

## PROPOSTA DE ARGUMENTO CONTRA O NATURALISMO METAFÍSICO

DOMINGOS FARIA

---

**Abstract.** In this paper I present a proposal to reformulate the argument of Alvin Plantinga (2011) against metaphysical naturalism. Contrary to Plantinga's argument, in this new version I propose to consider the probability of the reliability of cognitive faculties, not with regard to any kind of beliefs, but only with respect to metaphysical beliefs. I claim that those who accept naturalism have a defeater for the belief that their cognitive faculties are reliable with respect to metaphysical beliefs and, thus, they have a defeater for any of their metaphysical beliefs, including the belief in metaphysical naturalism. Therefore, those who accept naturalism have a defeater for naturalism; in other words, metaphysical naturalism is self-defeating.

**Keywords:** Plantinga; naturalism; theism; self-defeater; metaphysics.

---

O naturalismo metafísico é a perspectiva de que o mundo descrito pelas ciências naturais é tudo o que existe e 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 (2011) concebeu um argumento que procura mostrar que esta perspectiva naturalista, em conjunção com a teoria da evolução, é inconsistente e autoderrotante; assim, é irracional aceitar o naturalismo. Considero que esta conclusão é verdadeira, mas penso que o argumento de Plantinga tem premissas falsas (cf. Faria 2015); por isso, para defender a conclusão em questão proponho uma nova versão do argumento contra o naturalismo.

Na minha proposta de reformulação do argumento contra o naturalismo, estou a considerar a probabilidade da fiabilidade das faculdades cognitivas, não em relação a todo o tipo de crenças (tal como faz Plantinga), mas apenas em relação a crenças metafísicas. Por crenças metafísicas entendo as crenças sobre a natureza última da realidade, como sustentar que existem entidades universais ou, pelo contrário, que existem apenas particulares; como defender o niilismo mereológico ou, pelo contrário, o universalismo mereológico; como acreditar no naturalismo metafísico ou no seu oposto, entre outros.

Para formular o argumento recorro às seguintes abreviaturas: 'N' abrevia a proposição que *o naturalismo metafísico é verdadeiro*, 'E' abrevia a proposição que *nós e as nossas faculdades cognitivas surgiram por meio de processos assinalados pela teoria evolucionista contemporânea*, 'FM' abrevia a proposição de que *as nossas faculdades cognitivas são fiáveis com respeito a crenças metafísicas* (ou seja, que produzem

*Principia* 18(3): 361–370 (2014).

Published by NEL — Epistemology and Logic Research Group, Federal University of Santa Catarina (UFSC), Brazil.

uma preponderância considerável de crenças metafísicas verdadeiras em comparação com as falsas), e ' $P(\dots| \_)$ ' é uma abreviatura para '*a probabilidade condicional de ... dado \_*'. Tendo em conta essas abreviaturas, o argumento que proponho é o seguinte:

- (1)  $P(FM|N \& E)$  é baixa ou insondável.
- (2) Quem aceitar  $N \& E$  e também vir que (1) é verdadeira, tem um anulador para  $FM$ .
- (3) Este anulador não pode ser anulado com crenças metafísicas.
- (4) Quem tem um anulador para  $FM$ , tem um anulador para qualquer uma das suas crenças metafísicas, incluindo  $N$ .
- (5) Logo,  $N$  anula-se a si mesmo e não pode ser racionalmente aceite.

Analisando a premissa (1), por que razão afirmar que a probabilidade condicional de  $FM$ , dado  $N \& E$ , é baixa ou insondável? Para começar a argumentar a favor da baixa probabilidade de  $FM$ , dado  $N \& E$ , baseio-me na seguinte ideia fundamental: ao contrário das crenças perceptivas ou de memória, etc, possuir crenças metafísicas ou, mais concretamente, sustentar crenças metafísicas falsas em vez de verdadeiras parece ser de forma geral muito provavelmente irrelevante ou indiferente para a sobrevivência ou reprodução, i.e., para se ter um comportamento adaptativo. Assim, suponha-se que ' $CF$ ' abrevia a proposição de que *sustentarem-se crenças metafísicas falsas em vez de verdadeiras em geral não faz diferença significativa para haver comportamento adaptativo*. Então, qual é a  $P(CF|N \& E)$ ? Aqui temos duas opções: ou considerar que as crenças metafísicas não têm em geral qualquer resultado comportamental (opção A) ou considerar que as crenças metafísicas têm em geral algum resultado comportamental (opção B).

