contra-exemplos. Por exemplo, afirmam que relógios antigos (*clocks*) são móveis, e também que relógios tradicionais fixados em torres (como “*Big Ben*”) também são relógios antigos, mas não são móveis. A interpretação do experimento é que ao verificar uma proposição sobre categorias, indivíduos entendem “verdade” como “geralmente verdadeiro”. Assim, utilizam informações de suas representações prototípicas para exercer o julgamento.

Na classificação de itens, estes são escolhidos com base na semelhança com o protótipo, os candidatos passam por processo de seleção, logo a categoria se torna cada vez mais estável (STERNBERG, 2008).

O Experimento de pensamento das Terras gêmeas

O experimento de pensamento das Terras Gêmeas é descrito por Putnam (1973) como uma forma de argumentar a favor da tese de que os significados não estão na cabeça. O responsável pelo significado de um termo ou conceito, de acordo com o filósofo, não é um estado psicológico associado ao objeto, mas o referente do termo no mundo, o qual é público. Por exemplo, H₂O (a referência) determina o que o termo “água” significa, mesmo que o falante não conheça a estrutura atômica da água.

Putnam (1973) pede para supormos que exista um planeta chamado Terra Gêmea. Este planeta é idêntico a Terra, a não ser por uma leve diferença, o líquido chamado de “água” não é H₂O, mas outro, com fórmula química complicada. O filósofo simplifica a fórmula para XYZ. Este líquido é aparentemente indistinguível do H₂O, e apenas com análises científicas é possível descobrir a diferença. O XYZ substitui o H₂O em todos os casos na Terra Gêmea, seja para mares, chuvas, rios.

De acordo com Putnam (1973), se uma nave da Terra algum dia chegar à Terra Gêmea, os tripulantes de início iriam supor que a palavra “água” possui o mesmo significado da Terra. Apesar disso, esta suposição seria confirmada como errada quando descobrissem que a estrutura atômica do líquido da Terra Gêmea é XYZ. Chegariam à conclusão de que o termo “água” na Terra Gêmea significa XYZ.

Continuando sua argumentação, Putnam (1973) propõe uma volta no tempo para o ano 1750 em ambos os planetas. Uma pessoa e sua cópia, mesmo desconhecendo a estrutura atômica, e mesmo estando no mesmo estado psicológico, entenderiam o termo “água” de forma diferente. Afinal, para o filósofo, a referência H₂O ou XYZ determina o significado do termo.

Putnam (1973) aplica o mesmo procedimento a um exemplo com alumínio e molibdênio para seu experimento de pensamento, mas este não traz nenhuma novidade em relação ao exemplo da água. Entretanto, um último exemplo sobre Ulmeiros e Faias será importante para a análise seguinte. O filósofo pede para supormos que ele tenha um irmão gêmeo na Terra Gêmea, o qual, assim como ele, não sabe a diferença de um Ulmeiro e uma Faia. A única diferença entre os dois é que o gêmeo utiliza o termo “Ulmeiro” para se referir a Faias. Então, estando no mesmo estado psicológico, quando o gêmeo diz Ulmeiro, o significado da palavra é faia, mas quando o verdadeiro Putnam diz Ulmeiro, o significado é Ulmeiro. Com isso Putnam acredita mostrar que o significado dos termos não é definido por nossos estados psicológicos.

Essências e o argumento da Terra Gêmea

A maioria dos exemplos que Putnam (1973) considera é exatamente do tipo que faria seu argumento funcionar. Estes exemplos incluem termos cujo referente é único e claramente distinguível, com estrutura atômica única. Apesar disso, esse tipo de termo não é abundante em nossa linguagem, mesmo se reduzirmos o escopo para apenas categorias naturais. O problema do argumento de Putnam está ligado ao problema das essências. Para o experimento de pensamento de Putnam funcionar, teríamos que admitir que nossos conceitos podem ter uma referência em forma de essência no mundo. Mesmo que nosso conceito não capture a tal essência da referência, essa essência existe, pois é ela que define o significado. Porém são poucas as coisas no mundo que realmente possuem essências.

