

**UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO DE FILOSOFIA E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA  
MESTRADO EM FILOSOFIA**

**REALISMO, NATURALISMO E SEMÂNTICA MORAL**

**SILVIO KAVETSKI**

**FLORIANÓPOLIS  
2017**



Silvio Kavetski

## **REALISMO, NATURALISMO E SEMÂNTICA MORAL**

Dissertação de Mestrado submetida ao Programa de Pós-Graduação em Filosofia da Universidade Federal de Santa Catarina como requisito para a obtenção do Título de Mestre em Filosofia.

Orientador: Prof. Dr. Darlei Dall'Agnol

Florianópolis  
2017

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Kavetski, Silvio  
Realismo, naturalismo e semântica moral / Silvio  
Kavetski ; orientador, Darlei Dall'Agnol - Florianópolis,  
SC, 2017.  
164 p.

Dissertação (mestrado) - Universidade Federal de Santa  
Catarina, Centro de Filosofia e Ciências Humanas. Programa  
de Pós-Graduação em Filosofia.

Inclui referências

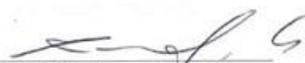
1. Filosofia. 2. Metaética. 3. Realismo. 4.  
Naturalismo. 5. Semântica Moral. I. Dall'Agnol, Darlei. II.  
Universidade Federal de Santa Catarina. Programa de Pós  
Graduação em Filosofia. III. Título.

Silvio Kavetski

**“REALISMO, NATURALISMO E SEMÂNTICA MORAL”**

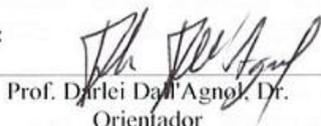
Esta dissertação foi julgada adequada para obtenção do Título de “Mestre em Filosofia”, e aprovada em sua forma final pelo Programa de Pós-Graduação em Filosofia.

Florianópolis, 24 de fevereiro de 2017.



Prof. Roberto Wu, Dr.  
Coordenador do Curso

**Banca Examinadora:**

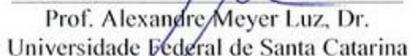


Prof. Darlei Dall'Agnol, Dr.  
Orientador

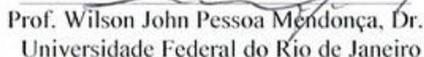
Universidade Federal de Santa Catarina



Prof.ª Milene Consenso Tonetto, Dr.ª  
Universidade Federal de Santa Catarina



Prof. Alexandre Meyer Luz, Dr.  
Universidade Federal de Santa Catarina



Prof. Wilson John Pessoa Mendonça, Dr.  
Universidade Federal do Rio de Janeiro



## AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer à minha família, Pedro, Amélia e Porfírio, pelo incentivo que sempre me deram para que eu pudesse estudar. Obrigado por acreditarem no meu trabalho. A vocês minha eterna admiração, gratidão, amor e respeito.

Agradeço em especial ao meu orientador, Professor Darlei Dall’Agnol. Obrigado pela atenção, paciência, dedicação, amizade e, principalmente, pelos ensinamentos que ajudaram a aprimorar tanto minhas competências filosóficas quanto morais.

À Professora Milene Consenso Tonetto (UFSC) e ao Professor Alexandre Meyer Luz (UFSC) por terem participado da banca de qualificação e de defesa. Agradeço pelas inúmeras sugestões e críticas para que eu pudesse melhorar o presente trabalho.

Ao Professor Wilson Mendonça (UFRJ) pela minuciosa leitura deste trabalho e por ter feito primorosas críticas na arguição de defesa. Obrigado por ajudar a melhorar minhas posições filosóficas e a compreender melhor o naturalismo moral.

Não poderia deixar de agradecer aos Professores Marciano Spica (UNICENTRO) e Ernesto Giusti (UNICENTRO), pela excelente educação filosófica durante o período de graduação e pelo incentivo constante a desenvolver pesquisas de pós-graduação. Se hoje tenho alguma virtude filosófica, certamente devo-as a vocês.

Aos amigos, especialmente Bruno Aislã, Luiz Helvécio, Kherian Gracher, Delvair Moreira e André Luiz, pelas discussões e momentos agradáveis que tornaram menos árdua a realização deste trabalho.

Agradeço, também, ao CNPq, pela bolsa de estudos.



## RESUMO

O naturalismo moral é a teoria metaética que sustenta que fatos e propriedades morais são fatos e propriedades naturais. Desde que G. E. Moore apresentou o seu argumento da questão aberta, tem havido várias críticas a essa teoria, o que fez com que os filósofos articulassem várias teorias metaéticas alternativas ao naturalismo, tais como intuicionismo, emotivismo, prescritivismo e a teoria do erro. Mas a partir da década de oitenta David Brink, Richard Boyd e Nicholas Sturgeon desenvolveram uma nova versão do naturalismo moral – o naturalismo não reducionista – que, argumentativamente, evita essas objeções e apresenta inúmeras outras vantagens. O resultado foi uma reascensão do realismo moral naturalista. No entanto, dois filósofos formularam uma objeção ao naturalismo não reducionista, que ficou conhecida como ‘Argumento da Terra Gêmea Moral’, que tem gerado bastante discussão. O objetivo deste trabalho é reconstruir as linhas principais desse debate mostrando que: o naturalismo não reducionista realmente tem boas respostas a algumas críticas frequentes, tais como o argumento da questão aberta de Moore, a reformulação de Hare deste argumento, à objeção construtivista, à crítica de relativismo, ao argumento do desacordo moral etc; e que, mesmo que o argumento da terra gêmea moral seja o seu principal problema, o que parece ser o caso, há algumas estratégias de respostas possíveis a favor do naturalista.

**Palavras-chave:** Realismo Moral; Naturalismo; Argumento da Terra Gêmea Moral; Semântica Moral.



## ABSTRACT

Moral naturalism is the metaethical theory that maintains that moral facts and properties are natural facts and properties. Since G. E. Moore presented his open question argument there have been several critiques to this theory, which made the philosophers articulate several alternative metaethical theories to naturalism as intuitionism, emotivism, prescriptivism and the error theory. But from the eighties David Brink, Richard Boyd and Nicholas Sturgeon developed a new version of moral naturalism – the non reductionist naturalism – that, arguably, avoids these objections and it have numerous another advantages. The result was a resurrection of naturalistic moral realism. However, two philosophers formulated an objection to non reductionist naturalism that became known as ‘Moral Twin Earth Argument’ that has generated much discussion. The objective of this work is to reconstruct the main lines of this debate showing that: the non reductionist naturalism really have good replies to some frequent critiques such as Moore’s open question argument, Hare’s reformulation of this argument, the constructivist’s objection, the critique of relativism, to the moral disagreement’s argument etc; and that, even if the moral twin earth argument be its main problem, what looks to be the case, there are some possible response strategies in favor of the naturalist.

**Keywords:** Moral Realism; Naturalism; Moral Twin Earth Argument; Moral Semantics.



## SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>13</b>
<b>I – NATURALISMO MORAL E SEUS CRÍTICOS .....</b>	<b>17</b>
INTRODUÇÃO .....	17
1. Realismo moral e naturalismo .....	19
2. Moore e o argumento da questão aberta .....	31
3. R. M. Hare e o prescritivismo .....	37
4. Mackie e a teoria do erro .....	63
OBSERVAÇÕES FINAIS .....	75
<b>II – NATURALISMO NÃO REDUCIONISTA.....</b>	<b>77</b>
INTRODUÇÃO .....	77
1. Os pressupostos .....	79
2. Semântica moral .....	96
2.1. A teoria de Boyd.....	96
2.2. Aplicação às propriedades morais .....	104
3. Explicações morais .....	112
3.1. “O problema com a ética” .....	113
3.2. As propriedades morais são explicativas .....	123
OBSERVAÇÕES FINAIS .....	132
<b>III – NATURALISMO E SEMÂNTICA MORAL.....</b>	<b>135</b>
INTRODUÇÃO .....	135
1. O argumento da Terra Gêmea Moral .....	136
1.1. O “novo” argumento da questão aberta .....	144
2. Algumas objeções .....	147
OBSERVAÇÕES FINAIS .....	152
<b>CONCLUSÃO .....</b>	<b>155</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>159</b>



## INTRODUÇÃO

Quando as pessoas falam sobre ética, é mais comum que tenham em mente questões sobre que tipo de ações são boas, más, corretas, estimáveis ou reprováveis. Elas podem dizer, por exemplo, que todas as ações boas devem ter algo em comum que as faz serem boas e, a partir disso, tentar encontrar o elemento comum que compartilham. Podem dizer que se trata de haver algo em suas consequências e que isso é importante para determinarmos se são certas ou erradas. Ou, podem ignorar as consequências e afirmar que se deve dar atenção a algumas regras mais gerais sobre que coisas são certas ou erradas. Ainda, podem dizer que as ações boas ou más tem algo a ver com o caráter de quem as executa. E assim por diante. O que elas estão fazendo é tentar teorizar sobre o modo como devemos viver. E fazer isso é fazer *ética normativa*. Além disso, elas podem ter em mente questões mais específicas. Podem discutir sobre se uma guerra é justa e sobre quais condições deve satisfazer, sobre se é justificado utilizarmos os animais como fonte de alimentação ou como objetos de vários tipos de experimentos, sobre se é permissível alguém escolher se quer viver ou morrer, sobre se é correto ou não interromper uma gravidez nos primeiros meses de gestação etc. Ao fazerem isso, elas estarão fazendo *ética aplicada*. Os filósofos costumam dizer que esses assuntos – sobre ética normativa e ética aplicada – são questões de *primeira ordem*. São questões que estão mais estritamente associadas ao modo como devemos guiar nossa conduta. Mas há questões de *segunda ordem* também. Quando as pessoas começam a discutir sobre assuntos de natureza mais geral, tais como se os valores morais são relativos a uma ou outra comunidade específica ou se são universais, se há verdades morais independentes do que as pessoas pensam ou se o *certo* e o *errado* dependem das pessoas, se o *bom* é expressão das atitudes morais ou se é algo externo e objetivo etc., diz-se que estão tratando de questões de ordem superior. Questões de segunda ordem constituem aquela parte da ética denominada *metaética*. ‘Qual é a função semântica do discurso moral?’, ‘Fatos e propriedades morais existem?’, ‘Há conhecimento moral?’, ‘Os juízos morais podem ser objetivamente corretos ou incorretos?’, ‘As propriedades morais são conceitualmente motivadoras ou não?’. A metaética é sobre coisas desse tipo.