Por um lado, seguindo a opção A, pode parecer implausível sustentar que de forma geral as crenças metafísicas estão diretamente ligadas a comportamentos específicos (quer sejam adaptativos ou não). Por exemplo, que comportamento um sujeito passará a ter caso comece a postular universais ou caso comece a acreditar no nominalismo de semelhanças, no tetradimensionalismo, no realismo modal, etc? A ideia aqui presente é a de que é possível que de certas propriedades neurofisiológicas  $NF$  sobrevenham ou se constituam conteúdos metafísicos (verdadeiros ou falsos), mas essas mesmas propriedades  $NF$  podem não gerar qualquer impulso neural para os músculos e para outras partes relevantes do organismo, não havendo assim comportamento. Por isso, se a evolução seleciona comportamentos adaptativos e se as crenças metafísicas em geral não tendem a causar qualquer comportamento (nem adaptativo nem não-adaptativo), então a verdade ou falsidade das crenças metafísicas parece ser irrelevante ou invisível à seleção natural. Assim, de acordo com a opção A, a  $P(CF|N \& E)$  é alta.

Por outro lado, segundo a opção *B*, considerando-se que de forma geral as crenças metafísicas estão diretamente relacionadas com comportamentos, e até mesmo que se admita a possibilidade das crenças metafísicas serem de alguma forma adaptativas (ou talvez um subproduto de algum traço ou comportamento adaptativo), ainda assim parece bastante plausível alegar que globalmente crenças metafísicas falsas têm os mesmos benefícios adaptativos que crenças metafísicas verdadeiras. Isto porque, dado *N & E*, e supondo que existem tais ligações entre crenças metafísicas e comportamentos, parece razoável aceitar que comportamentos que resultam de crenças metafísicas falsas geralmente não são suficientemente diferentes a nível adaptativo daqueles que resultam de crenças metafísicas verdadeiras para fazer diferença na seleção natural.

Mas por que razão pensar que em geral, dado *N & E*, as crenças metafísicas falsas são, adaptativamente, tão boas como as verdadeiras? Uma possível via de argumentação consiste em notar que, do ponto de vista naturalista, praticamente todos os nossos antepassados que sobreviveram, e por isso que possuíam um comportamento adaptativo, tiveram faculdades que produziam geralmente crenças metafísicas falsas. Isto porque as crenças metafísicas da maioria dos nossos antepassados nas sociedades antigas (que presumivelmente tinham consequências comportamentais) faziam referência a seres ou eventos sobrenaturais que dado *N* são crenças falsas. Assim, dado *N & E*, a probabilidade de *CF* aplicado aos nossos antepassados é alta. No entanto, as nossas faculdades não são significativamente diferentes comparativamente com as dos nossos antepassados. Por isso, dado *N & E*, a probabilidade de *CF* aplicada a nós também parece alta. Portanto, em geral não existe um benefício significativo, do ponto de vista naturalista e da sua conjunção com a seleção natural, em sustentar crenças metafísicas verdadeiras em vez de falsas. Ora, se este raciocínio está correto, então de forma geral sujeitos que têm crenças metafísicas falsas sobrevivem tão bem como se eles tivessem as correspondentes crenças metafísicas verdadeiras e, assim, ter crenças metafísicas falsas em vez de verdadeiras não faz em geral diferença para se ter um comportamento adaptativo, o que significa que, de acordo com a opção *B*, a  $P(CF|N \& E)$  é igualmente alta.

Assim, de uma forma ou de outra a  $P(CF|N \& E)$  parece ser alta. Ora, se isso estiver correto, então *FM* é improvável dado a conjunção de *N & E & CF*; pois, se *CF* é verdadeira, então é bastante razoável aceitar que a seleção natural tratará de igual forma as crenças metafísicas verdadeiras e as crenças metafísicas falsas e, por isso, a seleção natural não dará preferência a uma faculdade que produza uma preponderância considerável de crenças metafísicas verdadeiras a uma que produza crenças metafísicas falsas. Ou seja, temos razões fortes para duvidar que a seleção natural, dado *N & E*, selecionaria uma faculdade produtora de crenças metafísicas fiáveis, o que justifica a baixa probabilidade de *FM*, dado *N & E*, assinalada na premissa (1). Parece então que, dado *N & E*, não há qualquer razão para que a seleção natural nos

tenha dotado com uma fiável faculdade cognitiva que produza crenças metafísicas. Para justificar isso pode-se ainda recorrer ao seguinte teorema da probabilidade total:

$$(6) P(FM|N \& E) = [P(FM|N \& E \& CF) \times P(CF|N \& E)] + [P(FM|N \& E \& \sim CF) \times P(\sim CF|N \& E)]$$

