

SERÁ PROCEDENTE O ARGUMENTO DE PLANTINGA CONTRA O NATURALISMO METAFÍSICO?

Domingos Faria

Doutorando na Universidade de Lisboa

Bolsista Fundação para Ciência e Tecnologia (Portugal) (SFRH/BD/85051/2012)

Membro da LanCog Research Group

Natal, v. 22, n. 39
Set.-Dez. 2015, p. 121-139

Princípios
Revista de filosofia

E-ISSN: 1983-2109



Resumo: O naturalismo metafísico é a perspectiva de que não há uma pessoa como Deus, nem algo semelhante a Deus. Alvin Plantinga tem argumentado que esta perspectiva é auto-refutante e, por isso, irracional. Isto porque se o naturalismo metafísico fosse verdadeiro (em conjunção com a teoria da evolução), então teríamos fortes razões para duvidar da fiabilidade das nossas faculdades cognitivas e, por conseguinte, teríamos fortes razões para duvidar da verdade de qualquer das nossas crenças, inclusive do naturalismo metafísico. O meu objetivo neste artigo é examinar criticamente este argumento de Plantinga contra o naturalismo metafísico e analisar se é capaz de lidar com algumas objeções.

Palavras-chave: Alvin Plantinga; Naturalismo Metafísico; Argumento Evolucionista contra o Naturalismo; Filosofia Analítica da Religião.

Abstract: Metaphysical naturalism is the view that there is no such person as God, nor anything like God. Alvin Plantinga has argued that this view is self-defeating and, for that reason, it is irrational. This is because if metaphysical naturalism were true (in conjunction with evolutionary theory), then we would have compelling reasons for doubting the reliability of our cognitive faculties and, consequently, we would have compelling reasons for doubting the true of any of our beliefs, including metaphysical naturalism. My aim in this paper is to critically assess Plantinga's argument against metaphysical naturalism and to survey whether it is able to handle some objections.

Keywords: Alvin Plantinga, Metaphysical Naturalism, Evolutionary Argument against Naturalism, Analytical Philosophy of Religion.

O naturalismo metafísico é a perspectiva de que não há uma pessoa como o Deus da religião teísta, nem algo semelhante a Deus. Mas será racional aceitar o naturalismo metafísico? Alvin Plantinga concebeu um argumento que procura mostrar que esta perspectiva naturalista, em conjunção com a teoria da evolução, é inconsistente e auto-refutante; assim, é irracional aceitar o naturalismo. Uma vez que as primeiras versões deste argumento (cf. Plantinga 1991; 1993; 2000) suscitaram fortes objeções (cf. Beilby 2002), Plantinga voltou a formular recentemente o seu argumento (cf. Plantinga 2008; 2011; 2012) de modo a ultrapassar essas refutações. Neste artigo tenho como objetivo examinar criticamente se esta última formulação do argumento contra o naturalismo é procedente. Para concretizar esse objetivo começarei por expor a última versão do argumento de Plantinga na primeira secção deste texto. Após essa exposição, na segunda secção formularei e discutirei algumas objeções desafiantes para esta última versão do argumento. Com isto vou procurar mostrar que a principal premissa do argumento de Plantinga, a premissa (1), parece ser falsa.

1. Argumento de Plantinga contra o naturalismo metafísico

No seu argumento, Plantinga pressupõe que os naturalistas são todos ou quase todos materialistas acerca das pessoas humanas. Ora, de acordo com esta concepção materialista as pessoas são entendidas inteiramente como objetos materiais, sem qualquer alma ou eu imaterial. Deste modo, para o propósito da argumentação, assimila-se o materialismo ao naturalismo. Tendo em conta estes pormenores, o argumento original de Plantinga é o seguinte, onde N é *naturalismo*, E é *a atual teoria evolucionista*, e F é a proposição de que *as nossas faculdades cognitivas são fiáveis* (ou seja, que produzem uma preponderância considerável de crenças verdadeiras em comparação com as falsas). Assim, pode-se representar o argumento contra o naturalismo metafísico desta forma:

(1) $P(F|N\&E)$ é baixa.

- (2) Quem aceitar N&E e também vir que (1) é verdadeira, tem um anulador para F.
- (3) Este anulador não pode ser anulado.
- (4) Quem tem um anulador para F, tem um anulador para qualquer crença que julgue ser produzida pelas suas faculdades cognitivas, incluindo N&E.
- (5) Logo, N&E anula-se a si mesmo e não pode ser racionalmente aceite.

