

CONHECE-TE A TI MESMO: EXTERNALISMO E AUTOCONHECIMENTO DE ATITUDES PASSADAS

KNOW THYSELF: EXTERNALISM AND SELF-KNOWLEDGE OF PAST ATTITUDES

Danilo Fraga Dantas*

Resumo: Existe uma tese que dá conta da compatibilidade entre externalismo sobre conteúdo mental e autoconhecimento (BURGE, 1988). Esta tese, que explora a propriedade de autoverificação de autoatribuições do tipo “penso que p ”, porém, funciona apenas para autoatribuições de pensamentos cuja expressão é a primeira pessoa do tempo presente do modo indicativo em seu uso assertórico. Entre os casos problemáticos estão as autoatribuições no passado e as autoatribuições de atitudes proposicionais específicas. Tal falha do compatibilismo é apontada por Boghossian (1992) como prova da incompatibilidade entre externalismo e o autoconhecimento. Existe uma extensa bibliografia para dar conta das autoatribuições no passado. Em tais artigos (BURGE, 1995), defende-se a existência de uma função preservativa da memória que, salvo em situações adversas, garantiria a veracidade das autoatribuições no passado. Porém há pouca bibliografia sobre o outro problema da tese compatibilista, o das autoatribuições de atitudes proposicionais específicas. Este artigo propõe expandir o alcance da teoria da memória preservativa, para que ela dê conta do autoconhecimento de atitudes proposicionais passadas. Para tanto, propomos critérios que definam quando uma autoatribuição de pensamentos passados é verdadeira, mesmo nos casos em que atitude e conteúdo dos pensamentos não sejam idênticos, mas apenas similares. O critério para similaridade de conteúdos (que garantiria a verdade de autoatribuições deste tipo) foi proposto por Bernecker (2009), o critério para as atitudes é novo: a atitude que S pensa em t_2 como tendo tido em t_1 é similar à atitude em t_1 sse a atitude atual for implicada pela anterior.

Palavras-chave: Autoconhecimento. Memória. Metarrepresentação. Externalismo. Atitudes proposicionais.

Abstract: There is a thesis that assure the computability between externalism about mental content and self-knowledge (BURGE, 1988). However, this theses, that explore the autoverification property of claims of the type “I think that p ”, works only for assertive claims that are express in the simple present tense. Among the problematic cases are the claims in the past tense and claims about specific propositional attitude. This fails about the thesis of the compatibility is pointed by Boghossian (1992) as a prove of the incompatibility between externalism and self-knowledge. There is a wide bibliography that try to manage the claims in past tense. In these articles (BURGE, 1995), It's argued that the preservative function of memory assure the veracity of the claims in past

* Danilo F. Dantas (danfraga@gmail.com) é graduado em Comunicação e mestre em Comunicação e Cultura Contemporâneas pela Universidade Federal da Bahia (Ufba). Atualmente ele cursa o mestrado em Filosofia da Mente pela Universidade Federal do Rio Grande do Sul (Ufrgs).

tense (at last in normal situations). However, there isn't a lot of work about the problem of the claims about specific propositional attitude. This article propose a expanse the preservative memory theory, with the purpose that It manages the self-knowledge of past propositional attitudes. We propose some criteria that define when a claim of this type is true, mainly in the cases that the content of the attitude of the past thought is not identical with the actual thought. The criterion about the similarity of content is the Bernecker's one (2009). The criterion about similarity of attitudes is new: the attitude that S thinks in t_1 as if he have had in t_2 is similar with the attitude that he had in t_1 iff the actual attitude is entailed by the former.

Keywords: Self-knowledge. Memory. Meta-representation. Externalism. Propositional attitudes.

O autoconhecimento, mais especificamente a autoridade de primeira pessoa (APP), é um tema recorrente na bibliografia acerca do externalismo sobre conteúdo mental¹. O interesse na relação entre temas está fundado na suspeita da inconsistência entre eles. Externalismo é a posição segundo a qual o pensamento é parcialmente determinado por relações entre sujeito e o ambiente (físico ou social) no qual ele se encontra. De modo que o conteúdo mental não pode ser determinado apenas por propriedades intrínsecas ao sistema nervoso do sujeito (PUTNAM, 1975). A APP designa o acesso privilegiado (imediato, não-empírico, não inferencial ou mesmo discriminativo) do sujeito em relação aos seus próprios pensamentos. Tal autoridade é frequentemente expressa pelo princípio da transparência dos conteúdos mentais: se S pensa que p , S sabe (*a priori*) que é isso, e não outra coisa, o que pensa. Porém, se o conteúdo do pensamento é parcialmente determinado pelo ambiente e se o sujeito tem acesso privilegiado ao conteúdo de seus pensamentos, então o sujeito deveria ter acesso privilegiado ao ambiente em que se encontra. Esta é a conclusão, por *modus ponens*, do seguinte argumento, irrestritamente generalizável para qualquer proposição ou componente de proposição (FARIA, 2006: 108):

¹ Daqui em diante, “externalismo” refere-se a “externalismo sobre conteúdos mentais”.