Ora, o teorema (6) diz-nos que a probabilidade de *FM* sobre *N & E* é a média ponderada da probabilidade de *FM* sobre *N & E & CF* e *N & E & ~CF*, por sua vez ponderada pelas probabilidades de *CF* e *~CF* sobre *N & E*. Com a aplicação de (6) podemos constatar que se a  $P(CF|N \& E)$  for alta, como tentamos argumentar em cima, e se a  $P(FM|N \& E \& CF)$  for baixa, então a  $P(FM|N \& E)$  será igualmente baixa. Suponha-se, por exemplo, que a  $P(CF|N \& E)$  é 0.8, assim nesse caso a  $P(\sim CF|N \& E)$  é 0.2, e estipule-se que a  $P(FM|N \& E \& CF)$  é 0.3, então mesmo no pior dos cenários em que a  $P(FM|N \& E \& \sim CF)$  é 1, a  $P(FM|N \& E)$  é 0.44, i.e., inferior a  $\frac{1}{2}$ . Portanto, mesmo nos piores cenários a probabilidade em questão é baixa. Todavia, pode-se julgar que atribuir valores precisos para essas probabilidades é uma tarefa absurda ou pode-se alegar que devido às nossas limitações cognitivas não conseguimos determinar minimamente o valor dessas probabilidades. Por isso, talvez o mais razoável, dada a nossa atual situação epistémica, seja argumentar que a  $P(FM|N \& E)$  é insondável. Mas quer se aceite que a probabilidade em questão é baixa, quer se aceite que é insondável, fica-se comprometido com a primeira premissa.

O próximo passo do argumento contra o naturalismo consiste em notar que se uma pessoa reconhecer que a  $P(FM|N \& E)$  é baixa ou insondável, e também aceitar *N & E*, então ela tem um anulador para *FM*, isto é, uma razão para rejeitar *FM*. Seguindo Plantinga na clarificação do que é ter um anulador, pode-se dizer que um anulador para uma crença *C* que eu tenha é outra crença *C\** que passo a ter e que é tal que, dado que tenho *C\**, não posso continuar racionalmente a ter *C*. Por exemplo, olho para um campo e vejo o que me parece ser uma ovelha — assim, formo a crença *C* de que estou a ver uma ovelha. Todavia, suponha-se que o leitor identifica-se adequadamente como o proprietário desse campo e informa-me que não há qualquer ovelha no campo. Além disso, diz-me que o que eu vejo de facto é afinal um cão que é indistinguível de uma ovelha a esta distância. Por causa disso forma a crença *C\** de que não há ovelhas no campo e que afinal o que estou a ver é um cão. Ora, esta minha nova crença *C\** é um anulador da minha antiga crença *C*; ou seja, ao adquirir *C\** não posso continuar a sustentar *C*, tendo assim uma boa razão para a abandonar.

Do mesmo modo, um sujeito *S* que aceite *N & E* pode inicialmente ter a crença *C* de que *FM*. Mas por alguma razão, como aquelas apresentadas para sustentar a premissa (1), *S* pode também reconhecer e começar a aceitar a crença *C\** de que a  $P(FM|N \& E)$  é baixa ou insondável. Ora, *S* ao aceitar *C\** adquire uma boa razão para desconfiar seriamente ou para abandonar *C*, não a podendo continuar a sustentar racionalmente. Por isso, se *S* aceitar *N & E* e vir que a  $P(FM|N \& E)$  é baixa

ou insondável, então  $S$  tem um anulador para  $FM$ ; ou seja,  $S$  não pode continuar a acreditar racionalmente que as suas faculdades cognitivas são fiáveis com respeito a crenças metafísicas — devendo ser agnóstico quanto a  $FM$  ou aceitar a sua negação.