Vejam os uma breve fundamentação da premissa (1), pois é aquela que merece maior atenção. Para esta premissa, que afirma que *a probabilidade condicional de F dada a conjunção de N e E é baixa*, Plantinga começa por argumentar que uma *crença*, do ponto de vista do materialismo, é um evento ou estrutura no sistema nervoso. Ora, uma tal estrutura terá pelo menos dois tipos de propriedades: propriedades neurofisiológicas NF, como ter determinadas relações neuronais, e propriedades mentais, como ter um conteúdo *tal e tal* – assim, a crença será a crença que *p* para uma determinada proposição *p* (em que *p* será o conteúdo da crença). De um ponto de vista materialista, há três formas nas quais estas propriedades NF e mentais se podem relacionar: materialismo não-redutor lógico (MNL), materialismo não-redutor nomológico (MNN), e materialismo redutor (MR). De acordo com o materialismo não-redutor as propriedades mentais são sobrevenientes relativamente a propriedades NF, em que tal sobreveniência pode ser entendida desta forma:

- (S) Para quaisquer mundos possíveis *M* e *M** e quaisquer estruturas *E* e *E**, se *E* tem as mesmas propriedades NF em *M* que *E** tem em *M**, então *E* tem as mesmas propriedades mentais em *M* que *E** tem em *M**.

No entanto, este entendimento da sobreveniência é suscetível de duas interpretações. Por um lado, caso se interprete a sobreveniência como envolvendo necessidade lógica ampla, então em (S) quantificamos sobre mundos possíveis *simpliciter* – assim, temos a versão de materialismo não-redutor lógico (MNL). Por outro lado, caso se interprete a sobreveniência como envolvendo necessidade nômica (i.e. o tipo de necessidade que têm as leis da natureza), então em (S) quantificamos apenas sobre mundos nomicamente possíveis – deste modo, temos a versão de materialismo não-redutor nômico (MNN). Por fim, no caso do materialismo redutor (MR) a relação entre propriedades NF e propriedades mentais não é de sobreveniência, mas sim de identidade.

Após ter clarificado estas noções, Plantinga apresenta a justificação central para (1) a partir de uma *experiência mental*: imagine-se uma população de criaturas hipotéticas que se assemelha a nós na posse de crenças e suponha-se que o naturalismo é válido para elas, bem como admita-se que surgiram por processos referidos pela teoria da evolução contemporânea como a seleção natural e a mutação genética aleatória. Pode-se perguntar, então, qual é a $P(F|N\&E)$ com respeito a estas criaturas hipotéticas, considerando-se essa probabilidade segundo cada uma das três teorias anteriores acerca das relações das propriedades mentais e NF. Começando pelo MNL, questione-se: qual é a $P(F|N\&E\&MNL)$?

Suponha-se que as criaturas hipotéticas da experiência mental evoluíram; por isso, pode-se afirmar que o seu comportamento é adaptativo (pelo facto de promover a sobrevivência e o sucesso reprodutivo) e, dessa forma, as propriedades NF que causam esse comportamento também são adaptativas. Aliás, são adaptativas porque a evolução e o seu mecanismo de seleção natural recompensa o comportamento adaptativo e pune o comportamento desadaptado, bem como recompensa as propriedades NF que causam comportamento adaptativo e pune as propriedades NF que causam comportamento desadaptado. Todavia, não importa para a adaptabilidade dessas propriedades NF que o conteúdo que daí sobrevém seja verdadeiro ou falso, pois quer esse conteúdo seja

verdadeiro quer seja falso não faz diferença para a adaptação; por outras palavras, a seleção natural visa o comportamento adaptativo, mas não se importa com a crença verdadeira, tal como aliás parece ser de certa forma sugerido por Charles Darwin (1881, p. 315-316), por Patricia Churchland (1987, p. 548) ou por Thomas Nagel (1989, p. 79). Portanto, pelo facto dessas criaturas hipotéticas terem crenças adaptativas, que causam comportamento adaptativo devido a terem propriedades NF adaptativas, não há razão para pensar que o conteúdo sobreveniente a tais propriedades NF tenha de ser verdadeiro.