1. Se penso que este copo está cheio d'água, então água é um constituinte do ambiente com o qual interajo (diretamente ou através de minha comunidade linguística).	Externalismo
2. Sei que penso que este copo está cheio d'água.	Autoridade de 1ª Pessoa
3. Se sei que <i>p</i> , então <i>p</i> .	Lei de Parmênides
4. Penso que este copo está cheio d'água.	2, 3 <i>Modus ponens</i>
5. Água é um constituinte do ambiente com o qual interajo (diretamente ou através de minha comunidade linguística).	1, 4 <i>Modus ponens</i>

Porém, como é inaceitável esta conclusão, de que o conhecimento *a priori* dos conteúdos dos pensamentos seja condição suficiente de conhecimento *a priori* dos objetos exteriores, resta regredir, por *modus tollens* e negar a conjunção entre externalismo e APP. Assim sendo, aceitar o externalismo implicaria tomar a investigação empírica do ambiente físico e social condição necessária para o autoconhecimento. Porém, dada a dificuldade de aceitar que não haja algum tipo de APP, a única opção restante seria a de negar o externalismo. E, de fato, para alguns filósofos, tal argumento nada mais é que uma redução ao absurdo do externalismo (cf. BOGHOSSIAN, 1997 e BERNECKER, 2000).

Existe, porém, uma resposta canônica a este argumento: a tese compatibilista, que foi desenvolvida por filósofos como Burge, Davidson, Shoemaker entre outros. Segundo essa tese, o autoconhecimento é possível porque uma autoatribuição do tipo “eu penso que *p*” tem a propriedade de ser autoverificável: o simples fato de o sujeito pensar que pensa que *p* torna invariavelmente verdadeiro que ele pensa que *p*, pois, ao pensar que pensa que *p*, o sujeito pensa que *p* (BURGE, 1988). Dado o caráter autoverificante de tais autoatribuições, o operador epistemológico do conhecimento vem grátis: como não é possível se enganar num caso como estes, basta que o sujeito pense que pensa que *p* para saber que pensa que *p*, mesmo que ele não discrimine o conteúdo de *p*. Deste modo, para tornar o autoconhecimento compatível com o externalismo, o compatibilismo reformula a APP: neste caso, o sujeito saberia o que pensa, mas não discriminaria o conteúdo de seu pensamento. Esta posição tem a vantagem de exibir a disparidade entre autoconhecimento e conhecimento de terceira pessoa.

Porém, o compatibilismo tem sérias limitações: a principal delas é que a incorporação do conteúdo do pensamento de primeira-ordem à autoatribuição só acontece em pensamentos cuja expressão característica é a primeira pessoa do tempo presente do

modo indicativo em seu uso assertórico. Para todos os outros casos, a diferença de perspectivas (entre pessoas, tempos ou modos) fornece uma margem para erros na atribuição². Entre estes casos, o mais discutido é o das autoatribuições no tempo passado. Nestes casos, a mudança de perspectiva temporal dá margem para erros na autoatribuições, pois se S pensa que pensou que *p*, não se segue que S tenha pensado que *p*. Seria esperado, porém, que qualquer modelo de autoconhecimento funcionasse em casos como este, pois, apesar de distintas temporalmente, estas duas perspectivas são de um mesmo sujeito. Esta falha no compatibilismo é explorada por Paul Boghossian, em seu argumento da memória: “(o modelo de Burge) leva a dizer que, apesar de S não saber amanhã o que ele está pensando agora, ele sabe agora o que ele está pensando agora (...). Porém, há um mistério aqui: pois o que se segue deveria ser uma banalidade sobre memória e conhecimento: se S sabe que *p* em t_1 , e se em t_2 S lembra de tudo o que sabia em t_2 , então S sabe que *p* em t_2 . (...) A única explicação que sugiro (...) não é que ele esqueceu algo, mas que ele nunca soube. Os julgamentos autoverificantes de Burge não constituem conhecimento genuíno (1989: 22 – 23). Assim, Boghossian pretende demonstrar a inadequação do compatibilismo e a inconsistência entre externalismo e APP.

Outro caso de falha no compatibilismo é aquele em que a autoatribuições de diferentes atitudes proposicionais³. Boa parte das pesquisas sobre o autoconhecimento deixa este caso em segundo plano. Porém, dado que um pensamento é composto tanto por seu conteúdo quanto pela atitude em relação a este conteúdo, o autoconhecimento requer conhecimento de ambos componentes. O modelo compatibilista funciona para atribuições que utilizam “pensar⁴” como operador que indica uma atitude e neste caso não há espaço para erro decorrente da atitude na autoatribuição. Mas esta garantia de autoverificação desaparece tão logo se utilize atitudes não-genéricas, tal como crença, desejo, etc. Por exemplo, se S crê que não deseja *p*, não se segue que S não deseje *p*:

2 Esta é a mesma limitação do cogito cartesiano: /eu sou, eu existo/ (1996, 17) é necessariamente verdadeiro, ao ser pensado por qualquer sujeito, apenas na forma expressa em primeira pessoa, no tempo presente do indicativo.

3 Uma atitude proposicional é uma disposição de tomar um conteúdo de determinada forma. Neste artigo, porém, quando falamos em “atitudes proposicionais”, estamos sempre nos referindo aos casos em que uma atitude proposicional é atribuída (ou autoatribuída) a um sujeito.

4 Em seu uso filosófico dominante, “pensar” tem o sentido de ter em mente. Quando Descartes (1996, 19) diz ser uma coisa que pensa, ele dá uma lista de coisas que esta coisa pode fazer, e entre elas está duvidar e desejar.

pode ser o caso de (i) S ser desatento em relação a suas atitudes mentais; de (ii) S se autenganar ao ponto de crer não ter tal desejo, quando, na verdade, o tem; ou ainda de (iii) S ter um entendimento incompleto do conceito de desejo. De modo que, dado que o conhecimento das atitudes requer posse de conceitos e que os conceitos são parcialmente determinados pelo ambiente, (iii) seria indetectável *a priori*, o que coloca novamente em questão a compatibilidade entre externalismo e APP.

Em que estavas pensando?