Para clarificar e fundamentar ainda melhor esta premissa (2) vale a pena utilizar algumas analogias. Para isso começemos apenas pelo caso em que a probabilidade em questão na premissa (1) é baixa. Recorrendo a uma analogia, suponha-se que há uma droga, conhecida por  $XX$ , que destrói fiabilidade das faculdades cognitivas com respeito às crenças metafísicas e que 95% daqueles que a tomam perdem tal fiabilidade cognitiva ao fim de duas horas após a sua ingestão. Agora imagine-se que  $S$  acredita que tomou essa droga há mais de duas horas, bem como acredita que a  $P(FM|S$  tomou  $XX$  há mais duas horas) é baixa (pelo facto de conhecer essa droga e as suas consequências). Ora, se  $S$  sustenta essas duas crenças conjuntamente, então tais crenças dão-lhe um anulador para a sua crença inicial de que as suas faculdades cognitivas são fiáveis com respeito às suas crenças metafísicas. Por outras palavras, se  $S$  tem a crença de que tomou  $XX$  há mais de duas horas e também tem a crença de que a  $P(FM|S$  tomou  $XX$  há mais duas horas) é baixa (uma vez que tal droga destrói a fiabilidade cognitiva com respeito às crenças metafísicas numa grande percentagem de casos), então tais crenças constituem uma forte razão para  $S$  deixar de acreditar em  $FM$ . Algo semelhante sucede com outros casos, por exemplo, caso  $S$  acredite que é vítima do génio maligno cartesiano e que a  $P(FM|$  ser vítima do génio maligno) é baixa; ou caso  $S$  acredite que vive numa *matrix* e que a  $P(FM|$  viver numa *matrix*) é baixa. Nestes exemplos  $S$  tem um anulador para  $FM$ . De igual modo, no caso em questão no argumento contra o naturalismo, se  $S$  acreditar que  $N \& E$  e que a  $P(FM|N \& E)$  é baixa, então  $S$  tem um anulador para  $FM$ .

Mas se a probabilidade em questão na premissa (1) for insondável, consegue-se chegar à mesma conclusão; pois, quem aceitar  $N \& E$  e também vir que a  $P(FM|N \& E)$  é insondável, tem igualmente um anulador para  $FM$ . Para se tentar mostrar isso pode-se recorrer igualmente a várias analogias. Por exemplo, imagine-se que somos incapazes de determinar a fiabilidade de um instrumento científico; ora, se acreditamos que é impossível sabermos quão fiável esse instrumento é, então é razoável que não formemos crenças com base nesse instrumento uma vez que é bem possível que tal instrumento nos engane. Ou, voltando à analogia da droga  $XX$ , que abordamos em cima, pode-se supor que quem tomar uma determinada quantidade de  $XX$  fica com a fiabilidade das suas faculdades cognitivas com respeito às crenças metafísicas completamente destruída; mas agora imagine-se que ninguém sabe que quantidade específica é essa. Assim, a  $P(FM|S$  tomou  $XX)$  é insondável, podendo tanto ser alta como baixa. Ora, perante estas circunstâncias, se  $S$  acredita que tomou  $XX$  e que a  $P(FM|S$  tomou  $XX)$  é insondável, então  $S$  parece ter razões para em princípio suspeitar de  $FM$ , i.e., tem um anulador de  $FM$ . O mesmo parece suceder para quem acreditar em  $N \& E$  e vir que a  $P(FM|N \& E)$  é insondável. Assim, quer a probabi-

lidade em questão seja baixa, quer seja insondável, continua a haver um anulador para *FM*.

Além disso, de acordo com a premissa (3), esse é um anulador que não pode ser ele mesmo anulado recorrendo a crenças metafísicas. Pois, um *anulador de anulador*, i.e., um anulador para esse anulador que recorra a crenças metafísicas teria, por exemplo, de ter a forma de um argumento a favor de *FM*. Todavia, (i) um tal argumento seria epistemicamente circular e, do mesmo modo, (ii) quem tem um anulador para *FM* também tem um anulador para qualquer argumento a favor de *FM* que contenha alguma premissa metafísica. Assim, por um lado, suponha-se que concebo um argumento a favor de *FM* que contenha pelo menos uma premissa metafísica, de tal modo que passo a acreditar em *FM* com base nesse argumento. Ora, para ficar convencido de *FM* com base nesse argumento devo acreditar que essas premissas são verdadeiras. No entanto, ao proceder desse modo, tenho de começar por confiar que as minhas faculdades cognitivas são fiáveis na produção de crenças sobre as suas premissas com conteúdo metafísico. Ou seja, ao argumentar a favor de *FM* nessa base já estou a pressupor que *FM* é verdadeira. Deste modo, qualquer argumento a favor de *FM* que recorra a premissas metafísicas será viciosamente circular, uma vez que a confiança num tal argumento pressupõe que a conclusão do argumento é verdadeira.