Por exemplo, tome-se qualquer crença particular C possuída por uma dessas criaturas hipotéticas. Qual é a $P(C | N\&E\&MNL)$? Pode-se supor que C é adaptativa pelo facto das suas propriedades NF serem adaptativas. Porém, daí nada se segue sobre a verdade ou falsidade do conteúdo que sobrevém a essas propriedades NF, uma vez que tais propriedades NF são selecionadas pela evolução e seleção natural, não porque causam o conteúdo que causam, mas sim porque causam comportamento adaptativo. Assim, se o conteúdo sobreveniente é verdadeiro, tudo bem; mas se é falso, tudo bem na mesma – isto porque a sua falsidade não impede nem prejudica de forma alguma a adaptabilidade das propriedades NF. Desde que essas propriedades NF sejam adaptativas, não importa para a sobrevivência e reprodução que conteúdo é determinado por tais propriedades NF. Ou seja, C tanto pode ser verdadeira como falsa sem comprometer em nada a sua adaptabilidade. Deste modo, a probabilidade de C ser verdadeira, dado $N\&E\&MNL$, é de cerca de $\frac{1}{2}$. Agora, se imaginarmos que essa criatura hipotética tem apenas 100 crenças independentes e a probabilidade de cada uma delas é de $\frac{1}{2}$, a probabilidade que $\frac{3}{4}$ delas sejam verdadeiras (que é um requisito modesto para a fiabilidade) será por volta de 0.000001. Portanto, $P(F | N\&E\&MNL)$ é baixa, sendo que algo similar sucede com MNN pelas mesmas razões (apesar da sobreveniência em causa ser diferente).

E quanto ao MR em que se advoga que as propriedades mentais são apenas propriedades NF (ou complexas combinações de

propriedades NF)? Qual é a $P(F|N\&E\&MR)$? A probabilidade em questão será igualmente baixa; pois, considere-se qualquer crença particular C possuída por uma das criaturas hipotéticas da experiência mental. Ora, C é um evento neuronal que exhibe propriedades NF. Do mesmo modo, pode-se supor novamente que C é adaptativa pelo facto de ter propriedades NF que causam comportamento adaptativo. Mas, uma vez que esse evento neural é uma crença, essas propriedades NF (ou algum subconjunto dessas propriedades) também constituem o conteúdo que C de facto exhibe. Então, dado N&E&MR, qual é a probabilidade do conteúdo de C ser verdadeiro?

Esse conteúdo não tem de ser verdadeiro para que as propriedades NF causem o tipo adequado de comportamento. É verdade que para haver esse tipo adequado de comportamento, que promove a sobrevivência e o sucesso reprodutivo, é preciso de facto haver propriedades NF adaptativas. E segundo MR, essas propriedades NF (ou algumas combinações complexas dessas propriedades) também constituem a propriedade de possuir um conteúdo particular. Todavia, não importa, com respeito à adaptabilidade dessas propriedades NF, se o conteúdo que elas constituem é verdadeiro ou falso; pois, mesmo que seja falso não compromete de forma alguma a sua adaptabilidade.

Portanto, o conteúdo que é constituído tanto pode ser verdadeiro como falso (sem pôr em causa a adaptabilidade) e, se é isso o que sucede, então a probabilidade do conteúdo de C ser verdadeiro, dado N&E&MR, é de cerca de $\frac{1}{2}$. Ora, se tal probabilidade se aplica a cada uma das crenças independentes das criaturas hipotéticas, então teremos de classificar a fiabilidade das suas faculdades cognitivas (i.e. a probabilidade dessas faculdades produzirem pelo menos $\frac{3}{4}$ de crenças verdadeiras) como baixa, tal como nos casos anteriores, o que significa que a $P(F|N\&E\&MR)$ é baixa. Em suma, ao considerarmos cada uma das três teorias materialistas acerca das relações entre propriedades mentais e físicas, constatamos que a $P(F|N\&E)$ com respeito às criaturas hipotéticas

é baixa. Mas se é isso o que ocorre, o mesmo valerá para nós, ficando assim justificada a premissa (1).

O próximo passo do argumento de Plantinga, a premissa (2), evidencia que se uma pessoa reconhecer que a $P(F|N\&E)$ é baixa, e também aceitar N&E, então ela tem um anulador para F, isto é, uma razão para rejeitar F. Além disso, de acordo com a premissa (3), esse é um anulador que não pode ser ele mesmo anulado. Pois, um *anulador de anulador*, i.e., um anulador para esse anulador teria, por exemplo, de ter a forma de um argumento a favor de F. Todavia, (i) um tal argumento seria epistemicamente circular, uma vez que ao argumentar a favor de F já estou a pressupor que F é verdadeira e, igualmente, (ii) quem tem um anulador para F também tem um anulador para qualquer argumento a favor de F. Por fim, como está patente na premissa (4), se S tem um anulador (que não pode ser anulado) para F, então tem concomitantemente um anulador para qualquer crença que considere ser produzida pelas suas faculdades cognitivas, para qualquer crença que é um resultado dessas faculdades. Ora, como todas as suas crenças são um resultado ou são produzidas pelas suas faculdades cognitivas, S tem um anulador para qualquer crença que ele tenha. Deste modo, S tem um anulador universal para o que quer que seja em que ele acredite, caindo assim num ceticismo esmagador. Mas uma vez que N&E também é uma das suas crenças, então S tem igualmente um anulador para essa crença. Por isso pode concluir-se que um sujeito que aceite N&E, e vir que a $P(F|N\&E)$ é baixa, tem um anulador para N&E. Ou seja, N&E anula-se a si próprio, é auto-derrotante, não podendo ser racionalmente aceite.