Em resposta às críticas de Boghossian, Burge introduziu no modelo compatibilista a noção de memória preservativa, que garantiria o status de conhecimento para autoatribuições no tempo passado. Nas palavras de Burge: “a memória não precisa ser sobre um evento ou conteúdo passado. Ela pode simplesmente, ao preservar o pensamento passado, conectá-lo ao pensamento presente. Estes casos envolvem um tipo especial de função da memória – a memória preservativa – que preserva o conteúdo proposicional e a atitude relacionada a ele, em vez de se referir a objetos, atitude, conteúdos, imagens ou eventos” (BURGE, 1998: 357). Dado que a memória preservativa apenas funciona conectando o pensamento atual ao pensamento passado, sem identificá-los com objetos particulares, seu bom funcionamento anularia a diferença de perspectiva temporal, garantindo, assim, que as autoatribuições no passado incluam os pensamentos do qual tratam. Quando, em t_2 , S pensa que (em t_1) pensou que p , através de cadeias causais, a memória preservativa conecta pensamentos atual e anterior, garantindo que o conteúdo do pensamento de t_2 seja o mesmo do pensamento de t_1 . Burge não sabe que cadeias causais são estas, porém, através de um argumento transcendental, mostra que elas precisam existir, pois a memória preservativa é um pressuposto da racionalidade: numa inferência o sujeito deve, a cada passo, ter em mente os passos anteriores e, como todo pensamento se dá no tempo, para realizar qualquer inferência o sujeito precisa fiar-se na memória preservativa.

A noção de memória preservativa também é relevante para o autoconhecimento de atitudes passadas. Se não fossem retidas as atitudes passadas, de nada adiantaria reter os conteúdos. Por exemplo, suponha que S creia que não- q e sua memória lhe disponibilize o conteúdo “se p , então q ”. Se S deve crer ou não que não- p , depende de se sua atitude passada em relação a se p então q era de crença, não de temor ou desejo, etc.

(SHOEMAKER, 1988: 187). Porém, apesar de Burge afirmar que “ a memória preservativa normalmente retém conteúdo e atitude dos pensamentos anteriores (...)” (1998: 357), ao especular sobre o modo como esta atitude é retida, ele considera duas possibilidades: ou (i) a memória preservativa retém conteúdo e atitude, ou (ii) esta retém apenas o conteúdo, enquanto a memória de particulares se encarregaria das atitudes. Porém, como a memória de particulares é discriminativa, no caso de (ii), seria fácil criar um contra-exemplo do tipo de Boghossian que colocaria em questão a consistência do modelo compatibilista e a APP (BERNECKER, 1996 e GIBBONS, 2001). Assim, a memória preservativa deve reter conteúdo e atitude.

Deste modo, o modelo compatibilista suplementado pela noção de memória preservativa teria sucesso em explicar o autoconhecimento no passado. Este modelo tem sucesso principalmente nos casos em que conteúdo e atitude da autoatribuição são idênticos aos do pensamento original. Não se segue, porém, que todo autoconhecimento seja deste tipo. Bernecker argumenta que, se identidade de conteúdo e atitude fosse requisito para todo tipo de autoconhecimento, este seria algo muito raro, principalmente nos casos flexionados no passado. O motivo é que “autoconhecimento diacrônico apoia-se na memória e a memória não é um dispositivo passivo de reprodução de conteúdo, é um dispositivo ativo que processa o conteúdo estocado” (2009: 111). Os argumentos para a necessidade de tal processamento são: (i) a fim de expressar o mesmo pensamento, a memória precisa modificar o conteúdo passado (KAPLAN 1989), (ii) muitas lembranças não se referem a um pensamento em particular, mas a um modo de pensar num período do tempo e (iii) frequentemente a memória edita a informação de acordo com estereótipos culturais de estruturação da informação (BERNECKER 2009, 120). Assim, este parece ser um caso de autoconhecimento: em t_2 , S acredita justificadamente que (em t_1) acreditou que p , porém, em t_1 , de fato, S lamentou que p . Porém, a teoria da memória preservativa não dá conta de casos como este.

Os limites da autenticidade

Segundo Bernecker (2009, 111), uma análise preliminar do autoconhecimento no passado com atitudes não-factivas seria da seguinte forma: em t_2 , S sabe que pensou (em t_1) que p , em que “pensou” está por alguma atitude não-factiva (crença, desejo, intenção, etc), se e somente se:

- (1) S acredita justificadamente em t_2 que ele pensou (em t_1) que p ;
- (2) S pensou em t_1 que p^* ;
- (3) p é idêntico a, ou suficientemente similar a, p^* ;
- (4) a atitude que S acredita em t_2 ter tido (em t_1) em relação a p é a mesma que, ou suficientemente similar a, a atitude que S teve em t_1 em relação a p^* .

No autoconhecimento de atitudes factivas (conhecimento, lembrança, orgulho), além das condições (1) – (4), também valem as seguintes condições⁵:

- (5) p em t_2 é verdade
- (6) p^* em t_1 é verdade

A *condição da crença justificada* (1) expressa que conhecimento implica crença: saber ter pensado que p implica crer ter pensado que p , em que esta crença, evidentemente, está justificada. A *condição do conteúdo* (2) expressa que autoconhecimento implica verdade, ou seja, saber que se pensou p implica ter pensando p . A *condição do conteúdo* (3) exprime que, num caso de autoconhecimento no passado, o conteúdo do pensamento atual deve ser o mesmo ou suficientemente similar ao conteúdo do pensamento original. Do mesmo modo que (3), a *condição da atitude* (4) exprime que, no autoconhecimento no passado, a atitude do pensamento atual deve ser a mesma ou suficientemente similar à atitude do pensamento original. Os argumentos em favor de (4) são os mesmo em favor de (3) – a existência do tipo de caso exposto no final da última seção. As condições (5) e (6) expressam que autoconhecimento de atitudes factivas implica que a verdade do conteúdo dos pensamentos, tanto em t_1 quanto t_2 .