Quando dizemos que um animal é um lobo, por exemplo, estamos dizendo que existe um grupo de características de certos organismos que podemos classificar como “lobo”, e que o indivíduo em questão possui esse grupo de características. Mas, se voltarmos em sua história evolutiva, teremos indivíduos semelhantes que não saberemos dizer se são ou não lobos. Isto porque “lobo” é um conceito (ou categoria, seja clássica ou prototípica) inventado para nos referirmos a um conjunto de animais, mas cada indivíduo é único, e alguns podem não encaixar no nosso conceito. Ainda, se considerarmos que membros da mesma espécie são aqueles que podem se cruzar e gerar filhos férteis, então não podemos dizer que um pastor alemão e um poodle sejam da mesma espécie, pois o acasalamento dos dois não gera filhos férteis e, ainda, um pastor alemão gera filhos férteis com alguns lobos enquanto poodles não os geram. Poderíamos até mesmo utilizar métodos mais precisos de dizer o que é uma espécie, como a partir de análises de DNA; conseguiríamos assim fazer uma classificação mais precisa, mas nunca chegaríamos a categorias que sejam completamente mutuamente exclusivas, isso porque cada indivíduo é único. Uma criança aprendendo a falar chama todas as aves de “passarinho”, seus conceitos são imprecisos, mas também os conceitos filosóficos e científicos ainda não chegam à precisão absoluta, pois para isso teríamos que ter um nome diferente para cada organismo (*token*) no mundo. Isto não é um problema para a teoria clássica dos conceitos e categorias, pois para esta precisaríamos apenas ter uma definição arbitrária que fosse útil ou suficiente. Mas isto é um problema para o experimento de Putnam, pois se a definição de um conceito for arbitrariamente

criada não faz mais sentido dizer que ela é definida pelo mundo, ou pela referência em si.

Dennett (1995, p.201) afirma que nada que seja complicado o bastante para ser interessante pode ter uma essência. O filósofo usa o exemplo da vida. Poderíamos voltar em uma máquina do tempo e tentar achar um momento onde a vida começou, mas seria difícil dizer o que é vida neste momento. De acordo com Dennett, Darwin acreditava que essas características de perda de essência e foco em passos graduais geram uma revolução metafísica e epistemológica que surgiu junto com sua ciência. Desde Platão e Sócrates consideramos essencial definir nossos termos para compreender o mundo, mas isso é o que *nós* precisamos fazer, o *mundo* em si não possui essências que correspondem com nossos termos.

Putnam (1973) usa exemplos perfeitos para seu argumento, pois água é sempre H₂O. A análise do filósofo depende da noção de que a maioria de nossos conceitos possui um correspondente preciso no mundo. Mas em seu último exemplo, sobre as diferenças de Ulmeiros e Faias, o qual ele considera menos, o problema vem à tona. Será que existe uma essência de Ulmeiros que correlaciona com sua referência no mundo mesmo que o falante não a conheça? Um Ulmeiro é diferente de outro, o qual também é diferente de uma Faia, não há diferença essencial, cada árvore possui sua organização atômica própria. Para Putnam, os estados mentais de duas pessoas seriam o mesmo, mas estariam falando de coisas distintas. Porém, se Faias e Ulmeiros fossem espécies próximas o bastante, poderiam até haver Faias as quais são mais próximas em sua organização atômica a certo Ulmeiro do que outros demais Ulmeiros. Não há um sentido definitivo no mundo, em termos de essência, que faça um Ulmeiro ser absolutamente distinto de Faias; cada combinação de átomos no mundo é única. Para a referência dar o significado a nossos conceitos (da forma em que o experimento dele exige) precisaríamos de um conceito para cada combinação de átomos possíveis.

Passivelmente o externalista social Tyler Burge (1982) apresente uma saída para este problema, pela proposta de que a comunidade científica ou a sociedade determina o significado de nossos conceitos. Mas o que dizer de conceitos como a vida? Como poderíamos aplicá-lo ao experimento de pensamento da Terra Gêmea? Não existe uma definição científica do que a vida é. De toda forma, conseguimos usar o conceito de forma coerente. Tudo aponta para o fato de que cada um de nós constrói categorias imprecisas baseadas em protótipos criados, utilizando algumas das propriedades dos objetos, e testando esse protótipo em relação à concepção de outras pessoas. Por isso, o

conceito de vida é um bom exemplo, já que não podemos dizer precisamente o que ela é, mas podemos pensar em coisas que podem ser vida, e coisas que não são vida. Mas não há uma forma simples de colocar elementos dentro ou fora desta categoria, por exemplo, robôs avançados podem ou não serem considerados vida? O conceito é baseado no protótipo de cada ser humano e se aperfeiçoa na medida em que buscamos propriedades dos referentes e colocamos a representação ao teste com o protótipo do outro. Mas no fim o que dá o significado ao conceito não é nenhum referente com uma essência intrínseca, nem grupos de cientistas com definições precisas, mas sim o próprio protótipo impreciso que criamos.