Tendo assim em conta que N&E anula-se a si próprio, então há um conflito real entre naturalismo e evolução e, portanto, a sua conjunção não pode ser racionalmente aceite. Uma vez que a evolução é um dos pilares da ciência contemporânea, pode-se afirmar que há um conflito entre naturalismo e ciência. No entanto, como argumenta Plantinga, tal incoerência ou conflito não ocorre caso seja uma conjunção da evolução com o teísmo. Isto porque a teoria científica da evolução é compatível com um

Deus teísta que supervisione e orquestre o curso da evolução, tal como sustenta Sober (2011, p. 187-221), entre outros. É verdade que há alguns, como Dawkins (1986, p. 5), que defendem que o curso da evolução não é supervisionado ou orquestrado por ninguém. Todavia, tais alegações não fazem parte da teoria científica da evolução enquanto tal, mas são apenas acrescentos metafísicos de alguns naturalistas. Além disso, sendo T a proposição de que Deus nos criou por um processo de evolução, a $P(F|T\&E)$ será presumivelmente alta. Pois, de acordo com o teísmo, Deus criou-nos à sua imagem; ou seja, ele criou-nos de tal forma que, em certos aspetos, nos assemelhamos a ele – e um desses aspetos consiste em nos parecermos com Deus no sentido de podermos ter crenças fiáveis e conhecimento. Ora, se tal sucede, então T&E não se anula a si mesmo, ao contrário do que ocorre com N&E. Portanto, o verdadeiro conflito não é entre ciência e religião teísta, mas sim entre ciência e naturalismo. O argumento de Plantinga é válido, mas será sólido?

2. Objeções ao argumento de Plantinga

Como objeção imediata a esta argumentação pode-se alegar, tal como Dennett (2011, p. 35-36/51-52), que os nossos cérebros são *mecanismos sintáticos* que são concebidos pela evolução para localizar a verdade. Deste modo, a evolução por seleção natural, com o pressuposto de que não é orientada, tal como explica por que razão os corações são altamente fiáveis na função de bombear o sangue ou por que razão os olhos são altamente fiáveis a perceber o meio envolvente, também explica por que motivo as crenças que são provocadas por esses olhos ou por outros sentidos são altamente fiáveis na função de obter a verdade. Aliás, é precisamente a competência para rastrear a verdade dos mecanismos de fixação de crenças que explica a sua adaptabilidade tal como a competência para bombear o sangue por parte do coração explica a sua adaptabilidade. Por causa da adaptabilidade os corações são para a circulação do sangue e os cérebros são para ras-

tratar corretamente as condições relevantes do ambiente. E é exatamente isso, de acordo com Dennett, o que a evolução tem feito.

Porém, parece que Dennett está a confundir o problema que está em causa, pois não se está a questionar na premissa (1) como as coisas são ou o que a evolução tem feito, mas sim como as coisas seriam se a conjunção da evolução com naturalismo fosse verdadeira. Ou seja, o que se está a questionar é sobre a $P(F|N\&E)$ e não sobre a $P(F|@)$, onde @ é *o modo como as coisas são atualmente*. Ora, a crítica de Dennett parece incidir sobre esta última probabilidade ao procurar argumentar que a $P(F|@)$ é alta. Mas isso em nada objeta a premissa (1) uma vez que podemos concordar e até parece manifesto que, dado @, a probabilidade das nossas faculdades cognitivas serem fiáveis é alta; todavia, dado N&E, essa probabilidade é baixa. Portanto, a crítica de Dennett parece não acertar no alvo.

No entanto, se deixarmos de lado @ e nos focarmos apenas em N&E, podemos pensar que o conteúdo sobreveniente ou constituído pelas propriedades NF terá de ser tendencialmente verdadeiro para que as criaturas hipotéticas tenham comportamento adaptativo. Pois, suponha-se que o conteúdo sobreveniente ou constituído pelas propriedades NF é falso. Ora, se o conteúdo for falso, provavelmente poderá levar a um comportamento desadaptado uma vez que podemos imaginar que tais criaturas têm, por exemplo, a crença de que os cogumelos *amanitas phalloides* (que são venenosos) são bons para comer. E se têm crenças como essas, então poderá levar a um comportamento incompatível com a sua sobrevivência. Deste modo, parece que é preciso haver crenças verdadeiras (ou uma preponderância de crenças verdadeiras) para que essas criaturas tenham um comportamento adaptativo bem-sucedido. E se isto suceder, então a premissa (1) é falsa.