O bom funcionamento das condições (3) e (4), porém, depende da definição de “similaridade” em termos menos ambíguos. Apesar parecer uma relação diádica, a “similaridade” é, de fato, uma relação entre dois indivíduos, um padrão de comparação e um grau de desvio permitido. O problema é que, quando completamente definida, a noção de similaridade torna-se supérflua. “Dizer que duas coisas são similares por terem uma propriedade específica em comum é dizer nada mais que elas têm tal propriedade em comum” (GOODMAN, 1972: 445). Deste modo, para que as condições (3) e (4)

⁵ Deste modo, enquanto o autoconhecimento no passado em geral tem a direção de ajuste mente(segunda-ordem)-mente(primeira-ordem), o autoconhecimento de atitudes factivas passadas tem em adição direção de ajuste mente(primeira-ordem)-mundo.

funcionem corretamente, seria necessário substituir a noção de similaridade por alguma outra noção menos problemática.

Conhece-te teus conteúdos

Segundo Bernecker (2009, 114), o autoconhecimento no passado requer que o conteúdo do pensamento atual seja implicado pelo conteúdo que o pensamento original. Eis a tese da implicabilidade (*ibidem*):

Tese da Implicabilidade: um pensamento de segunda-ordem em t_2 se qualifica como uma instância de autoconhecimento de um pensamento de primeira-ordem em t_1 se o conteúdo de segunda-ordem é implicado pelo conteúdo do pensamento de primeira-ordem.

A tese da implicabilidade é compatível com a estratégia compatibilista original ($p \rightarrow p$) e, como a implicação preserva a verdade, respeita o requisito de autenticidade do conhecimento. Porém, se a tese da implicabilidade for entendida a partir da noção de implicação material, esta se tornaria liberal demais para modelar o autoconhecimento no passado. Uma proposição falsa implica materialmente qualquer coisa, porém seria um resultado intolerável que uma autoatribuição do tipo “eu pensei que q ” se qualificasse como autoconhecimento de qualquer pensamento com conteúdo falso que o sujeito teve no passado. Outros paradoxos da implicação material são: toda proposição implica uma proposição verdadeira e dada qualquer duas proposições, uma implica ou é implicada pela outra⁶. Mesmo deixando de lado tais paradoxos, a implicação material ainda seria liberal demais para modelar o autoconhecimento no passado. Isto porque, uma vez que toda proposição implica infinitas outras proposições que não tratam do mesmo tema, neste caso, seriam possíveis muitos casos esdrúxulos de autoconhecimento. Por exemplo, seria um caso de autoconhecimento se, em t_2 , S pensasse que (em t_1) pensou que Sócrates morreu de causas não-naturais, quando, de fato, em t_1 , S pensou que Jabir ibn Hayyan descobriu o ácido sulfúrico.

Bernecker (2009, 115) sugere, então, que a noção de implicação a ser utilizada na tese

⁶ Nem a adoção da implicação estrita resolveria este problema, pois uma contradição implica estritamente qualquer coisa ($(p \wedge \neg p) \rightarrow q$) e um teorema é implicado por qualquer coisa ($p \rightarrow (q \vee \neg q)$).

da implicabilidade é a da lógica relevante desenvolvida por A.R. Anderson e N.D. Belnap (1975). Numa lógica relevante, p implica q se p é relevante para q , ou seja, se p for necessário numa prova de q a partir de p . Em tal sistema, é falso que $(p \rightarrow (q \rightarrow q))$, pois p é irrelevante na prova de $(q \rightarrow q)$. A intuição por trás desta regra é que, se p é necessário na derivação de q , p e q são tematicamente relacionados. Dado que, em lógica proposicional, identidade de temas é expressa por identidade de variáveis proposicionais, para que p implique q , p e q devem compartilhar ao menos uma variável proposicional (num cálculo de predicados de primeira ordem, p e q devem compartilhar ao menos uma constante individual, num cálculo de segunda ordem, p e q devem compartilhar ao menos uma constante individual ou predicativa, etc). Assim, a autoatribuição “em t_1 , pensei que Sócrates morreu de causas não-naturais” pode se qualificar como autoconhecimento do pensamento de que Sócrates morreu envenenado, mas não do pensamento de que Jabir ibn Hayyan descobriu o ácido sulfúrico. Isto porque “Sócrates morreu de causas não-naturais” (Ns) e “Sócrates morreu envenenado” (Es) compartilham a mesma constante individual (s) e seus predicados são tais que $E \subseteq N$ – o que não é o caso para “Sócrates morreu de causas não-naturais” e “Jabir ibn Hayyan descobriu o ácido sulfúrico”.

Em *Metaphysics of Memory* (2008, 158), quando ainda não havia desenvolvido a tese da implicabilidade, Bernecker propôs uma lista de “princípios da similaridade semântica”. No contexto de sua teoria atual, tais princípios funcionam como exemplos de em que o conteúdo de um pensamento é implicado pelo conteúdo de outro. Os princípios propostos por Bernecker são: (a) *princípio da substituição de artigo*, segundo o qual (i) pronomes demonstrativos e possessivos podem ser substituídos por artigos definidos ou indefinidos, (ii) artigos definidos podem ser substituídos por artigos indefinidos e (iii) artigos definidos e indefinidos podem ser omitidos; (b) *princípio da instanciamento universal*, segundo o qual proposições universais podem ser substituídas por proposições sobre indivíduos, desde que o indivíduo em questão pertença ao domínio do discurso; (c) *princípio da generalização existencial*, segundo o qual proposições sobre indivíduos podem ser substituídas por proposições existenciais, desde que o indivíduo em questão pertença ao domínio do discurso; (d) *princípio da eliminação da conjunção*, segundo o qual proposições conjuntivas podem ser substituídas por um de seus conjuntos; (e) *princípio da universalização conceitual*, segundo o qual conceitos podem ser substituídos por outros conceitos, desde que o primeiro seja uma nota do segundo; (f)

princípio da omissão, segundo o qual pode-se omitir informação contextualmente irrelevante; (g) *princípio da eliminação da negação*, segundo o qual uma dupla negativa pode ser substituída por uma declaração afirmativa.