O presente trabalho é sobre metaética. Mais especificamente, é sobre o problema de saber se as propriedades morais são propriedades naturais ou não. Considerar este problema é se perguntar sobre a

verdade do *naturalismo moral*. Esta teoria metaética se compromete com a tese de que fatos e propriedades morais são fatos e propriedades naturais. Assim, não há uma espécie de “reinos dos valores” em que fatos e propriedades como *bom*, *correto*, *dever* e *justiça* subsistem e, em algum sentido, se distinguem de tudo aquilo que podemos identificar como sendo fatural. Para o naturalista moral, fatos e propriedades morais são objetivos e fazem parte do mundo natural, de modo que não são, em algum sentido, dependentes daquilo que as pessoas pensam sobre o certo e o errado.

O naturalismo moral teve seus altos e baixos. Quando Moore, munido de um argumento – o argumento da questão aberta – acusou grande parte das teorias éticas de cometerem uma suposta falácia, o naturalismo moral sofreu grande descrença. Os eticistas tenderam ou a endossar um realismo não naturalista, como o intuicionismo, ou a desenvolver algum tipo de não cognitivismo, como o emotivismo e o prescritivismo. Nesse cenário, várias outras críticas ao naturalismo moral surgiram. Mas, a partir das décadas de setenta e oitenta, vários avanços na filosofia da linguagem, metafísica, filosofia da mente e epistemologia possibilitaram uma reascensão do naturalismo moral. Desenvolveu-se um novo naturalismo – o naturalismo não reducionista, como ficou conhecido. Este naturalismo, aparentemente, não é vulnerável ao argumento de Moore ou às críticas posteriores das outras tendências metaéticas. Por isso, o naturalismo moral voltou a atrair uma grande quantidade de adeptos.

No entanto, parece que mesmo esse novo naturalismo não está livre de problemas. Uma tese característica do naturalismo se refere à semântica dos termos morais. Dado que o defensor dessa corrente metaética se compromete com a ideia de que os fatos e propriedades morais são naturais, então ele sustenta que a referência dos termos morais é similar à referência dos termos não morais. Isso tem sido um problema desde Moore e o naturalismo não reducionista, com sua nova teoria semântica, supostamente evitaria objeções que reivindicassem alguma disparidade semântica entre os termos morais e os termos não morais. Mas graças a dois filósofos americanos, esse tipo de problema – sob uma nova roupagem, considerando avanços em várias áreas da filosofia – voltou a ganhar notoriedade. Em uma série de artigos, Terence Horgan e Mark Timmons (doravante H&T) apresentaram um argumento que ficou conhecido como o ‘Argumento da Terra Gêmea Moral’. Este argumento tem a finalidade de mostrar que o *novo* naturalismo é falso, e H&T acreditam que faz isso com sucesso.

Deixarei para falar dos detalhes desse argumento no momento adequado. Por hora, umas poucas palavras bastam. O ponto de H&T consiste em argumentar que, para falantes competentes de uma dada língua, parece que os termos morais não possuem as propriedades semânticas que o naturalismo moral assume possuírem. Os dois autores se utilizam de um experimento mental baseado no cenário construído por Hilary Putnam, o famoso experimento Terra *vs.* Terra Gêmea. Mas, o que está em questão não é a referência de ‘água’, mas de ‘bom’. H&T argumentam que se imaginarmos que a referência de ‘bom’ na Terra é *x* e na Terra Gêmea é *y*, e considerarmos um possível desacordo entre terráqueos e terráqueos gêmeos sobre algum assunto moral, a nossa intuição será de que não há um mero desacordo verbal, mas *moral* entre os falantes. Isso significa que parece que a referência de ‘bom’ não depende somente do mundo natural, como o naturalista argumenta, mas de *algo a mais* (que poderia ser uma atribuição dos próprios agentes). Por isso, o naturalismo moral seria falso, pois supostamente não conseguiria explicar a referência das propriedades morais. Parece que H&T propuseram um problema difícil para o naturalista e, desde a publicação dos seus artigos, no início da década de noventa, tem havido uma significativa discussão desse problema. Assim, sustentar o naturalismo hoje implica na necessidade de se ter uma boa resposta para esse problema. Ou, se perguntar sobre a verdade do naturalismo moral significa, em algum momento, ter de pensar sobre o argumento proposto por esses dois filósofos.

O objetivo deste trabalho é reconstruir as linhas principais desse debate mostrando que: o naturalismo não reducionista realmente tem boas respostas a algumas críticas frequentes, tais como o argumento da questão aberta de Moore, a reformulação de Hare deste argumento, à objeção construtivista, à crítica de relativismo, ao argumento do desacordo moral etc; e que, mesmo que o argumento da terra gêmea moral seja o seu principal problema, há algumas estratégias de respostas possíveis a favor do naturalista. Dado que a discussão desse argumento é uma parte um pouco específica do assunto geral do naturalismo moral e pressupõe que se saiba uma série de outras coisas, não deverei ir direto ao ponto. Vou dedicar o primeiro capítulo para fazer algumas coisas bem introdutórias. Primeiro, irei apresentar brevemente o que é o naturalismo moral. Depois, vou dedicar um tempo para falar de algumas críticas frequentemente feitas a essa teoria, tais como o argumento da questão aberta, a teoria do erro e as várias objeções apresentadas por R. M. Hare ao desenvolver o seu prescritivismo universal.

No segundo capítulo quero me concentrar somente no naturalismo não reducionista com o intuito de mostrar suas principais características. Num primeiro momento, falarei de alguns avanços na filosofia em geral que possibilitaram o seu desenvolvimento e tentarei dizer em que sentido algumas das objeções usuais ao naturalismo não se aplicam ao naturalismo não reducionista (por isso da importância de apresentar essas objeções no primeiro capítulo). Depois, irei apresentar a teoria semântica de R. Boyd. A razão para fazer isso é que Boyd foi um dos principais filósofos a desenvolver a semântica do naturalismo não reducionista e, além disso, o argumento de H&T dirige-se especificamente a essa pressuposição semântica; então, é bom ter claro o alvo do argumento. Para finalizar, vou apresentar outra estratégia característica do naturalismo não reducionista para defender que propriedades morais são naturais, a ideia de que propriedades morais desempenham um papel explicativo. Tendo apresentado as noções gerais do naturalismo, algumas críticas e a estratégia do naturalismo não reducionista, poderei focar na objeção de H&T.

Assim, no terceiro capítulo, finalmente, vou apresentar o argumento de H&T e apontar alguns de seus problemas que parecem favorecer o naturalismo moral. Como o leitor perceberá, não irei esboçar uma posição definitiva a respeito desse problema. Irei me limitar a apresentar o argumento da terra gêmea moral e o suposto novo argumento da questão aberta que deriva do experimento mental de H&T, assim como algumas estratégias de resposta. Com essas três partes, é possível ter uma boa noção do debate sobre o naturalismo moral em geral e do naturalismo não reducionista em específico, além de um aprofundamento maior em um de seus principais problemas que vem sendo discutido contemporaneamente.

## I – NATURALISMO MORAL E SEUS CRÍTICOS

“Cada coisa pode ser chamada boa por bondade divina como princípio primeiro, exemplar, efetivo e final de toda bondade”.  
(Tomás de Aquino, *Sobre a Bondade de Deus*).