Uma das tentativas para Plantinga (2012, p. 110-111) responder a esta crítica consiste em recorrer à distinção entre *indicadores* e *crenças*. Para isso imaginemos a seguinte situação: um sapo está sentado numa folha de nenúfar, uma determinada mosca passa por perto, e a língua do sapo ao mover-se rapidamente captura essa

mosca. Nesta situação para o sapo ter um comportamento adaptativo bem-sucedido, i.e. para capturar a mosca, terá de ter certos *indicadores*, como estruturas que registam a aproximação da mosca, a distância entre mosca e sapo, a velocidade da mosca etc. Assim, para o comportamento do sapo ser adaptativo é necessário a existência de indicadores e é necessário que eles indiquem com precisão. Todavia, nada disso precisa de envolver um conteúdo verdadeiro. Ou seja, mesmo que se suponha que um tal sapo pode ter crenças, ele não precisa de acreditar que a mosca está a tal e tal distância e que voa a tal e tal velocidade. Pelo contrário, o que é necessário é haver indicadores precisos. Deste modo, caso esse sapo tenha crenças não importa o que ele acredita, ou se os conteúdos sobrevenientes ou constituídos pelas propriedades NF são verdadeiros ou falsos, para ter um comportamento adaptativo nessa situação.

Nos seres humanos e nas criaturas hipotéticas da experiência mental também existem várias estruturas que indicam ou registam vários estados de coisas no corpo ou no ambiente causando reações apropriadas, isto é, há indicadores. Por exemplo, se uma pessoa tiver uma certa infeção causada por microrganismos, então estruturas do sistema imunitário podem registar a presença desses corpos estranhos que, por sua vez, irão causar uma resposta adequada (como a produção de anticorpos e o aumento da temperatura) para destruir tais microrganismos. Porém, essa pessoa (ou o seu sistema imunitário) não precisa de acreditar que o seu corpo foi invadido por micróbios ou que tem uma infeção. Portanto, temos aqui comportamento adaptativo, em que se resolve uma infeção e uma invasão de micróbios, sendo completamente indiferente se a pessoa possui crenças verdadeiras ou falsas relativamente a isso.

Deste modo, indicadores precisos são uma condição necessária para o comportamento adaptativo. No entanto, como ilustram os exemplos acima, indicadores não são crenças, nem requerem crenças e, além disso, as crenças (tanto verdadeiras como falsas) parecem não ser necessárias para o comportamento adaptativo. Assim,

se há comportamento adaptativo, então há indicadores precisos a funcionar apropriadamente. Mas se há indicadores precisos a funcionar apropriadamente, não importa o que o organismo em questão acredita. Logo, desde que os indicadores sejam precisos, o conteúdo da crença pode ser seja o que for.

Mesmo admitindo que os indicadores são necessários para o comportamento adaptativo, pode insistir-se na objeção à premissa (1) ao defender-se que certos tipos de conteúdos causam (ou parcialmente causam) um determinado comportamento. Por exemplo, suponha-se que sou um dos sujeitos da experiência mental e que as minhas faculdades cognitivas surgiram de acordo com N&E. Imagine-se também que quero beber um refrigerante e a minha esposa diz-me que há uma lata de refrigerante no frigorífico; dessa forma, adquiero uma crença com o conteúdo de que *há uma lata de refrigerante no frigorífico*. Ora, parece que é por causa de acreditar em tal conteúdo que me leva (em parte) ao comportamento de procurar refrigerante no frigorífico em vez de o ir buscar a outro lugar qualquer. Ou seja, pelo menos em parte, é em virtude do conteúdo dessa crença que tenho o comportamento de me dirigir para o frigorífico em vez de me dirigir, por exemplo, para a máquina de lavar roupa. Portanto, tendencialmente é por causa do conteúdo de uma crença C que C causa em parte o comportamento que causa. Se é isso o que sucede, caso um sujeito tenha crenças falsas, como a crença de que os cogumelos *amanitas phalloides* são um alimento saudável, então provavelmente pode causar comportamentos que põem em causa a sua sobrevivência e reprodução. Por outras palavras, parece que crenças falsas tendem a causar comportamentos desadaptados, enquanto crenças verdadeiras tendem a conduzir a comportamentos adaptativos. Por isso, se isto estiver correto, a $P(F|N\&E)$ certamente não será baixa.