O que há em comum entre a tese da implicabilidade e os princípios da similaridade semântica é que, em ambos os casos, a proposição consequente tem uma quantidade de informação menor que antecedente. E, de fato, este é um requerimento do autoconhecimento não-inferencial. Neste caso, além da exigência de que o conteúdo do pensamento atual seja implicado (de modo relevante) pelo conteúdo do pensamento passado, não deve haver qualquer premissa adicional necessária ou utilizada pelo agente para derivar o pensamento atual a partir do passado. Assim sendo, a *condição do conteúdo* (3) para autoconhecimento no passado deve ser revisada da seguinte forma:

(3') p é implicado por p^* ;

em que “implicado” deve ser entendido na linha da lógica relevante. No caso de autoconhecimento não-inferencial existe ainda o requerimento de que nenhuma premissa adicional seja necessária ou utilizada pelo agente para derivar p e p^* . Satisfeitas estas condições, há autoconhecimento, seja o conteúdo da autoatribuição qualquer fração do conteúdo do pensamento passado, não importando quão pequena seja, ou qualquer implicação relevante do conteúdo do pensamento original, não importa quão irrelevante esta pareça, tal atribuição pode ser qualificada como autoconhecimento (BERNECKER 2008, 162).

Conhece-te tuas atitudes

Para tratar da similaridade entre atitudes, Bernecker propõe um critério que, apesar de independente da tese da implicabilidade, espera seguir a ideia por trás desta tese: para que uma autoatribuição a respeito do passado seja qualificada como autoconhecimento, a atitude que, em t_2 , S pensa ter tido em t_1 e a atitude que ele, de fato, teve em t_1 devem ou ser a mesma ou compartilhar direção de ajuste e polaridade (2009: 118). Deste modo, a *condição da atitude* (4) deveria ser revisada da seguinte forma:

(4') a atitude que, em t_2 , S pensa como tendo tido, em t_1 , em relação a p e a atitude que, de fato, em t_1 , S teve em relação a p^* devem ou ser a mesma ou devem (i) compartilhar direção de ajuste e (ii) polaridade.

guma para não. Para entender como esta condição funciona, suponha que, em t_2 , S acredite ter acreditado (em t_1) que ele, no futuro, seria um bom filósofo. Porém, a atitude que S realmente teve em t_1 em relação ao conteúdo “que eu seja um bom filósofo”, foi, de fato, a expectativa. Segundo os critérios de Bernecker, este poderia ser um caso de autoconhecimento, pois crença e expectativa têm direção de ajuste mente-mundo e polaridade positiva.

Porém, para que o autoconhecimento seja qualificado como não-inferencial, Bernecker adiciona alguns requerimento: (iii) a polaridade da atitude em t_2 não pode ser mais intensa que a da atitude em t_1 ; (iv) se a atitude em t_2 for factiva, então a atitude em t_1 também deve o ser; e (v) a atitude em t_1 não pode ser um componente da atitude em t_2 , nem vice-versa. Neste sentido, o exemplo acima, não seria um caso de autoconhecimento não-inferencial. Isto porque a polaridade da crença é mais intensa que a da expectativa. Porém, nem de longe este é um resultado aceitável: o exemplo acima não só é um caso de autoconhecimento não-inferencial, como pode ser interpretado como tendo atitudes idênticas, pois, em seu sentido mais comum, uma expectativa nada mais é que uma crença sobre o futuro. (Na verdade, segundo os critérios de Bernecker para o conteúdo, tal exemplo não seria um caso de autoconhecimento de qualquer tipo, pois que p será o caso não implica que p seja o caso). Este, porém, não é o único problema enfrentado pela proposta de Bernecker.

O requisito (i) da condição (4') funciona para atitudes com direção de ajuste mente-mundo (conhecimento, crença, suposição etc.) e mundo-mente (desejo, esperança, intenção, etc), porém não pode funcionar para atitudes que não têm direção de ajuste. Se uma atitude tem direção mente-mundo, suas condições de satisfação exigem que o conteúdo da atitude espelhe um estado de coisas do mundo. Se uma atitude tem direção de ajuste mundo-mente, por outro lado, suas condições de satisfação envolvem uma modificação do mundo a ponto dele corresponder ao conteúdo proposicional da atitude. Porém, para algumas atitudes proposicionais (orgulho, vergonha, culpa, etc), não se coloca em questão a satisfação de seu conteúdo, simplesmente porque esta é dada como certa. Por exemplo, se S envergonha-se de p e p não é e nem nunca será o caso, não se diz do pensamento de S que ele é falso ou que não pode ser satisfeito, mas que ele é mal-formado. De atitudes como estas, diz-se que elas não têm direção de ajuste. Porém,

uma vez que tais atitudes não têm direção de ajuste, não pode ser o caso de qualquer outra atitude ter *a mesma* direção de ajuste que elas. Mas, existem casos de autoconhecimento de atitudes não idênticas, em que a atitude passada não tem direção de ajuste: em t_2 , S diz que (em t_1) acreditou que p e, em t_1 , ele envergonhou-se de p . Segundo Bernecker, este não é um caso de autoconhecimento, porém quem se orgulha de algo deve crer que este algo é o caso.