### INTRODUÇÃO

Se o naturalismo moral é verdadeiro, então a afirmação de Tomás de Aquino deve de ser falsa. O naturalismo moral nega que haja alguma relação entre a moralidade e qualquer tipo de entidade supernatural. ‘*Naturalismo*’ é um termo normalmente aplicado a várias posições em diferentes áreas da filosofia. E, como é de se esperar de uma doutrina filosófica, sua própria definição é um assunto de disputa. Mas, ainda assim, é possível ter um consenso razoável sobre o que é o naturalismo, pelo menos num nível mais genérico e sem entrar em muitos detalhes. A discussão aqui será sobre o naturalismo *moral*, especificamente. Em metaética, o naturalismo moral pode ser definido como a tese de que fatos e propriedades morais são fatos e propriedades naturais; ou, como a ideia de que as propriedades morais não representam nenhum tipo de propriedades supernaturais ou não naturais<sup>1</sup>. É uma tese controversa, obviamente. Há quem defenda que fatos e propriedades morais são não naturais chamando atenção para as dissimilaridades semânticas, ontológicas, metafísicas e lógicas entre propriedades morais e não morais. Há quem diga que essas diferenças são apenas aparentes e que podemos dar sentido à naturalização da moral se compreendermos o que o *ser natural* realmente significa. Mas é difícil tomar essas teses e definições como sendo informativas em algum sentido a menos que se

---

<sup>1</sup> Gostaria de deixar claro que estou usando ‘fato’ e ‘propriedade’ de forma neutra com o objetivo de facilitar a discussão. Certamente, ‘fato’ e ‘propriedade’ são coisas diferentes. Uma propriedade é um atributo, uma característica ou um aspecto de determinado objeto, por exemplo. Ter quatro patas é uma propriedade dos cavalos, mas o ‘ter quatro patas’ é apenas uma atribuição linguística a um objeto, e não um fato. Um fato é o modo pelo qual um objeto é ou o acontecimento de um evento. ‘Sócrates beber cicuta’ é um fato. ‘Fato’ e ‘propriedade’ são categoricamente distintos. Estarei seguindo Terence Horgan e Mark Timmons ao usar ambos os termos de forma intercambiável para facilitar a discussão. Cf. HORGAN, T. & TIMMONS, M. *Troubles on Moral Twin Earth: Moral Queerness Revived*. *Synthese*, 92, 1992a.

diga o que significa um fato ou propriedade moral o ser natural, supernatural e como são formuladas e defendidas essas ideias. Essas são coisas que deverei fazer no decorrer desse trabalho.

Neste primeiro capítulo, gostaria de fornecer uma caracterização breve do naturalismo moral e apresentar algumas das principais críticas a essa teoria. Vou dividir a tarefa em quatro partes. Na primeira, quero explicar o que é o naturalismo moral e quais seriam as supostas vantagens dessa posição. Defender o naturalismo é, normalmente, defender o realismo moral, embora defender o realismo moral não implique em defender o naturalismo. Então tentarei explicar o que é o realismo moral e em que sentido o naturalismo constitui razões a favor do realismo. Farei uma série de distinções que, suponho, ajudarão a entender melhor o assunto em questão. Começarei pelas teses mais gerais, o naturalismo metafísico e o naturalismo metodológico, com a intenção de mostrar em que sentido o naturalismo moral é e não é extensão dessas teses. Terei que dar exemplos de algumas teorias éticas, obviamente. O principal problema não é dizer o que é o naturalismo, mas dizer em que sentido as propriedades morais são naturais. Isso requer uma caracterização do que significa alegar que algo é natural. Este é um assunto que também terei que abordar nessa primeira seção. Irei falar sobre algumas interpretações, supor que há uma que é a mais usual e, por fim, apresentar o critério de David Copp para determinar se uma dada propriedade é natural ou não. O critério certamente é discutível, mas o objetivo de apresentá-lo aqui como o padrão é apenas para tentar fornecer um primeiro direcionamento. Falarei também sobre algumas vantagens que os naturalistas frequentemente reivindicam à sua posição. Por fim, farei uma distinção entre dois tipos de naturalismo, aquilo que se costuma chamar reducionista e naturalismo não reducionista. Isso nos dará uma noção razoável do que é o naturalismo antes que comecemos a discutir as teorias de forma mais específica.

A partir da segunda seção, vou começar a apresentar algumas críticas ao naturalismo moral. Naturalmente, começarei pela clássica, o argumento da questão aberta de Moore. Moore sustentou que propriedades morais não são naturais com base no seu argumento da questão aberta. É importante falar sobre sua posição, pois ela influenciou grandemente o desenvolvimento posterior da metaética.

Na terceira seção, darei atenção a uma teoria específica, o prescritivismo universal de R. M. Hare. Hare apresentou uma reformulação do argumento de Moore e avançou uma série de outras críticas ao naturalismo. Aqui falarei sobre o emotivismo moral e como Hare reage a essa teoria apresentando o prescritivismo como alternativa

não cognitivista. Irei apresentar várias objeções de Hare ao naturalismo, tais como a “Regra de Hare”, que é a ideia de que não se pode derivar conclusões sobre *dever ser* a partir de premissas sobre o *ser* somente, sua reformulação do argumento de Moore, a tese da prescritividade e a ideia de que o naturalismo moral implica em relativismo.

Na quarta e última seção, irei apresentar outra objeção bastante conhecida ao naturalismo moral, a teoria do erro de J. Mackie. Essa teoria sugere que os fatos e propriedades morais não são de tipo natural, como o naturalista defende. Na verdade, sugere que não são de tipo algum. Para o teórico do erro, fatos e propriedades morais são uma espécie de ficção. Dividirei em duas partes a explicação dessas ideias, assim como faz Mackie. A primeira é a parte conceitual. Aqui, explico o que significa dizer que propriedades morais são objetivas e categoricamente prescritivas. E a segunda é a parte ontológica. Aqui, apresento os argumentos de Mackie para manter que fatos morais não existem e que as sentenças morais implicam numa espécie de erro. Todo esse percurso nos dará uma ideia do que é o naturalismo moral e de algumas de suas principais críticas.

## 1. Realismo moral e naturalismo

Poucos de nós diriam que coisas como mesas, cadeiras, automóveis, livros e montanhas não são reais. Elas simplesmente estão aí, existem, dizemos. Mas pode parecer estranho pensar que coisas como *bondade*, *incorreção* e *justiça* existam ou sejam reais. Afinal, não podemos vê-las, senti-las, cheirá-las ou tocá-las. Mas, isso é simplificar a questão. Não vemos ou tocamos os números, os elétrons, o tempo e o amor e, no entanto, não dizemos que essas coisas não existem ou não são reais. Obviamente, isso nos leva a perguntar sobre a dimensão da realidade, ou a que tipo de condições algo tem que satisfazer para ser considerado real. E isso pode ser um pouco complicado. Alguns podem dizer que algo é real se é passível de investigação pelas ciências naturais. Outros podem dizer que algo é real se é parte da melhor explicação do mundo. Outros, ainda, que algo é real se tem a propriedade de causar alguma coisa no mundo. Com isso, começamos a perceber que a pergunta sobre a realidade das propriedades morais não é trivial.

Há uma boa quantidade de filósofos que têm tentado mostrar que os fatos e propriedades morais são reais e que existem independentemente de pensarmos sobre elas. Eles são frequentemente conhecidos como realistas morais. Um realista moral sustenta que os

enunciados morais descrevem fatos e que são verdadeiros ou falsos na medida em que descrevem esses fatos corretamente ou incorretamente. ‘O assassinato é incorreto’ é tão similar quanto ‘O gato tem quatro patas’, pois supostamente é uma descrição da realidade. Há uma ação, o assassinato, que tem a propriedade da *incorreção*, e o ter esta propriedade é independente daquilo que pensamos sobre o assassinato, se é errado ou não. Podemos tornar a definição do realismo moral mais precisa considerando o que David Brink diz.

*Realismo Moral:* (1) Há fatos ou verdades morais, e (2) esses fatos ou verdades são independentes das evidências para eles<sup>2</sup>.

Como ele mesmo aponta, os oponentes da primeira tese são os não cognitivistas e os niilistas. Os primeiros, grosso modo, sustentam que os enunciados morais, em primeiro lugar, não predicam a instanciação de propriedades e, portanto, não são verdadeiros ou falsos e que, em segundo, expressam algum tipo de estado não cognitivo dos sujeitos. Os emotivistas dizem que expressam as emoções ou os sentimentos de aprovação ou desaprovação. Os subjetivistas que expressam as opiniões morais de uma pessoa, grupo ou comunidade. Os expressivistas que expressam o endosso de um sistema de normas. Para um niilista moral, por outro lado, nosso discurso moral pode ser organizado de tal forma a dar a impressão de que propriedades morais são reais, mas, no fundo, fatos morais não existem.

O realista, portanto, não é um não cognitivista ou um niilista. Ele mantém que fatos morais existem. Como ele faz isso? Antes de responder essa pergunta, falemos sobre a segunda tese que constitui a definição do realismo moral, a tese da independência, por assim dizer. Russ Shafer-Landau tem uma boa caracterização da tese da independência. Ele diz:

Realistas acreditam que há verdades morais que são obtidas independentemente de qualquer perspectiva preferida, no sentido de que *os padrões morais que fixam os fatos morais não são verdadeiros em virtude de sua confirmação a*

---

<sup>2</sup> BRINK, D. O. *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press, 1989, p. 17.

*partir de nenhuma perspectiva atual ou hipotética*<sup>3</sup>.

Quer dizer, aquilo que *faz* com que determinada ação ou pessoa tenha tal propriedade moral não é a nossa avaliação ou o nosso código moral, mas o próprio estado de coisas. Ao fazermos juízos morais estamos tentando fornecer uma descrição precisa da realidade, que é externa e independente.