Suponhamos que gregos antigos pudessem encontrar os robôs mais avançados da atualidade, possivelmente classificá-los-iam como seres vivos. A referência encaixa no protótipo, esses robôs se movem e parecem ter ações com propósito. Não há nada que realmente seja “vida” no universo para que eles pudessem estar certos ou errados em sua interpretação. O protótipo representacional deles definiria a palavra e o significado.

Para uma comparação ao experimento de pensamento de Putnam, suponhamos que na Terra Gêmea existissem (além dos animais), máquinas muito avançadas, e que as pessoas chamassem estas de vida. Se pessoas da terra original fossem levadas à Terra Gêmea, algumas variações dos protótipos construídos por estas pessoas iriam aceitar essas máquinas como vida, enquanto outras não iriam aceitar. Possivelmente nenhum cientista iria poder definir quem está certo, e nem a comunidade, pois se trata de um termo sem definição precisa e que não está ligado uma referência exata no mundo (assim como muito outros termos). A referência não iria mudar em nada essencialmente, seriam robôs, mas o significado do conceito iria variar de pessoa para pessoa, dependendo da extensão que o protótipo que ela desenvolveu durante a vida abrange.

Considerações finais

O exemplo principal de Putnam do experimento das Terras Gêmeas é forte e mostra seu ponto de vista. Entretanto, apesar de mostrar que o significado pode não estar na cabeça, ele não exclui a possibilidade contrária. Por exemplo, seria possível responder ao experimento com exemplo da água, afirmando que a mudança atômica do líquido não faz nenhuma diferença para o significado do termo da palavra água. O termo se refere à forma líquida e transparente dele, e principalmente à utilidade que

possui. Assim, apesar da diferença atômica, iriam interagir com o líquido da mesma forma (beber, usar para lavar, banhar), logo o significado da palavra seria o mesmo. Já no exemplo da vida, apresentado neste artigo, a interpretação do significado na mente baseada em protótipos representacionais é a única possível, pois neste caso nem a referência e nem a definição científica são determinantes para o significado dos conceitos.

Agradecimentos

Alfredo Pereira Júnior, Jonas Gonçalves Filho e CAPES.

Referências

- BURGE, T. “Other Bodies”. In: WOODFIELD, A. (Ed.) *Thought and Object: Essays on Intentionality*. Oxford: Clarendon, 1982, p. 97-120.
- DENNETT, D. C. *Darwin’s Dangerous Idea: evolution and the meanings of life*. New York: Penguin Books, 1995.
- GRAYLING, A. C. “Epistemology”. In: BUNNIN L., TSUI-JAMES E.P. (Eds.) “The Blackwell Companion to Philosophy”, Oxford: Blackwell Publishing Ltd, 2007, p.37-60.
- MARKMAN, A. B. “Conceptual Representations in Psychology”. In: NADEL, L. (Ed.) “Encyclopedia of Cognitive Science”. New York: John Wiley & Sons, 2005. p. 1-4.
- MURPHY, G. L.; LASSALINE, M.E. “Hierarchical Structure in Concepts and the Basic Level of Categorization”. In: LAMBERTS, K.; SHANKS, D. (Eds). *Knowledge, Concepts, and Categories*. Cambridge: The MIT Press, 1997, p. 93-131.
- MURPHY, G. L. *The Big Book of Concepts*. Cambridge: The MIT Press, 2002.
- HAMPTON, J. A. “A demonstration of intransitivity in natural categories”. *Cognition* v.12, p. 151–164, 1982.
- PUTNAM, H. “Meaning and Reference”. *The Journal of Philosophy*, v.70, p.699-711, 1973.
- STERNBERG, Robert J. *Psicologia Cognitiva*. 4ed. Porto Alegre: Artmed, 2008.