O requisito (ii) da condição (4') está correto: no autoconhecimento, uma atitude negativa não pode representar uma atitude positiva, ou vice-versa. Não é um caso de autoconhecimento, por exemplo, se, em t_1 , S duvide que p e, em t_2 , S diz que (em t_1) acredite em p .

Como foi demonstrado pelo exemplo da expectativa, o requisito (iii) não funciona para todo caso de autoconhecimento. Existem casos em que, se tomada independentemente do conteúdo, a polaridade da atitude da autoatribuição em t_1 é mais intensa que a da atitude em t_1 . Porém, tal conclusão contra-intuitiva não tem lugar se se levar em consideração a relação entre atitude e conteúdo. Pois, apesar da polaridade da crença ser mais forte que a da expectativa, seu conteúdo no futuro, iguala sua polaridade à da expectativa. O que este exemplo nos mostra é que nenhuma proposta que trate atitudes proposicionais independente de seus conteúdos pode funcionar. Isto se dá porque, apesar algumas atitudes terem critérios específicos quanto ao conteúdo (o conteúdo da lembrança, por exemplo, deve estar no passado), e estes critérios nem sempre são o caso para toda atitude que represente a atitude original.

O requisito (iv) da condição (4') está correto: dado que atitudes factivas têm condições mais estritas que atitudes não-factiva, se a atitude em t_2 for factiva, para que o autoconhecimento seja não-inferencial, deve também ser o caso da atitude em t_1 ser factiva.

Outro problema da proposta de Bernecker está no requisito (v). Na verdade, há algo estranho em sua exposição⁷: na mesma página, Bernecker afirma que “uma atitude completa poderia ser representada como sendo uma das atitudes simples que a compõe” e que “a atitude em t_1 não é um componente da atitude em t_2 , nem vice-versa” (2009:

7 Salvo o caso de má interpretação.

119). A única interpretação que mantém a consistência do texto é a seguinte: Bernecker defende que, para o autoconhecimento em geral, uma atitude complexa pode ser pensada como uma de suas atitudes componentes, mas, para o autoconhecimento não-inferencial, isto não pode ser o caso. Assim, este exemplo deveria lido assim: “suponha que eu saiba em t_1 que p . Em t_2 , eu digo ter acreditado em t_1 que p . *Se eu sei que conhecimento implica crença*, não há razão alme garantir o conhecimento (*não-inferencial*) de minha atitude passada” (*ibidem* – grifo e comentário acrescentados). Porém não parece haver qualquer justificativa para a introdução do condicional “se eu sei que conhecimento implica crença”. Se o conteúdo dos conceitos são parcialmente determinados pelo ambiente (natural ou social), porque seria necessário, para que uma atitude implicasse outra, que o sujeito tenha conhecimento da implicação? Se há atitudes complexas, sua relação com as atitudes que a compõem é uma relação lógica, não psicológica.

Por fim, além de incompleta, há a suspeita de que a proposta de Bernecker seja incorreta. Eis um caso que levanta tal suspeita: se, em t_2 , S diz que (em t_1) desejou p , mas, de fato, em t_1 , ele teve a intenção de p . Segundo Bernecker, esta autoatribuição de S é um caso de autoconhecimento⁸, caso as outras condições estejam satisfeitas. Mas, nem de longe este é um caso de autoconhecimento não-inferencial. Em *Ética a Nicômaco* (III.1.1110a14-20), Aristóteles, discute casos em que o sujeito tem a intenção de fazer algo, mas, como o faz devido a uma coerção externa, não deseja seu fim. Se este for o caso do exemplo acima, tal autoatribuição não parece ser um caso de autoconhecimento (o que seria um contra-exemplo à teoria).

Existe, porém, uma proposta para a similaridade entre atitudes, que, além de mais elegante, resolve todos os contra-exemplos apresentados. Esta proposta é uma versão da tese da implicabilidade, o que foi rejeitado por Bernecker: “uma vez que atitudes não implicam umas as outras, não podemos utilizar a tese da implicabilidade para explicar a noção de similaridade entre atitudes” (2009: 118). Bernecker, porém, não dá qualquer motivo para que atitudes não impliquem atitudes. E, em determinada parte do texto, implicitamente, afirma o contrário: “existem atitudes simples e existem atitudes

⁸ Isto porque desejo e intenção tem a mesma direção de ajuste (mundo-mente) e polaridade (positiva); a polaridade do desejo não é mais intensa que a da intenção, nenhuma das duas atitudes é factiva e desejar que p não faz parte de intencionar p , nem vice-versa.

complexas que são compostas das simples” (*ibidem*). Porém, o simples fato de uma atribuição de atitude proposicional utilizar um conceito de atitude complexo basta para que ela implique atribuições que utilizam cada um de seus conceitos componentes. Além disso, que atitudes implicam atitudes é a posição padrão, vide, por exemplo, a tese clássica de que conhecimento implica crença. Que atitudes implicam atitudes é defendido, por exemplo, por John Searle e Daniel Vanderveken: “certos pares de forças ilocucionárias F_1 , F_2 são de tal modo que não é possível um falante⁹ executar um ato ilocucionário da forma $F_1(p)$ sem executar um ato da forma $F_2(p)$. Nestes casos, dizemos que F_1 implica ilocucionalmente F_2 e escrevemos $F_1 \Rightarrow F_2$ ” (1985: 129 – 130). Deste modo, não parece haver qualquer razões para rejeitar a possibilidade de uma tese da implicabilidade para atitudes. O que se propõe é a seguinte:

Tese da implicabilidade (para atitudes): a atitude proposicional que S pensa (em t_2) como tendo tido (em t_1) em relação a p é suficientemente similar à atitude que S teve (em t_1) em relação a p^* se e somente se a atitude proposicional do estado psicológico posterior for implicada pela atitude proposicional do estado anterior.