Várias teorias contrastam com a tese da independência. Consideremos três delas. Um relativista moral, como Gilbert Harman, por exemplo, mantém que “algo ser moralmente certo ou errado, bom ou mal, justo ou injusto, etc., é uma questão relativa – relativa a uma ou outra moralidade ou padrão moral de referência”<sup>4</sup>. Poderíamos dizer que, de acordo com o relativismo, algo não *é* certo ou errado, mas é *tornado* certo ou errado em dependência a um código moral específico. E, dado que os códigos morais são formados a partir daquilo que os agentes consideram certo ou errado, para o relativismo, não há fatos morais independentes de nossas mentes. Um construtivista, como John Rawls, sustenta que algo ser correto ou não depende da aceitação de pessoas livres e razoáveis num estado de equilíbrio reflexivo em que há uma coerência máxima de crenças<sup>5</sup>. Assim, os fatos morais não estão dados, mas são construídos à medida que há acordo refletido dos indivíduos capazes de discutir razoavelmente. E um prescritivista, como Richard Hare, defende que juízos morais são prescrições universalizáveis. Isso significa que algo ser bom depende da aceitação universal deste algo – querendo isso dizer que todos aceitariam como bom independente da posição que ocupam – e que, se há assentimento a este algo, então o indivíduo que assentiu não pode não se comprometer com o conteúdo da avaliação moral. O ponto é que são os sujeitos que determinam o que é prescritivamente universalizável. Então, as questões sobre o certo e o errado não são independentes de nossas crenças e opiniões morais. O realista, portanto, não endossa nenhuma dessas três teorias metaéticas. Podemos dizer que essas teorias são antirrealistas, assim como as teorias não cognitivistas e niilista citadas acima.

---

<sup>3</sup> SHAFER-LANDAU, R. *Moral Realism: A Defense*. Clarendon Press: Oxford, 2003, p. 15. Itálico do próprio autor.

<sup>4</sup> HARMAN, G. *Moral Relativism Explained*. 2012, p. 1. (Artigo não publicado).

<sup>5</sup> Cf. RAWLS, J. *A Theory of Justice* (Revised Edition). Cambridge, Massachusetts: Harvard University Press, 1971/1999, p. 40-46.

Na verdade, como sugeriu Fischer<sup>6</sup>, há vários motivos, *a partir do senso comum*, para se endossar o realismo moral. Quando dizemos que algo é certo ou errado, normalmente não estamos meramente querendo expressar a nossa preferência. Se preferimos café forte e sem açúcar e alguém nos diz que o *melhor* é café fraco e com açúcar, provavelmente não iniciaremos uma discussão para tentar provar que o melhor é realmente café forte e sem açúcar. Podemos até citar algumas vantagens, mas normalmente dizemos que é uma questão de preferência. No entanto, se somos moralmente comprometidos com o juízo moral de que a tortura é errada e alguém nos diz que acha tal prática apazível, as coisas mudam. Certamente não diremos que se trata apenas de mera preferência. Comumente achamos que a tortura, de algum modo, realmente tem a propriedade de ser errada. E assim com qualquer posição moral. Outra razão para o realismo é a convergência das opiniões morais. As pessoas discordam sobre muitas coisas, mas há grande acordo sobre vários juízos morais, como o de que o assassinato é errado, o de que o roubo é errado, que é bom promover a igualdade, e assim por diante. Pode-se dizer que a melhor explicação para isso é de que há propriedades morais. Outra razão é a verdade moral. Comumente supomos que a verdade sobre algum assunto independe das opiniões ou crenças dos sujeitos, que é algo externo a isso. Que a porta está quebrada independe do meu estado de espírito. No mesmo sentido, é intuitivo que a violência sexual é errada independe das crenças e almejos das pessoas. Ainda, a ideia de que há desacordo moral também pode contar como uma razão a favor do realismo (ao invés de a desfavor, como é usualmente sustentado). Quando alguém diz que o aborto é correto e outro alguém diz que é errado, nós consideramos intuitivamente que o desacordo entre as duas partes é genuíno, que alguém tem de estar errado. Isso é mais bem explicado pela ideia de que há fatos morais, pois, caso contrário, não faria o menor sentido uma parte tentar convencer a outra de seu ponto. Essas são algumas razões a favor do realismo a partir do senso comum.

Mas, esses são apenas argumentos intuitivos, quer dizer, são razões que parecem verdadeiras sem a necessidade de teorizações mais sofisticadas. Não *provam* nada a favor do realismo moral. Poder-se-ia começar a fazer isso se se pudesse mostrar que, realmente, há fatos morais. Por isso, voltemos à primeira tese do realismo. Como vimos, é a tese de que há fatos morais. Isso não quer dizer nada se não vier acompanhado de uma resposta à pergunta de como o realista faz isso.

---

<sup>6</sup> FISHER, A. *Metaethics: An Introduction*. Acumen, 2011, p. 56-60.

Basicamente, há duas alternativas. A primeira consiste em dizer que há fatos e propriedades morais, que nós temos acesso a esses fatos e propriedades, mas que eles (as) não são coisas como pedras, mesas e árvores, quer dizer, coisas que podemos observar e tocar. Estão para além dos nossos cinco sentidos. São um tipo *sui generis* de instâncias de propriedades. O nosso acesso a elas se dá através de uma faculdade de conhecimento denominada de intuição, neste caso de intuição *moral*. Apreendemos que a propriedade da incorreção se aplica ao assassinato diretamente, sem a necessidade de alguma investigação ou raciocínio indutivo ou dedutivo. Nós simplesmente temos a habilidade de reconhecer as instâncias de propriedades morais. Assim, não se requer nenhum argumento para que reconheçamos as bases daquilo a partir do que a moralidade é constituída. Uma das características centrais deste tipo de abordagem é o pluralismo de opiniões morais, pois os indivíduos podem ter intuições diferentes sobre assuntos mais controversos. Essa proposta é denominada de não naturalismo ou *intuicionismo moral*.

A segunda alternativa consiste em sustentar que os fatos e propriedades morais são naturais. A esta proposta denomina-se *naturalismo moral*. Podemos definir o naturalismo moral do seguinte modo:

*Naturalismo Moral:* É a tese de que há fatos e propriedades morais e que estes fatos e propriedades morais são naturais.

Essa é apenas uma definição geral, mas, para início de conversa, é bastante útil. Um naturalista moral é alguém que acredita que uma descrição completa da moralidade e suas exigências nos pode ser dada com base inteiramente em uma investigação natural da realidade. É claro que para começarmos a entender o que isso significa é preciso dizer um pouco mais sobre o que é uma investigação naturalista da realidade. Para isso, precisamos notar que o naturalismo moral é, *em grande medida*, extensão de uma tese metafísica mais geral, o *naturalismo metafísico*. O naturalismo metafísico é uma visão ontológica que, como apontam M. De Caro e D. Macarthur, se compromete com a seguinte tese:

*Naturalismo Metafísico:* O mundo consiste em nada mais do que entidades que explicações científicas bem sucedidas nos comprometem<sup>7</sup>.

Quando se fala em ‘explicações científicas’ o que se tem em mente são as ciências naturais. Neste sentido, essas explicações científicas não nos comprometem com a existência de entidades supersensíveis como Deus, alma ou espíritos. Explicações científicas tratam de entidades *naturais*<sup>8</sup>. Assim, o naturalismo metafísico é a ideia de que tudo o que existe faz parte do mundo natural. Horgan e Timmons tomam a tese do naturalismo metafísico de forma restrita. Eles não falam em ‘explicações científicas’, mas em ‘assunto da física’<sup>9</sup>. Desse modo, o naturalismo metafísico seria a ideia de que tudo o que existe é parte do mundo físico natural. As entidades que existem no mundo são assunto de que a física se ocupa.

O naturalismo metafísico é distinto do naturalismo metodológico. Embora não seja uma condição necessária, grande parte dos filósofos que assumem o naturalismo metafísico assume, também, o naturalismo metodológico. Este último pode ser definido assim:

*Naturalismo Metodológico:* A investigação científica é, em princípio, a nossa única fonte genuína de conhecimento ou entendimento. Todas as outras alegadas formas de conhecimento (e.g., o conhecimento *a priori*) ou entendimento são ou ilegítimas ou, em princípio, redutíveis ao conhecimento ou entendimento científico<sup>10</sup>.

Se tomarmos apenas a tese mais geral do naturalismo, o naturalismo metafísico, para caracterizarmos o naturalismo moral, então teorias emotivistas, subjetivistas, prescritivistas, construtivistas, relativistas e expressivistas são teorias naturalistas. Isso porque o naturalismo metafísico nega que exista um reino de entidades não naturais ou supersensíveis e essas teorias não estão pressupondo que a existência de

---

<sup>7</sup> DE CARO, M & MACARTHUR, D. Science, Naturalism, and the Problem of Normativity. In: DE CARO, M & MACARTHUR, D. *Naturalism and Normativity* (eds). Columbia University Press: New Yourk, 2010, p. 1-19, p. 4.

<sup>8</sup> O problema de se dizer isso é definir o que é ‘natural’. Tratarei disso adiante.

<sup>9</sup> HORGAN, T. & TIMMONS, M. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, Volume 16, 1991, p. 448.