Todavia, Plantinga (2012, p. 113-114) tenta argumentar que uma objeção como essa não parece ser forte. Pois, se supusermos que o materialismo é verdadeiro (como se está a proceder ao associar o materialismo ao naturalismo), então as crenças são estruturas neurais com propriedades NF e com um conteúdo proposi-

cional, tal como já vimos. Mas, dado o materialismo, é por causa das propriedades NF, e não em virtude do seu conteúdo, que a crença causa o comportamento que causa. Ou seja, os comportamentos devem-se às propriedades NF e aos seus impulsos neurais, e não a um certo conteúdo. Assim, se o materialismo for verdadeiro, o conteúdo sobreveniente ou constituído é irrelevante para o poder causal da crença com respeito ao comportamento. Aqui a ideia básica de Plantinga é a seguinte:

(#) Se uma crença C tivesse as mesmas propriedades NF mas um conteúdo diferente, então ainda teria o mesmo efeito causal sobre o comportamento.

Ora, isto significa que o conteúdo não é o que causa o comportamento. Assim, se a contrafactual (#) for verdadeira, não sucede que crenças falsas causem comportamento desadaptado em virtude de terem um conteúdo falso, nem que crenças verdadeiras causam comportamento adaptativo pelo facto de terem conteúdos verdadeiros. Desta forma, a verdade ou falsidade da crença é indiferente ou invisível para a seleção natural e para haver um comportamento adaptativo. Com (#) pode-se, então, continuar a defender a premissa (1).

No entanto, considero que esta resposta de Plantinga parece bastante implausível. Em primeiro lugar, a antecedente da contrafactual (#) é metafisicamente impossível dado MNL e dado MR. Já vimos que de acordo com MNL a sobreveniência expressa em (S) é entendida como envolvendo necessidade lógica ampla. Assim, para todos os mundos possíveis, se num certo mundo M há uma estrutura com determinadas propriedades NF_1 das quais sobrevém um conteúdo específico C_1 , então em qualquer outro mundo M^* que tenha uma estrutura com essas propriedades NF_1 daí sobrevém o mesmo conteúdo C_1 . Por outras palavras, não sucede que, ao ter-se em M e em M^* estruturas com as mesmas propriedades NF_1 , as propriedades NF_1 gerem em M o conteúdo C_1 e essas mesmas propriedades NF_1 gerem em M^* um diferente conteúdo C_2 . Deste

modo, não há qualquer mundo possível em que uma estrutura tenha as mesmas propriedades NF_1 mas conteúdos diferentes.

Um resultado similar verifica-se com o MR; pois, se a tese da necessidade da identidade defendida por Kripke (2012) for correta, i.e., a tese de que, para todo x e y , se $x=y$ então necessariamente $x=y$, a identidade entre propriedades NF e propriedades de conteúdo obtém-se em todos os mundos possíveis, supondo que tal identidade ocorre no mundo atual. Deste modo, dado MR, não existe nenhum mundo possível em que uma crença tenha as mesmas propriedades NF mas um conteúdo diferente. É verdade que o próprio Kripke tenta refutar MR ao defender que é metafisicamente possível haver mundos em que tal identidade não se obtém e, por isso, não haveria de toda identidade entre essas propriedades, o que tornaria MR numa teoria implausível. Contudo, como já procurei defender (cf. Faria 2014), essa argumentação de Kripke contra MR não parece ser procedente. De qualquer forma, ao utilizar-se a tese da necessidade da identidade e tendo em conta MR, a antecedente de (#) é metafisicamente impossível. Assim, considerando-se tanto MR como MNL, a contrafactual (#) se for verdadeira, será apenas vacuamente verdadeira, porque a sua antecedente é metafisicamente impossível. Além disso, se a semântica das contrafactuals de Lewis (1973) e Stalnaker (1968) estiver em ordem, qualquer contrafactual com a mesma antecedente de (#) será de igual modo trivial ou vacuamente verdadeira, como é o caso da seguinte:

(*) Se uma crença C tivesse as mesmas propriedades NF mas um conteúdo diferente, então *não teria* o mesmo efeito causal sobre o comportamento.

Com isto pode-se questionar: Que razão há para se preferir a contrafactual (#) em vez desta última contrafactual? Afinal são as duas vacuamente verdadeiras. Claro que Plantinga pode tentar argumentar que a semântica tradicional de Lewis-Stalnaker não é apropriada, talvez porque é intuitivo sustentar que algumas contra-

factuais, com a mesma antecedente impossível, podem ser não-vacuamente verdadeiras e outras não-vacuamente falsas – tal como sugerem Nolan (1997) ou Salerno (2007), entre outros. Por exemplo, enquanto a contrafactual “se 2 tivesse sido maior do que 3, então 3 teria sido menor do que 2” parece não-vacuamente verdadeira, a contrafactual “se 2 tivesse sido maior do que 3, então 3 teria sido maior do que 2” parece não-vacuamente falsa. Mas, mesmo adotando uma semântica alternativa, o problema permanece: por que razão a contrafactual (#) seria não-vacuamente verdadeira enquanto a contrafactual (*) seria não-vacuamente falsa? Plantinga não oferece qualquer razão para se acreditar na verdade não-trivial de (#) em vez de (*), limitando-se a postular que essa contrafactual (#) é verdadeira. Todavia, isso parece ser bastante insuficiente para afastar a objeção de que os conteúdos causam em parte os comportamentos.