Mas, por outro lado, quem aceitar *N & E* e, assim, tiver um anulador para *FM* terá igualmente um anulador para as premissas metafísicas de um argumento a favor de *FM*. Ou seja, ver que a  $P(FM|N \& E)$  é baixa ou insondável dá a *S*, caso aceite *N & E*, uma razão para duvidar de *FM*. Porém, se tal sucede, então isso também leva a que *S* não tenha confiança naquilo que é produzido pelas suas faculdades cognitivas com respeito a crenças metafísicas. Ora, uma vez que um determinado argumento que contenha alguma premissa metafísica a favor de *FM* é produzido ou está de alguma forma dependente da confiança nas suas faculdades cognitivas que produzem crenças metafísicas, e como a fiabilidade destas faculdades está posta em causa, então *S* tem uma razão para suspeitar de tal argumento.

A única possibilidade do naturalista anular o anulador de *FM* parece ser construir um argumento que recorra apenas a premissas empíricas ou *a posteriori* e que tenha como conclusão *FM*. Mas haverá um tal argumento? Do ponto de vista naturalista parece que não há tal argumento, havendo pelo contrário vários indícios a favor de *CF* e contra *FM*. Por exemplo, de um ponto de vista naturalista, tal como já notamos em cima, pode-se apelar a considerações históricas de que os nossos antepassados tinham maioritariamente crenças metafísicas falsas mas comportamentos adaptativos, o que torna improvável que os nossos antepassados tenham *FM*. Mas como nós temos faculdades cognitivas semelhantes, então também temos uma razão para desconfiar que temos *FM*. Do mesmo modo, ao longo de toda a história até aos nossos dias a maior parte das pessoas sustentaram e ainda sustentam crenças meta-

físicas religiosas, que para o naturalista são crenças falsas. Porém, várias explicações naturalistas das crenças religiosas, como as de David Wilson (2002), tentam mostrar que tais crenças são adaptativas uma vez que são crenças que promovem a solidariedade e cooperação entre os elementos de um determinado grupo. Deste modo, dado o naturalismo, as pessoas têm em geral crenças metafísicas falsas mas adaptativas, o que constitui uma razão para o naturalista desconfiar de FM. De qualquer modo deixo o desafio para o naturalista tentar formular um argumento, só com premissas empíricas ou *a posteriori*, que anule o presente anulador de FM e que mostre que CF é falsa ou improvável.

O último passo do argumento diz-nos que se S tem um anulador para FM, então tem concomitantemente um anulador para qualquer crença metafísica que seja produzida por tais faculdades, para qualquer crença que é um resultado dessas faculdades que produzem crenças metafísicas. Ora, como todas as suas crenças metafísicas são um resultado ou são produzidas por tais faculdades, S tem um anulador para qualquer crença metafísica que ele tenha. Mas uma vez que N também é uma das suas crenças metafísicas, então S tem igualmente um anulador para essa crença. Por isso pode concluir-se que um sujeito que aceite  $N \& E$ , e vir que  $P(FM|N \& E)$  é baixa ou insondável, tem um anulador para N. Ou seja, N anula-se a si próprio, não podendo ser racionalmente aceite. Em suma, se virmos que (1) é verdadeira e aceitamos  $N \& E$ , então adquirimos um anulador para FM e, por conseguinte, obtemos uma razão para suspeitar de todas as crenças que sejam produzidas por tais faculdades. Uma vez que o naturalismo é produzido por essas faculdades, então também adquirimos um anulador para o naturalismo. Portanto, ao acreditar-se no naturalismo adquire-se uma boa razão para se deixar de acreditar no naturalismo; i.e., o naturalismo é autoderrotante.