<sup>10</sup> DE CARO, M & MACARTHUR, D. 2010, *Op. Cit.*, p. 4.

fatos e propriedades morais depende de um reino supernatural. Mas isso não é o caso. É um erro dizer que essas teorias são naturalistas. O naturalista moral está reivindicando que fatos e propriedades morais são naturais num sentido diferente dessas visões antirrealistas. Ele quer dizer que existem fatos e propriedades morais objetivos e que esses fatos morais são naturais. Neste sentido, eles são independentes do que as pessoas pensam. O fato de a água ferver a tal temperatura é um fato natural, independente de pensarmos que a água ferve a uma temperatura diferente ou que nem mesmo ferve. O mesmo acontece com a moralidade. Que certos atos tenham certas propriedades morais é independente de pensarmos que tenham tais propriedades. É neste sentido que a segunda tese do realismo moral deve ser entendida.

Não é difícil dizer o que é o naturalismo: é a ideia de que fatos e propriedades morais são naturais. O que é difícil é mostrar o que significa dizer que fatos morais são naturais, ou, o que é ‘natural’. É comum dizer que a moralidade pertence ao domínio daquilo que é *normativo*. A normatividade tem a ver com o que devemos fazer e com a avaliação de certas coisas, como em ‘Deve-se promover a igualdade’ ou ‘Ajudar a reduzir a pobreza mundial é moralmente bom’. Se se assumir o naturalismo moral, então se deve mostrar como essas coisas que pertencem à normatividade podem ser “acomodadas” em uma visão natural de mundo, pois se tudo o que existe são entidades naturais, então os fatos e propriedades morais também devem ser entidades naturais. Em geral, há vários modos de se argumentar o que ‘propriedades naturais’ significam. David Copp fez uma espécie de taxonomia dessas estratégias. Ele diz que há quatro abordagens principais na literatura: o modelo reducionista, o modelo de fornecer definições ostensivas, o modelo de definições metafísicas e o modelo de definições epistêmicas<sup>11</sup>.

Sem entrar em muitos detalhes, pois iremos discutir de forma mais aprofundada no decorrer do trabalho, consideremos o que, segundo Copp, significa cada uma dessas estratégias. A estratégia reducionista consiste em identificar uma base de propriedades que supostamente são indiscutivelmente naturais e depois exigir, para qualquer propriedade que é qualificada como moral, que seja adequadamente relacionada com a base natural de propriedades identificada. Essa base natural tem de poder ser expressa através de predicados que refiram propriedades naturais. Então, qualquer suposta propriedade natural tem de poder ser

---

<sup>11</sup> Cf. COPP, D. Why Naturalism? *Ethical Theory and Moral Practice*, Vol. 6, 2003, p. 182-185.

relacionada com a base natural e poder ser expressa através de uma terminologia que refira a propriedades naturais<sup>12</sup>. O modelo das definições ostensivas consiste em tomar ‘objetos naturais’ como sendo tudo aquilo que nos rodeia de alguma forma, tais como mesas, árvores e montanhas, e tentar definir as propriedades morais como sendo objetos deste tipo, objetos que podemos apontar e dizer ‘Olhe. Aquilo é uma propriedade natural!’<sup>13</sup>. O modelo metafísico possui pelo menos quatro variantes. O primeiro é dizer que propriedades naturais são descritivas ou propriedades fatuais. O segundo é dizer que propriedades naturais são propriedades causais. Esta é a estratégia de tentar explicar o mundo natural dizendo que representa uma ordem causal. O terceiro é dizer que o mundo natural é um múltiplo espaço-temporal. Assim, as propriedades naturais são tudo aquilo que está localizado no espaço e no tempo. E a quarta consiste em dizer que o mundo natural é o mundo materialista ou fisicalista. Assim, propriedades naturais são físicas ou materiais<sup>14</sup>. O modelo das definições epistêmicas é a ideia de que o mundo natural é tudo aquilo que pode ser estudado pelas ciências naturais. Propriedades naturais são aquelas que as ciências naturais estudam<sup>15</sup>. Todas essas estratégias tem os seus problemas específicos, mas isso é algo que não iremos nos concentrar agora.

Como temos visto, parece que a estratégia de definições epistemológicas é a mais usual. Em geral, os filósofos normalmente concordam que ‘natural’ se refere, em algum sentido, àquilo que é estudado pelas ciências naturais. Considere as seguintes passagens.

Por ‘natureza’, portanto, eu significo e tenho significado aquilo que é o assunto das ciências naturais e também da psicologia. Pode ser dito que isso inclui tudo o que existiu, existe ou irá existir no tempo<sup>16</sup>.

O que eu acho plausível (mesmo que não uma verdade conceitual) é que fatos morais não podem

---

<sup>12</sup> Cf. COPP, 2003, *Op. Cit.*, p. 182.

<sup>13</sup> Cf. *Ibidem*, p. 183.

<sup>14</sup> Cf. *Ibidem*, p. 183-184.

<sup>15</sup> Cf. *Ibidem*, p. 185.

<sup>16</sup> MOORE. G. E. *Principia Ethica*. (Revised Edition). Cambridge: Cambridge University Press, 1993, p. 92.

ser naturais se não podem ser investigados empiricamente<sup>17</sup>.

Fatos e propriedades naturais são presumivelmente algo como aqueles fatos e propriedades coletados e estudados pelas ciências naturais e sociais (consideradas de forma geral)<sup>18</sup>.

O natural é seja o que for o objeto de estudo pelas ciências naturais. [...] [Uma] ciência é uma ciência natural caso os seus princípios fundamentais são descobertos *a posteriori*, através da confiança primária nas evidências empíricas<sup>19</sup>.

A vaga e pré-teórica ideia que o naturalista filosófico tenta articular e defender é que tudo – incluindo qualquer particular, eventos, fatos, propriedades e assim por diante – é parte do mundo natural e físico que a ciência investiga<sup>20</sup>.

Podemos tomar essa visão como a predominante. Mas, como nota Copp, é importante ter claro que o naturalismo moral mantém apenas que as propriedades morais são naturais, portanto, empíricas, e não que as propriedades morais devem poder ser explicadas por teorias científicas. Por isso, ele apresenta um critério mais preciso para se poder dizer que uma propriedade é natural. O critério é o seguinte:

*Propriedade Natural:* Uma propriedade é natural se, e somente se, qualquer proposição sintética sobre a sua instanciação que pode ser conhecida, pode ser conhecida apenas empiricamente<sup>21</sup>.

Proponho que compreendamos nesse sentido as propriedades naturais, pois as teorias naturalistas que iremos considerar adiante aparentemente concordam com essa definição.

---

<sup>17</sup> STURGEON, N. Moore on Ethical Naturalism. *Ethics*, 113 (3), p. 528-556, 2003, p. 543.

<sup>18</sup> BRINK, 1989, *Op. Cit.*, p. 22.

<sup>19</sup> SHAFER-LANDAU, R. Ethics as Philosophy: A Defense of Ethical Non naturalism. In: HORGAN, T. & TIMMONS, M. (eds). *Metaethics after Moore*. New York: Oxford University Press, 2006, p. 209-232, p. 212-213.

<sup>20</sup> TIMMONS, M. *Morality without Foundations*. Oxford: Oxford University Press, 1999, p. 12.

<sup>21</sup> Cf. COPP, 2003, *Op. Cit.*, p. 185.

É importante notarmos duas coisas agora. Acima afirmei que o naturalismo moral é, *em grande medida*, extensão do naturalismo metafísico. Utilizei essa expressão em itálico porque, em primeiro lugar, não é uma condição necessária o ser um naturalista metafísico para se ser um naturalista moral e, em segundo lugar, porque o ser um naturalista metafísico não é condição suficiente para se ser um naturalista moral. Deixe-me explicar isso. No que diz respeito ao primeiro ponto, é possível ser um naturalista moral sem se comprometer com o naturalismo metafísico. É possível sustentar que as propriedades morais são naturais e, ao mesmo tempo, admitir a existência de entidades supernaturais, tais como números, por exemplo. Pode-se admitir até mesmo Deus e sustentar o naturalismo moral, desde que se pressuponha que a moralidade é independente desta divindade. E em relação ao segundo ponto, cabe dizer que o naturalismo metafísico *não implica* o naturalismo ético. Acima, já consideramos isso brevemente. É possível ser um naturalista metafísico e ser um expressivista, um construtivista ou um relativista. Isso significa que se pode sustentar que não há propriedades supersensíveis, como os valores morais, e defender que esses valores são uma construção humana e não que são fatos naturais objetivos<sup>22</sup>.

Até agora estivemos falando em naturalismo moral em geral. Mas há pelo menos dois tipos de naturalismo. O primeiro deles consiste em dizer que há uma relação de identidade analítica entre propriedades morais e propriedades naturais. Podemos chamar esta abordagem de *naturalismo analítico* e defini-la do seguinte modo.

*Naturalismo Analítico:* Termos fundamentais como ‘bom’ tem definições naturalistas analiticamente verdadeiras<sup>23</sup>.