Em segundo lugar, Plantinga ao recorrer a (#) está a pressupor que o epifenomenalismo é a teoria mais provável, dado o materialismo; porém, isso parece igualmente implausível. De acordo com o epifenomenalismo apenas as propriedades NF têm poderes causais sobre o comportamento, sendo o conteúdo ou as propriedades mentais causalmente inertes. Ora, se o epifenomenalismo for o mais provável, dado o materialismo (bem como N&E), então a premissa (1) será certamente verdadeira. Mas, supondo que EP abrevia o epifenomenalismo e que M abrevia o materialismo, por que razão considerar que a $P(EP|M)$ é alta? Parece haver mais razões a favor da baixa probabilidade de EP, dado M, do que a favor de uma alta probabilidade; pois, por exemplo, se considerarmos o MR, as propriedades mentais ou de conteúdo são apenas propriedades NF e, por conseguinte, não há qualquer razão para pensar que o conteúdo não entra na cadeia causal que leva ao comportamento. Isto porque se propriedades de conteúdo C são idênticas a propriedades NF e, além disso, se uma crença X causa uma ação A em virtude de ter propriedades NF, então X também causa A em virtude de ter C. Aliás, filósofos como Kim (1996: 56) ou Polger (2004, p. 35) defendem que MR é a única teoria com

uma explicação robusta de como os fenómenos mentais podem causar fenómenos físicos. Assim, dado M (interpretado como MR), a probabilidade de EP será baixa.

Ora, se é verdade que a $P(EP|M)$ é baixa e que, dado M (pelo menos interpretado como MR), as propriedades mentais são causalmente eficazes sobre o comportamento, a objeção anterior (de que crenças falsas tendem a causar comportamentos desadaptados e que crenças verdadeiras tendem a conduzir a comportamentos adaptativos) parece ser restaurada. Claro que Plantinga pode continuar a insistir que aquilo que importa, dado N&E, é apenas que as propriedades NF causem comportamento adaptativo, sendo completamente irrelevante o conteúdo constituído ser verdadeiro ou falso. Todavia, se há causalidade mental, então parece ser de alguma forma relevante que o conteúdo seja tendencialmente verdadeiro em vez de falso, pelas razões apontadas na objeção anterior; pois, ter conteúdos falsos pode levar provavelmente a comportamento desadaptados enquanto caso se tenha conteúdos verdadeiros parece haver uma maior probabilidade para se ter comportamentos adaptativos. Por isso, estou inclinado a supor que a $P(F|N\&E)$ não parece ser baixa, mas talvez tendencialmente elevada, nomeadamente no que concerne às crenças que são mais relevantes para a sobrevivência e reprodução. Outros argumentos substancialmente diferentes para esta mesma conclusão foram avançados por Collin (2013), Tooley (2012), Ramsey (2002), Fodor (2002), Fales (2002), Sober (1998), entre outros, e não parece que Plantinga tenha oferecido uma resposta completamente satisfatória a essas críticas.

Perante tais objeções, considero que o argumento original de Plantinga contra o naturalismo não é procedente porque tem premissas falsas, nomeadamente a premissa (1). No entanto, será que por causa destes problemas o argumento contra o naturalismo está condenado ao fracasso? Penso que não; pois, mesmo que todas estas críticas sejam procedentes, é possível formular uma nova versão do argumento contra o naturalismo, tal como procurei fazer num outro lugar (cf. Faria, 2015), de tal forma que se consegue

contornar essas objeções. Assim, com uma nova proposta de formulação do argumento pode-se continuar a defender a irracionalidade do naturalismo metafísico de uma forma plausível.

Agradecimentos

Agradeço aos meus orientadores Pedro Galvão e Ricardo Santos, bem como aos colegas do LanCog que me ajudaram a discutir as ideias presentes neste artigo. Também agradeço as sugestões propostas pelos revisores da revista *Princípios*. Através de uma Bolsa de Doutorado (SFRH/BD/85051/2012) beneficieei, na realização deste trabalho, do apoio da Fundação para a Ciência e a Tecnologia.

Referências

BEILBY, J. *Naturalism defeated?: essays on Plantinga's evolutionary argument against naturalism*. Ithaca: Cornell University, 2002.