Ao contrário do argumento original de Plantinga, nesta reformulação quem aceita  $N \& E$  não adquire um anulador para todas as suas crenças de qualquer tipo, nem cai num ceticismo global e esmagador. Pelo contrário, nesta minha proposta de argumento contra o naturalismo apresenta-se uma conclusão mais fraca, nomeadamente que quem aceita  $N \& E$  tem um anulador apenas para as suas crenças metafísicas; por isso,  $N \& E$  conduz apenas a um ceticismo local. Ora, isso permite anular não só o naturalismo como também os argumentos ateológicos apresentados pelos naturalistas metafísicos. Por exemplo, frequentemente os naturalistas utilizam o argumento indiciário do mal ou o argumento da ocultação divina como tentativas de mostrar a inexistência de Deus. Mas se o naturalista tem um anulador para qualquer uma das suas crenças metafísicas, então tem igualmente um anulador para esses argumentos ateológicos uma vez que tais argumentos têm pelo menos uma premissa que expressa uma crença metafísica. Portanto, não seria racional o naturalista recorrer a tais argumentos.

Globalmente penso que esta proposta de reformulação (ou outra formulação si-

milar) do argumento contra o naturalismo metafísico é mais plausível e procedente do que a versão original de Plantinga. Sobretudo porque a premissa central (1) do argumento reformulado parece ser mais intuitiva, mais fácil de aceitar e não tão controversa como a premissa central de Plantinga. No entanto, admito que a presente reformulação do argumento pode ser suscetível a várias críticas desafiantes e há muito trabalho a fazer no futuro para se tentar responder com alguma plausibilidade a todas essas possíveis críticas.

Por exemplo, uma forma bastante desafiante de se criticar o argumento aqui proposto consiste em tentar defender que as principais alternativas ao naturalismo, como é o caso do teísmo, sofrem dos mesmos problemas que se apontou ao naturalismo. Ou seja, se o argumento proposto for sólido, então o teísmo depara-se com o mesmo tipo de autoderrotação que se aplica ao naturalismo. Esta crítica, conhecida como ‘objeção *Tu Quoque*’, já foi apontada ao argumento original de Plantinga (cf. Mirza 2011) e penso que pode ser de igual forma adaptada ao argumento aqui reformulado. Assim, questiona-se: qual é a probabilidade de *FM* assumindo que Deus existe e que a história cristã é verdadeira? Ou seja, abreviando ‘*T*’ o teísmo e a história cristã, qual é a  $P(FM|T)$ ? Ora, se Deus existe e se a história cristã é verdadeira, então também é provável que Deus tenha criado pessoas humanas e não-humanas com livre-arbítrio. Além disso, se tal história for verdadeira, algumas pessoas, como Satanás, rebelaram-se livremente contra Deus e a partir de então causaram grande parte dos males que existem no mundo. Por exemplo, dado *T*, é possível que Satanás cause nos seres humanos uma disfunção massiva nas faculdades cognitivas que produzem crenças metafísicas. Deste modo, seja ‘*S*’ a abreviatura da proposição de que *Satanás faz com que as faculdades cognitivas dos seres humanos que produzem crenças metafísicas não funcionem apropriadamente*. Ora, dada a seguinte fórmula de probabilidade total:

$$(7) P(FM|T) = [P(FM|T \& S) \times P(S|T)] + [P(FM|T \& \sim S) \times P(\sim S|T)]$$

Se a história cristã é verdadeira, então parece que a  $P(S|T)$  será presumivelmente alta e a  $P(FM|T \& S)$  será baixa e, assim, a  $P(FM|T)$  será baixa. Ou caso se advogue que é absurdo atribuir probabilidades precisas a tais casos ou que a nossa situação epistémica não nos permite saber quão provável isso pode ser, pode-se defender que a probabilidade em questão é insondável. Mas, de qualquer forma, se a probabilidade em questão é baixa ou insondável, então o teísta tem um anulador para *FM* e para qualquer uma das suas crenças metafísicas, de forma análoga ao que sucede nas premissas (2) e (4). Ora, como a crença no teísmo é ela própria uma crença metafísica, então o teísta tem um anulador para o teísmo. Portanto, o teísmo anula-se a si mesmo e não pode ser racionalmente aceite.