Note que essa é uma tese semântica. O naturalismo moral é uma tese metafísica. Qual é a relação entre ambos? De acordo com Horgan e Timmons<sup>24</sup>, tradicionalmente, o naturalismo como tese metafísica e o naturalismo como tese semântica foram estritamente associados, de modo que defender o naturalismo metafísico era defender o naturalismo analítico e vice versa. Isso pode ser explicado tomando a estratégia

<sup>22</sup> Esses dois pontos são notados por Copp. Cf. COPP, 2003, *Op. Cit.*, p. 180.

<sup>23</sup> Cf. HORGAN, T. & TIMMONS, M. Troubles for New Wave Moral Semantics: The Open Question Argument Revived. *Philosophical Papers*, Volume 21, No. 3, 1992b, p. 154.

<sup>24</sup> Cf. *Idem*.

tradicional do naturalismo para acomodar as propriedades morais ao mundo natural. Argumentava-se que os predicados morais eram *sinônimos* aos predicados usados para se referir a propriedades naturais. Assim, se se assumisse o utilitarismo, por exemplo, o predicado ‘moralmente correto’ significava ‘capaz de maximizar o prazer’ ou ‘capaz de, ao menos, diminuir a dor’. Se os dois predicados – ‘moralmente correto’ e ‘capaz de maximizar o prazer’ – são sinônimos, então devem expressar a mesma propriedade. E dado que ‘capaz de maximizar o prazer’ é uma propriedade natural<sup>25</sup>, então ‘moralmente correto’ também deve ser. Até certo momento da história da filosofia moral, não havia distinção nenhuma entre duas formas de naturalismo moral e, por conseguinte, o naturalismo analítico era a tese padrão do naturalismo moral.

O segundo tipo de naturalismo é a tese de que há uma relação de identidade sintética entre propriedades morais e propriedades naturais. Essa proposta é comumente conhecida como *naturalismo não reducionista*. Pode ser definido assim:

*Naturalismo Sintético*: Termos morais fundamentais como ‘bom’ tem definições naturalistas sinteticamente verdadeiras<sup>26</sup>.

O naturalismo sintético foi desenvolvido como uma alternativa do naturalismo analítico. Como veremos, este é vulnerável a uma série de críticas que aquele não é. A ideia básica é de que propriedades morais são naturais, mas que não se relacionam analiticamente. Neste sentido, o naturalista sintético é *não eliminativista* no sentido de que mantém o vocabulário moral, mas sustenta que os termos morais referem a propriedades naturais. Outra característica dessa abordagem é que propriedades morais são *constituídas* por propriedades naturais. Assim, o naturalista pode manter que não há uma única propriedade natural a que a propriedade moral corresponde, mas um conjunto indefinido

---

<sup>25</sup> Por que ‘prazer’ é uma propriedade natural? A seguinte suposição é razoável. Os psicólogos e neurocientistas têm progredido cada vez mais na análise de alguns sentimentos ou estados mentais, de modo que podem mapear que tipo de regularidade cerebral ocorre quando alguém tem determinado estado mental. Portanto, se poderia dizer que ‘prazer’ é um conjunto determinado de estados mentais que pode ser medido, controlado e até mesmo previsto. Isso claramente faz parte da constituição natural do mundo.

<sup>26</sup> Cf. *Ibidem*, p. 167.

delas. Essas são apenas algumas características gerais do naturalismo sintético. Uma explicação adequada requer muito mais que isso, algo que será feito adiante. Por hora, é suficiente saber que há uma tese metafísica mais geral, o naturalismo metafísico, que, grosso modo, sustenta que “tudo o que há” faz parte do mundo natural; que isso tem grande influência para a tese do naturalismo moral, que sustenta que fatos e propriedades morais são naturais; e que há dois tipos de naturalismo, o naturalismo analítico e o naturalismo sintético.

A versão naturalista do realismo moral é uma posição muito atrativa. Oferece objetividade aos juízos morais, nos certifica, supostamente, de que podemos ter conhecimento moral, que os juízos morais são passíveis de verdade e falsidade, que pelo menos alguns são verdadeiros e que, portanto, existem respostas corretas aos problemas morais. Talvez seja bem mais atrativa do que a versão não naturalista do realismo moral. Ela não se compromete com a existência de certas propriedades *sui generis*. Assim, não assume o compromisso de fornecer uma explicação de que tipo de propriedades são essas. Simplesmente advoga que as propriedades morais “estão aí”. Nesse sentido, não precisa assumir o problema de explicar que tipo de faculdade epistêmica é essa denominada intuição. Pressupõe que o conhecimento das propriedades morais é similar ao conhecimento de qualquer propriedade que comumente dizemos que conhecemos no que diz respeito às fontes de justificação. Assumir o naturalismo moral é dizer que a moralidade está livre de mistérios, por assim dizer, pois, ao postular que as propriedades morais são naturais, não se compromete com a dificuldade de explicar a existência de um domínio de valores morais que é distinto do domínio dos fatos não morais.

Mas ao mesmo tempo em que a posição parece ser atrativa, ela tem de lidar com uma série de problemas. O naturalismo moral tem sido atacado desde que se começou a fazer metaética. As estratégias usualmente consistem em dizer que há algum tipo de diferença entre propriedades morais e naturais, ou que propriedades morais não são objetivas, que dependem de construção humana, que são propriedades de um tipo muito peculiar, que propriedades morais têm certas características que as propriedades naturais não têm, e assim por diante.

Penso que já tenho uma boa ideia do que é o naturalismo moral. Por isso, a partir de agora gostaria de apresentar uma série de críticas comumente direcionadas a esta teoria ética. Essas críticas serão principalmente voltadas ao naturalismo reducionista. O objetivo será mostrar que a versão analítica do naturalismo parece se comprometer com vários problemas. Depois disso poderemos ver como o naturalismo

não reducionista se sai diante desses problemas. Começamos pela objeção clássica, o argumento da questão aberta de G. E. Moore.

## 2. Moore e o argumento da questão aberta

A pouco mais de um século, Moore chamou a atenção para uma suposta falácia que teria sido cometida pela maioria dos autores que se propuseram a escrever algo sobre ética até então. Essa falácia consistia em se definir *bom* em termos de alguma propriedade natural. Para sustentar que isso era uma falácia, ele lançava mão de um argumento, que ficou conhecido como ‘argumento da questão aberta’. Todos nós sabemos que ‘falácia naturalista’ não é uma das melhores expressões que se poderia ter usado. Isso porque o suposto erro não é cometido somente por aqueles que tentam definir *bom* como alguma propriedade natural, mas como alguma propriedade metafísica também (ou como qualquer outra propriedade que se queira definidora de *bom*). O ponto importante é que, mesmo que hoje se diga que Moore não denunciou falácia alguma<sup>27</sup>, o seu argumento foi um marco na história da ética. Quando ele o apresentou, muitos eticistas ou o endossaram<sup>28</sup> ou foram ao menos inspirados a pensar que deveria haver algo errado em se dizer que propriedades morais são naturais. Isso influenciou, em grande medida, o desenvolvimento de outras teorias metaéticas. Além disso, vários filósofos, posteriores a Moore, que buscaram fornecer uma refutação do naturalismo, tinham algo do argumento da questão aberta “em seu sangue”. E isso é válido até mesmo hoje em dia<sup>29</sup>.

Como afirma Sturgeon, o argumento da questão aberta é difícil de interpretar e várias vezes confuso<sup>30</sup>. Isso porque o próprio Moore foi bastante crítico a respeito de seu próprio argumento. Uma caracterização rigorosa da sua posição certamente requer algum grau de exegese da sua obra. Não é isso que irei fazer aqui. Apesar de Moore ter posteriormente

---

<sup>27</sup> DARWALL, S., GIBBARD, A. & RAILTON, P. *Moral discourse & Practice: Some Philosophical Approaches*. New York, Oxford: Oxford University Press, 1997, p. 3.

<sup>28</sup> Ver, por exemplo, AYER, A. J. *Language, Truth and Logic*. Harmondsworth: Penguin Books, 1971 e STEVENSON, C. L. *Ethics and Language*. New Heaven, Yale University Press, 1958.

<sup>29</sup> A reformulação do argumento da questão aberta de Terence Horgan e Mark Timmons, que discutiremos adiante, têm rendido uma quantidade considerável de artigos hoje.

<sup>30</sup> Cf. STURGEON, N. Ethical Naturalism. In COPP, D. (ed) *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, 2006, p. 94.

recusado algumas coisas que havia primeiramente assumido no *Principia*, existe um consenso geral bem difundido sobre o que é o argumento da questão aberta<sup>31</sup>.

O ataque de Moore ao naturalismo ético pretendeu mostrar que *bom* não pode ser analisado ou definido em termos de um conjunto de propriedades quaisquer, sejam elas naturais ou não. É importante notar que a preocupação de Moore não era com uma investigação linguística do termo, ou seja, com como usamos a expressão na linguagem natural, mas com uma investigação conceitual. Há três pressuposições que são centrais ao argumento de Moore. Em primeiro lugar, Moore pensava que se *bom* é uma propriedade natural, então deve haver uma especificação de qual propriedade é, quer dizer, se deve saber a que tipo de propriedade natural a propriedade moral se refere. Ele próprio exemplificava citando, por exemplo, John Stuart Mill ao dizer que *bom* é *prazer*<sup>32</sup>. Em segundo lugar, ele assumia que, se o naturalista estivesse certo e houvesse uma identidade entre propriedades morais e naturais, então os termos, moral e natural, deveriam ser sinônimos. Além disso, ainda a respeito desta segunda pressuposição, se dois termos fossem sinônimos, qualquer falante competente deveria poder saber disso. E, em último lugar, Moore sustentava que tinha um teste para saber se dois termos eram sinônimos ou não. Bastava substituir um pelo outro numa mesma sentença e, se o significado ou proposição fosse preservado, então isso seria evidência de que os termos em questão eram sinônimos. Passemos, então, de modo mais específico ao seu argumento.