CHURCHLAND, P. Epistemology in the age of neuroscience. *Journal of Philosophy*. v. 84, n. 10, 1987, p. 546-553.

COLLIN, J. Semantic inferentialism and the evolutionary argument against naturalism. *Philosophy Compass*. v. 8, n. 9, 2013, p. 846-856.

DARWIN, C. Letter to William Graham, Down, July 3, 1881. In: DARWIN, F. (Ed.). *The life and letters of Charles Darwin including an autobiographical chapter*. Londres: John Murray, 1881. v. 1. p. 315-316.

DAWKINS, R. *The blind watchmaker: why the evidence of evolution reveals a universe without design*. Nova York: Norton & Company, 1986.

DENNETT, D. Truths that miss their mark: naturalism unscathed / Habits of imagination and their effect on incredulity. In: PLANTINGA, A.; DENNETT, D. (Ed.). *Science and religion: are they compatible?* Nova York: Oxford University Press, 2011. p. 35-36 / 51-52.

FALES, E. Darwin's Doubt, Calvin's Calvary. In: BEILBY, James. (Ed.). *Naturalism defeated?: essays on Plantinga's evolutionary argument against naturalism*. Ithaca: Cornell University, 2002. p. 42-58.

FARIA, D. Será procedente o argumento de Kripke contra a teoria da identidade Tipo-Tipo?. *Revista Portuguesa de Filosofia*. v. 70, n. 1, 2014, p. 112-131.

FARIA, D. Proposta de argumento contra o naturalismo metafísico. *Principia: an international journal of epistemology*. v. 18, n. 3, 2015, p. 361-370.

FODOR, J. Is science biologically possible?. In: BEILBY, James. (Ed.). *Naturalism defeated?: essays on Plantinga's evolutionary argument against naturalism*. Ithaca: Cornell University, 2002. p. 30-42.

KIM, J. *Philosophy of mind*. Boulder: Westview Press, 1996.

KRIPKE, S. *O nomear e a necessidade*. Trad. Ricardo Santos e Teresa Filipe. Lisboa: Gradiva, 2012.

LEWIS, D. Counterfactuals and comparative possibility. *Journal of Philosophical Logic*. v. 2, 1973, p. 418-446.

NAGEL, T. *The view from nowhere*. Nova York: Oxford University Press, 1989.

NOLAN, D. Impossible worlds: a modest approach. *Notre Dame Journal for Formal Logic*, v. 38, 1997, p. 325-527.

PLANTINGA, A. An evolutionary argument against naturalism. *Logos*. v. 12, 1991, p. 27-48.

PLANTINGA, A. Is naturalism irrational? In: PLANTINGA, A. *Warrant and Proper Function*. Nova York: Oxford University Press, 1993. p. 216-237.

PLANTINGA, A. Naturalism and lack of knowledge. In: PLANTINGA, A. *Warranted Christian Belief*. Nova York: Oxford University Press, 2000. p. 227-240.

PLANTINGA, A. Against naturalism. In: PLANTINGA, A.; TOOLEY, M. *Knowledge of God*. Oxford: Wiley-Blackwell, 2008. p. 1-69.

PLANTINGA, A. The evolutionary argument against naturalism. In: PLANTINGA, A. *Where the conflict really lies: science, religion and naturalism*. Nova York: Oxford University Press, 2011. p. 307-350.

PLANTINGA, A. The evolutionary argument against naturalism. In: STUMP, J.; PADGETT, A. (Ed.). *The Blackwell companion to science and Christianity*. Oxford: Blackwell, 2012. p. 103-115.

POLGER, T. *Natural minds*. Cambridge: MIT, 2004.

RAMSEY, W. Naturalism defended. In: BEILBY, James. (Ed.). *Naturalism defeated?: essays on Plantinga's evolutionary argument against naturalism*. Ithaca: Cornell University, 2002. p. 15-29.

SALERNO, J. Why counterpossibles are not trivial. *The Reasoner*. v. 1, n. 1, 2007, p. 1-7.

SOBER, E. Plantinga's probability arguments against evolutionary naturalism. *Pacific Philosophical Quarterly*. v. 79, n. 2, 1998, p. 115-129.

SOBER, E. Evolution without naturalism. In: KVANVIG, J. (Ed.). *Oxford studies in Philosophy of Religion*, Volume 3. Nova York: Oxford University Press, 2011. p. 187-221.

STALNAKER, R. A theory of conditionals. In: *Studies in logical theory, American Philosophical Quarterly Monograph Series*. v. 2, 1968, p. 98-112.

TOOLEY, M. Plantinga's new argument against materialism. *Philosophia Christi*. v. 14, n. 1, 2012, p. 29-48.

Artigo recebido em 17/09/2015, aprovado em 15/12/2015