Será esta objeção procedente? Para escapar a esse problema pode-se conceber, do ponto de vista teísta, pelo menos duas possíveis estratégias argumentativas: por um

lado, se o teísmo e a história cristã forem verdadeiros, pode-se dizer que 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. Além disso, também faz parte da história cristã a ideia de que mesmo numa situação em que os seres humanos se afastam da imagem de Deus (p.e. por causa das suas ações livres ou das ações doutros seres, como Satanás), Deus pode restaurar essa imagem nos seres humanos com a colaboração destes (i.e. restaurar o ser humano do pecado ou restaurar as suas faculdades). Assim, sendo ‘*R*’ a abreviatura de que *Deus restaura a ‘imagem de Deus’ no ser humano*, então mesmo que a  $P(FM|T \& S \& \sim R)$  seja baixa, a  $P(FM|T \& S \& R)$  presumivelmente não será baixa. Ora, se tal for o caso, então não há a partir daí um anulador para *FM* e, por conseguinte, não há um anulador para o teísmo.

Por outro lado, mesmo concedendo que a  $P(FM|T \& S)$  pode ser baixa ou insondável, inclusive com a conjunção de *R*, pode-se tentar argumentar que a crença no teísmo é diferente das restantes crenças metafísicas, no sentido que pode ser produzida por outras faculdades em vez de ser (completamente) produzida por *FM*. Por exemplo, pode-se defender que, dado *T*, a crença teísta é baseada nas experiências religiosas e que, por isso, não tem a mesma origem que as restantes crenças metafísicas. Ora, se Deus existe, então poderia garantir que as crenças baseadas nas experiências religiosas são fiáveis. Ou, mais concretamente, pode-se defender que é provável existir um *sensus divinitatis*, uma faculdade que em determinadas circunstâncias e funcionando apropriadamente produz crenças teístas fiáveis, caso Deus exista e caso pretenda que os seres humanos adquiram conhecimento de Deus (cf. Plantinga 2000, 188-190). Deste modo, se alguma destas linhas argumentativas for procedente, consegue-se salvaguardar a crença teísta da autoderrotação ao mesmo tempo que se defende que o naturalismo é autoderrotante.

## Agradecimentos

Agradeço os comentários e críticas de vários colegas e amigos, nomeadamente de Pedro Galvão, Ricardo Santos, Adriana Silva Graça, José Manuel Mestre, Diogo Santos, Pedro Dinis, Raimundo Henriques, Elton Marques, Fernando Furtado, João Carlos Silva, João Pinheiro, Vítor Guerreiro, entre outros, que me ajudaram a repensar algumas ideias neste artigo. Discuti versões preliminares deste artigo no 1º Congresso Português de Filosofia, bem como no Seminário LanCog. Através de uma Bolsa de Doutoramento (SFRH/BD/85051/2012) beneficiei, na realização deste trabalho, do apoio da Fundação para a Ciência e a Tecnologia.

## Referências

Faria, D. 2015. A Irracionalidade do Naturalismo Metafísico: um exame crítico ao argumento de Plantinga. *Revista Portuguesa de Filosofia*. Forthcoming.

*Principia* 18(3): 361–370 (2014).

- Mirza, O. 2011. The Evolutionary Argument Against Naturalism. *Philosophy Compass* 6: 78–89.
- Plantinga, A. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press.
- . 2011. *Where the Conflict Really Lies: Science, Religion, and Naturalism*. Oxford: Oxford University Press.
- Wilson, D. 2002. *Darwin's Cathedral: Evolution, Religion and the Nature of Society*. University of Chicago Press.

DOMINGOS FARIA  
LanCog Research Group CFUL  
Universidade de Lisboa  
df@domingosfaria.net

**Resumo.** Neste artigo proponho uma reformulação do argumento de Alvin Plantinga (2011) contra o naturalismo metafísico. Ao contrário do argumento de Plantinga, nesta nova versão proponho considerar a probabilidade da fiabilidade das faculdades cognitivas, não em relação a todo o tipo de crenças, mas apenas em relação a crenças metafísicas. Argumento que quem aceita o naturalismo tem um anulador para a crença de que as suas faculdades cognitivas são fiáveis com respeito a crenças metafísicas e, por conseguinte, tem um anulador para qualquer uma das suas crenças metafísicas, inclusive para a sua crença no naturalismo metafísico. Assim, quem aceita o naturalismo tem um anulador para o naturalismo; ou seja, o naturalismo metafísico é autoderrotante.

**Palavras-chave:** Plantinga; naturalismo; teísmo; autoderrotação; metafísica.