Podemos colocá-lo na forma canônica do seguinte modo<sup>33</sup>.

P1. Suponha-se que o predicado '*bom*' é sinônimo, ou analiticamente equivalente, ao predicado natural '*N*'.

C1. Então, é parte do significado da afirmação de que "*x* é *N*" que "*x* é *bom*".

---

<sup>31</sup> É importante dizer que aqui o foco é o argumento da questão aberta somente. Não será feita uma análise cuidadosa da evolução e mudanças da posição de Moore e de outras críticas que ele supostamente apresentou ao naturalismo.

<sup>32</sup> Cf. MILL, J. S. *Utilitarianism*. New York, Bantan Books, 1993.

<sup>33</sup> Embora faça leves alterações, adoto o modelo de A. Miller de apresentar o argumento. Cf. MILLER, A. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press, 2011, p. 13-14.

P2. Mas, então, alguém que se pergunte seriamente sobre um  $x$ , que é  $N$ , se é *bom*, estaria cometendo uma confusão conceitual.

P3. Dada qualquer propriedade natural  $N$ , é sempre uma *questão aberta* se um  $x$ , que é  $N$ , é ou não *bom*. Quer dizer, é sempre uma questão *significante* perguntar de um  $x$ , que é  $N$ , se é ou não *bom*.

C2. Assim, não pode ser o caso que ‘*bom*’ seja sinônimo ou analiticamente equivalente com o predicado natural ‘ $N$ ’.

C3. Portanto, a propriedade de *ser bom* não pode, como uma questão de necessidade conceitual, ser idêntica à propriedade de *ser  $N$* .

Começemos por P1 e C1. O que Moore nos pede para supor de início é que a posição naturalista é verdadeira. Primeiro ele supõe que o predicado ‘*bom*’ é sinônimo de algum predicado natural ‘ $N$ ’. Uma relação de sinonímia entre dois termos ou conceitos estabelece que ambos possuem o mesmo significado, ou algo muito próximo a isso. É por isso que existe a distinção entre sinônimos *perfeitos*, tais como ‘cão’ e ‘cachorro’, ‘automóvel’ e ‘carro’, ‘após’ e ‘depois’, etc., e sinônimos *imperfeitos*, como ‘feliz’ e ‘alegre’, ‘cidade’ e ‘município’ etc. É importante notar que a sinonímia não capta precisamente, desde um ponto de vista filosófico, a identidade de significado entre dois termos, como podemos ver no caso dos sinônimos imperfeitos. É por isso que Moore introduz uma noção filosoficamente mais precisa para caracterizar a relação entre *ser bom* e possuir uma propriedade natural  $N$ , que é a noção de *analiticidade*. Esta noção estabelece que há uma relação *logicamente idêntica* entre dois conceitos quando se diz que um é outro. Imagine que digo que ‘Todo solteiro é não casado’. O que estou expressando com esta sentença é uma relação de analiticidade entre ‘solteiro’ e ‘não casado’. ‘Solteiro’ e ‘não casado’ compartilham, necessariamente, todas as características constitutivas. Se dois objetos, termos ou conceitos são analiticamente equivalentes, então um não pode ter uma propriedade sequer que outro não tenha. Nesse sentido, o que Moore pede para supor é que é o caso que ‘*bom*’ e uma propriedade

natural '*N*' instanciam exatamente a mesma propriedade. Apesar de a expressão 'sinônimo' não ser filosoficamente precisa para este caso, Moore a utiliza, pois ele acreditava que se dois termos tem exatamente as mesmas propriedades então eles devem ser sinônimos. Se supormos que '*bom*' e '*N*' são analiticamente equivalentes, então é certo que temos que supor também que '*x é N*' significa o mesmo que '*x é bom*'. Portanto, C1 se segue de P1 por razões óbvias.

Passemos a P2. Aqui Moore nos pede para considerar o que se segue da primeira conclusão. Se pressuponho, como o naturalista faz, que '*bom*' é analiticamente equivalente a uma propriedade natural '*N*', então não posso, sem cometer um erro conceitual, perguntar, de '*N*', se é '*bom*'. Esta é uma pressuposição razoável, pois se admito que *x* é analiticamente equivalente a *y*, quer dizer, que tem exatamente todas as propriedades de *y*, então estaria cometendo uma confusão conceitual ao perguntar se *x* é *y*. Confusão porque de algum modo teria que supor que *x* tem alguma propriedade que *y* não tem, e vice-versa, para a questão fazer sentido, mas não posso supor isso já que admiti de início que *x* e *y* possuem uma relação de analiticidade. Podemos considerar o exemplo do próprio Moore. Se assumirmos que *bom* denota a propriedade natural *prazer*, então quando digo 'O *prazer é bom*' estou dizendo o mesmo que 'O *prazer é prazer*'. Portanto, comete uma confusão conceitual quem pergunta do *prazer* se é *bom*, pois estaria perguntado algo como 'O *prazer é prazer?*'. P2 está como segunda premissa, mas não deixa de ser a conclusão da conjunção entre P1 e C1. O cometer a confusão conceitual ao perguntar sobre a suposta propriedade natural que define *bom* se é *bom*, se segue se assumirmos que o naturalismo é verdadeiro.

No entanto, P3 vem nos mostrar que quando falantes competentes perguntam sobre a suposta definição de *bom* se é *bom*, eles não estão cometendo uma confusão conceitual. De fato, a questão, supostamente, é genuína. Esta é a premissa central do argumento da questão aberta. Moore achava que, se ela é verdadeira, então o naturalismo tem que ser falso. Podemos dizer que o raciocínio de Moore é algo parecido com o seguinte. Suponha que 'solteiro' e 'não casado' são analiticamente equivalentes. Então, para falantes competentes, não é uma questão *significante* 'É *x*, que é solteiro, não casado?', pois não tem como um *x* ser 'solteiro' e não ser 'não casado'. A expressão usada frequentemente para se referir a este tipo de questão é de que é uma *questão fechada*. Se o naturalismo é verdadeiro e as propriedades morais são analiticamente equivalentes às propriedades naturais, então o mesmo deveria ocorrer.

Quer dizer, a questão ‘É  $x$ , que é  $N$ , *bom*?’<sup>34</sup> deveria ser fechada, pois, se o naturalismo é verdadeiro, não tem como  $x$  ser ‘ $N$ ’ e não ser ‘*bom*’. Mas o problema é que ‘É  $x$ , que é  $N$ , *bom*?’ não parece ser uma questão fechada. Falantes competentes não associam naturalmente a propriedade de *ser bom* à propriedade de *ser N*. Isso significa que, no caso das propriedades morais, sempre faz sentido perguntar sobre a definição de *bom* se é *bom*. Nas palavras de Moore, é uma *questão aberta*<sup>35</sup>.

Às vezes se explica essa premissa do argumento da questão aberta utilizando-se da teoria da linguagem pressuposta por Moore<sup>36</sup>. Moore era adepto da teoria descritivista do significado. Em termos muito simples, esta teoria sustenta que o significado de um termo é dado através do conjunto de descrições definidas que os falantes competentes associam a este termo. Assim, o significado de ‘Platão’, por exemplo, é conferido por descrições definidas tais como ‘O discípulo de Sócrates’, ‘O homem que escreveu os *Diálogos*’, ‘O filósofo que formulou o Mito da Caverna’ e assim por diante. Todas essas descrições, e mais um número indefinido que se referem a Platão, designam um mesmo objeto, a saber, Platão; e o significado de ‘Platão’ é conferido por esse número indefinido de descrições. Nesse sentido, diz-se que, dado que Moore é um adepto da teoria descritivista do significado, então para dois termos serem analiticamente equivalentes, ou sinônimos, os falantes competentes de uma dada língua devem poder associar exatamente o mesmo conjunto de descrições definidas tanto a ‘*bom*’ quanto a ‘ $N$ ’. Mas os falantes competentes não associam o mesmo conjunto de descrições definidas à propriedade moral e à propriedade natural. Por isso, é uma questão aberta se a suposta propriedade natural que é definidora da propriedade moral é *boa* ou não.

Se P3 é verdadeira, se ‘É  $x$ , que é  $N$ , *bom*?’ é uma questão aberta, se os falantes competentes associam descrições definidas diferentes às propriedades morais e naturais, então temos, obviamente, C2. Quer dizer, propriedades morais e propriedades naturais não podem ser analiticamente equivalentes. Se não podem, e se o naturalismo se compromete com a tese de que podem – ou melhor, que *são* –, como explícito em P1, C1 e P2, então C3 se segue. *Bom*, então, não pode ser uma propriedade natural. Esse é o argumento da questão aberta.