Deste modo, a *condição da atitude* (4) deveria ser revisada da seguinte forma:

(4'') a atitude que S pensa em t_1 como tendo tipo (em t_1) em relação a p deve ser implicada pela atitude que, de fato, S teve em t_1 em relação a p^* .

Em princípio, tal proposta só funcionara para atitudes complexas. Porém, existe uma hipótese de que, sendo crença e o desejo as atitudes básicas, que corresponderiam às grandes categorias tradicionais da cognição e da volição, toda atitude poderia ser completamente analisada numa conjunção de crenças e desejos (2002, 40). Em *Intencionalidade* (*idem*), Searle testou tal hipótese e mostrou que, salvo algumas exceções (expectativa), ela é falsa. No decorrer da prova, porém, Searle demonstra que, apesar de não poderem sempre ser analisada em termos de crença e desejo, toda atitude contém alguma crença ou desejo. “Em poucas palavras, a hipótese defendida por nossa breve discussão não é a de que todas ou mesmo várias formas de intencionalidade podem ser

9 Não parece indevido deixar de lado as diferenças advindas do caráter pragmático dos atos ilocucionários e o caráter lógico das atitudes proposicionais. O próprio Searle, em outro texto, afirma que: “na realização de cada ato ilocucionário com um conteúdo proposicional, expressamos um certo estado intencional com este conteúdo proposicional” (2002, 12).

reduzidas a crenças e desejos – o que é claramente falso –, mas sim a de que todos os estados intencionais, mesmo aqueles desprovidos de uma direção de ajuste e aqueles que não têm uma proposição completa por conteúdo, não obstante *contém* uma crença ou um desejo, ou ambos” (*idem*: 49). Deste modo, toda atitude implica alguma crença, desejo, ou ambos, ou outra atitude menos complexa.

O método de análise de atitudes é, a princípio, bastante singelo: “tome um tipo específico de estado intencional com um conteúdo proposicional específico. Em seguida, pergunte a si mesmo em que deve acreditar e o que deve desejar para estar imbuído desse estado intencional com esse conteúdo” (*idem*: 46). Seguem algumas análises¹⁰ de Searle (*idem*, 42) e algumas consequências das análises. O caso mais simples é o da expectativa, que nada mais é que uma crença a cerca do futuro:

$$\text{Expectativa}_S(p) \leftrightarrow \text{Cren}_S(Fp)$$

O desapontamento é mais complicado. Se S está desapontado porque p , S deve ter tido antes a expectativa de que não- p e agora acreditar que p :

$$\text{Desapontamento}_S(p) \rightarrow (\text{Cren}_S(p) \wedge (P(\text{Cren}_S(F\neg p)) \wedge \text{Dej}_S(\neg p)))$$

$$\text{Logo: Desapontamento}_S(p) \rightarrow (\text{Expectativa}_S(\neg p))$$

O lamento é também relativamente simples. O pesar, porém, impõe uma restrição extra ao lamentar, pois o conteúdo do pesar deve dizer respeito a coisas que tenham a ver com a pessoa que está pesarosa.

$$\text{Lamentar}_S(p) \rightarrow (\text{Cren}_S(p) \wedge \text{Dej}_S(\neg p))$$

$$\text{Pesar}_S(p) \rightarrow (\text{Cren}_S(p) \wedge (\text{Dej}_S(\neg p) \wedge \text{Cren}_S(p \text{ está relacionado a } S)))$$

$$\text{Logo: Pesar}_S(p) \rightarrow \text{Lamentar}_S(p)$$

A culpa acrescenta ao lamento o elemento da responsabilidade. O remorso é como a culpa, só que dirigido ao próprio agente. Neste sentido, o remorso envolve o culpar a si mesmo:

10 No que se segue: “ $\text{Cren}_S(p)$ ” (S crê que p) e “ $\text{Dej}_S(p)$ ” (S deseja que p) serão tomados como operadores modais que operam sobre proposições completas, ou sobre fórmulas que contenham proposições completas. No mais, nas análises das atitudes, utiliza-se os operadores modais clássicos “ $\Box p$ ” (é necessário que p), “ $\Diamond p$ ” (é possível que p), “ Fp ” (será o caso que p), “ Pp ” (foi o caso que p), etc.

$Culpar_S X \text{ por } (p) \rightarrow (Cren_S (p) \wedge (Dej_S (\neg p) \wedge Cren_S (X \text{ é responsável por } p)))$

$Remorso_S (p) \rightarrow (Cren_S (p) \wedge (Dej_S (\neg p) \wedge Cren_S (S \text{ é responsável por } p)))$

Logo: $(Culpar_S X \text{ por } (p) \rightarrow Lamentar_S (p)) \wedge (Culpar_S S \text{ por } (p) \leftrightarrow Remorso_S (p))$

A esperança requer uma incerteza quanto à possibilidade de que o estado esperado de fato se verifique. Deste modo:

$Esperança_S (p) \rightarrow ((\neg Cren_S (p) \wedge \neg Cren_S (\neg p)) \rightarrow (Cren_S (\diamond p) \wedge Dej_S (p)))$

O orgulho e a vergonha exigem alguma relação com o agente, embora esta relação não precise ser tão forte quanto a responsabilidade. Além disso, o orgulho envolve satisfação e o desejo de tornar conhecido, enquanto a vergonha envolve insatisfação e o desejo de ocultar. Assim:

$Orgulho_S (p) \rightarrow ((Cren_S (p) \wedge Dej_S (p)) \wedge (Cren_S (p \text{ está relacionado a } S) \wedge Dej_S (p \text{ seja sabido pelos outros})))$

$Vergonha_S (p) \rightarrow ((Cren_S (p) \wedge Dej_S (\neg p)) \wedge (Cren_S (p \text{ está relacionado a } S) \wedge Dej_S (\neg(p \text{ seja sabido pelos outros}))))$

$Satisfação_S (p) \rightarrow Cren_S (p) \wedge Dej_S (p)$

Logo: $Orgulho_S (p) \rightarrow Satisfação_S (p)$ ¹¹

Conclusões

O modelo compatibilista suplementado pela noção de memória preservativa dá uma explicação plausível de boa parte dos casos de autoconhecimento no passado, inclusive daqueles que têm atitudes proposicionais não-genéricas. Este modelo, porém, não consegue dar conta de outros casos de autoconhecimento, a saber, daqueles em que não há identidade, mas apenas similaridade, entre conteúdo ou atitude entre autoatribuição e o pensamento original. Para dar conta destes casos, Bernecker (2009) propôs critérios para definir quando conteúdo e atitude são similares o suficiente para haver autoco-

¹¹ É fácil perceber também como essas análises levam em consideração a estrutura formal dos estados intencionais de segunda (terceira, quarta, enésima) ordem. O procedimento é simples: substituir p pelo resultado da análise de outro estado intencional.

hecimento. No caso da similaridade de conteúdo, foi aceita a tese da implicabilidade, que modela a relação de similaridade a partir da relação de implicação relevante. Em relação à similaridade entre atitudes, entretanto, a proposta de Bernecker dava margem a muitos contra-exemplos e, por isso, foi rejeitada. Em seu lugar, foi proposta a tese da implicabilidade (para atitudes), que dá conta de todos os casos não explicados pela proposta anterior.

É o caso das atitudes sem direção de ajuste. Se, em t_2 , S diz que (em t_1), ao passar por uma rua escura, acreditou que era possível que ele fosse roubado, e, em t_1 , ao passar por uma rua escura, ele, de fato, temeu ser roubado, segundo os critérios expostos e defendidos aqui, há autoconhecimento. Isto porque um homem que teme que p , deve acreditar que é possível que p e deve querer que seja o caso que não- p . O temor é analisado da seguinte forma:

$$\text{Temor}_S(p) \rightarrow \text{Cren}_S(\diamond p) \wedge \text{Dej}_S(\neg p)$$

Espera-se que, a partir destas análises e da tese da implicabilidade, seja possível explicar todos os casos de autoconhecimento no passado com atitudes e conteúdos não-identicos, inclusive aqueles que o modelo de Bernecker falha em explicar. Deste modo, apesar desta proposta ser completamente independente do externalismo ou da teoria da memória preservativa, espera-se ela ajude a resolver o impasse entre externalismo e APP, além de contribuir para a parca bibliografia sobre as atitudes proposicionais.

Referências

- ANDERSON, A. R., & BELNAP, N. D. *Entailment: the logic of relevance and necessity*. Princeton: Princeton University Press, 1975.
- ARISTÓTELES; *The complete works of Aristotle*. Princeton: Princeton University Press, 1984.
- BERNECKER, Sven. Self-knowledge and the bounds of authenticity. *Erkenn*, 71, p. 107–121, 2009.
- _____. *The metaphysics of memory*. Dordrecht, Holanda: Springer, 2008.
- _____. Knowing the world by knowing one's own mind. *Synthese*, 123, p. 1-34, 2000.
- _____. Externalism and the attitudinal component of self-knowledge. *Noûs*, 30, p. 262–275, 1996.
- BOGHOSSIAN, Paul. What the externalist can know a priori. *Proceedings of the Aristotelian Society*, 47, p. 161-75, 1997.

- _____. Externism and inference. *Philosophical Issues*, 2, p. 11-28, 1992.
- _____. Content and self-knowledge. *Philosophical Topics*, 17, p. 5-26, 1989.
- BURGE, Tyler. Memory and self-knowledge. In P. LUDLOW & N. MARTIN (Eds.). *Externalism and self- knowledge*. Stanford: CSLI Publications, 1998. p. 351–370.
- _____. Content preservation. *Philosophical Issues*, 6, p. 271– 300, 1995.
- _____. Individualism and self-knowledge. *Journal of Philosophy*, p. 85, 649–663, 1988.
- DESCARTES, René. *Meditations on first philosophy*. Cambridge: Cambridge University Press, 1996.
- FARIA, Paulo. A preservação da verdade. *O que nos faz pensar*, 20, p. 101-126, 2006.
- GIBBONS, John. Externalism and knowledge of the attitudes. *The Philosophical Quarterly*, 51, p. 13 - 28, 2001.
- GOODMAN, Nelson. Seven strictures on similarity. In GOODMAN, Nelson (Ed.). *Problems and projects*. Indianópolis: Bobb-Merrill, 1972. p. 437–446.
- PUTNAM, Hilary. The meaning of meaning. *Philosophical Papers*, Vol. II: Mind, Language, and Reality, Cambridge: Cambridge University Press, 1975.
- SEARLE, John. *Intencionalidade: um ensaio em filosofia da mente*. São Paulo: Martins Fontes, 2002.
- SEARLE, J & VANDERVEKEN, D. *Foundations of illocutionary logic*. Cambridge: Cambridge University Press, 1985.
- SHOEMAKER, Sidney. On knowing one's own mind. *Philosophical Perspectives*, 2, p. 183-209, 1988.

Artigo recebido em: 16/12/09

Aceito em: 27/03/10