---

<sup>34</sup> Ou, como Horgan e Timmons sugeriram, ‘É  $x$ , que é *bom*,  $N$ ?’ Cf. HORGAN & TIMMONS, 1992b, *Op. Cit.*, p. 155.

<sup>35</sup> Cf. MOORE, 1993, *Op. Cit.*, §§14-16, p. 69ss.

<sup>36</sup> Cf. BRINK, D. O. Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy*, 18 (2), 2001, p. 154-176.

Mas, como vimos na seção anterior, há uma distinção entre o naturalismo moral e o naturalismo analítico, no sentido de que nem todo naturalismo moral é naturalismo analítico. Resta ainda saber por que, ao supostamente refutar o naturalismo analítico, Moore estaria refutando o naturalismo moral. Horgan e Timmons explicam isso dizendo que, na época de Moore, adotar o naturalismo moral metafísico era adotar o naturalismo analítico<sup>37</sup>. Isso porque havia uma aceitação do critério de identidade entre dois termos através da sinonímia, ou, pra ser mais preciso, analiticidade<sup>38</sup>. Então para se poder dizer que dois termos mantêm uma relação de identidade esses termos tinham que manter uma relação de identidade analítica. Isso significa, em última instância, que a verdade do naturalismo moral depende da verdade do naturalismo analítico. Por isso, Moore achou que, refutando o naturalismo analítico estaria refutando o naturalismo moral. Seu argumento mais geral, então, pode ser resumido do seguinte modo.

P1. A verdade do naturalismo moral depende da verdade do naturalismo analítico.

P2. O naturalismo analítico é falso.

C1. Portanto, o naturalismo moral é falso.

A segunda premissa é supostamente verdadeira com base no argumento da questão aberta.

O argumento da questão aberta tem lá os seus problemas. Algum tempo depois de Moore o ter apresentado, W. Frankena, por exemplo, publicou um influente artigo em que sustentava que o argumento da questão aberta era circular. Grosso modo, ele dizia que para o argumento de Moore funcionar, em P3 ele já deveria pressupor a falsidade do naturalismo, pois, senão, não faria sentido dizer que 'É *x*, que é *N*, bom' é uma questão aberta. Mas a falsidade do naturalismo é justamente a conclusão que o argumento da questão aberta almeja chegar. Por isso, seria circular<sup>39</sup>. Mas o interesse aqui não é discutir a

---

<sup>37</sup> Cf. HORGAN & TIMONS, 1992b, *Op. Cit.*, p. 154s.

<sup>38</sup> Obviamente, não se tinha ainda as teorias externalistas da referência e, portanto, não se tinha descoberto que o serem analíticos não é condição necessária para dois termos terem uma relação de identidade.

<sup>39</sup> Cf. FRANKENA, W. The Naturalistic Fallacy. *Mind*, No. 48, 1938, p. 464-477.

validade dessas objeções. Apresentarei as razões comumente utilizadas para recusar o argumento de Moore na segunda parte deste trabalho. Por hora, é importante notar que o argumento exerceu grande influência. Hare apresentou uma reformulação deste argumento em seu *The Language of Morals*<sup>40</sup>. No entanto, Hare não concordou com a posição positiva de Moore, quer dizer, a tese que dizia o que as propriedades morais são. Moore assumiu uma forma de realismo não naturalista; aquilo que na seção anterior apresentei brevemente como intuicionismo. Hare sempre foi um crítico do intuicionismo e, por isso, discordou de Moore nesse ponto. O ter concordado com a parte negativa da teoria de Moore e discordado da positiva obrigou Hare a desenvolver outra teoria ética para justificar o que os juízos morais são. Ao fazer isso, ele forneceu uma série de objeções a mais ao naturalismo moral. Consideremos, a partir de agora, as ideias de Hare.

### 3. R. M. Hare e o prescritivismo

Há quem diga que as mais influentes críticas do século XX ao naturalismo moral não estão no *Principia Ethica*, mas em *The Language of Morals*. Não sei se podemos medir o alcance teórico de Moore ou Hare e dizer precisamente qual deles representa a mais robusta oposição ao naturalismo. Mas é bastante razoável supor que Hare desempenhou um importante papel na tentativa de mostrar que essa teoria ética é falsa. Dado que houve uma tendência muito maior a aceitar a parte negativa da teoria de Moore, quer dizer, seu argumento da questão aberta, e recusar a parte positiva, a metaética acabou tomando rumos não cognitivistas depois de Moore. O emotivismo floresceu de forma significativa das páginas do *Ethics and Language* de C. L. Stevenson e do *Language, Truth and Logic* de A. J. Ayer. Hare concordava com as pressuposições gerais emotivistas a respeito da não factualidade semântica e do não cognitivismo psicológico dos juízos morais. Mas achava que essa abordagem implicava num erro irreversível, o de não conferir racionalidade à atividade moral. Por isso, no início da década de cinquenta começou a desenvolver uma teoria que ficou conhecida como prescritivismo universal. Nesse percurso, Hare apresentou várias críticas ao naturalismo moral que, supostamente, eram suficientes para mostrar que essa teoria é falsa.

---

<sup>40</sup> HARE, R. M. *The Language of Morals*, New York: Oxford University Press, 1952.

Gostaria de dedicar essa seção para apresentar algumas das principais teses de Hare e mostrar em que sentido essas teses consistem em objeções ao naturalismo moral. Já que o ponto de partida para o desenvolvimento de sua teoria fora a recusa do emotivismo, é importante começar por aqui. Vou apresentar algumas ideias gerais do emotivismo e mostrar qual foi a razão principal para Hare o ter recusado. Assim, passarei para a sua defesa de que a moralidade é uma atividade racional. Veremos que há uma pressuposição importante, a saber, a ideia de que juízos morais não são descrições de fatos da realidade. Isso significa que Hare recusa o naturalismo. A partir disso, poderemos passar para o desenvolvimento do prescritivismo e seus argumentos contra o naturalismo, o que será o assunto principal dessa seção.

Começemos, então, pelo emotivismo. Esta teoria ética foi proposta pela primeira vez por Ayer em seu livro *Language, Truth and Logic* emergindo de sua aderência ao positivismo lógico. Os positivistas tinham o objetivo de reformar a filosofia a partir de uma teoria do significado denominada verificacionismo. Ayer, em particular, tinha um objetivo bastante ambicioso, que não era resolver todos os problemas da filosofia, mas mostrar quais *poderiam* ser resolvidos e quais *não* poderiam. Através disso seria possível, ele acreditava, fazer uma separação clara entre aquilo que tinha significado e aquilo que não tinha e, nesse sentido, descartar o que não era genuinamente significativo. A teoria verificacionista do significado é basicamente composta pelo princípio de verificabilidade, que reivindica o seguinte.

*Princípio de Verificabilidade:* Uma sentença tem significado se, e somente se, é ou analítica ou empiricamente verificável<sup>41</sup>.

Mas o que isto quer dizer? Ora, uma sentença é analítica quando a sua verdade depende unicamente da definição de seus termos. ‘Todo solteiro é não casado’ é analítica, pois ‘solteiro’ é definido em termos de ‘qualquer um que seja não casado’. É característico desse tipo de sentença a não necessidade de recorrência à experiência para determinar a sua verdade. Basta atentar apenas para os termos que estão sendo usados. O outro extremo são as sentenças sintéticas. Uma sentença é sintética quando a sua verdade depende de como o mundo é, quer dizer, quando há a necessidade de recorrer aos fatos para determinar se é

---

<sup>41</sup> Cf. AYER, 1971, *Op. Cit.*, p. 7.

verdadeira ou falsa. Considere ‘Está chovendo’ ou ‘O quarto é escuro’. Obviamente trata-se de sentenças sintéticas, pois tenho de olhar para o mundo para determinar se está ou não chovendo e se o quarto é ou não escuro. Além disso, o princípio fala de ‘empiricamente verificável’. Uma sentença é empiricamente verificável se há um conjunto de observações que mostram que é verdadeira. Tomemos novamente ‘Está chovendo’. Posso olhar lá fora e ver se realmente há nuvens, se está caindo gotas d’água, se o chão está ficando molhado, etc. Temos que, uma sentença tem significado se, e somente se, ou é verdadeira em virtude da definição dos termos ou se há um conjunto de observações empíricas que evidenciam que é ou não verdadeira. Qualquer sentença que escapar a essa fronteira não tem significado.

Agora, consideremos o que isso tem a ver com as sentenças morais. Tomemos o juízo moral ‘O assassinato é errado’. A sentença obviamente não é analítica, pois não faz parte da definição de ‘assassinato’ a propriedade de ser ‘errado’. A sentença também não é empiricamente verificável, pois a propriedade ‘errado’ não expressa uma proposição ou um fato que pode ser perceptível através dos sentidos<sup>42</sup>. Na sentença ‘O quarto é escuro’ há duas coisas importantes, a saber, a propriedade de *ser um quarto* e a propriedade de *ser escuro*. Se há um objeto, ou neste caso um conjunto de fatores, que instancia a propriedade de ser um quarto e de ser escuro, então a sentença é verdadeira; se não, falsa. Isso é determinado através de verificação empírica. Mas no caso de ‘O assassinato é errado’ pode até haver uma ação que instancia a propriedade de ser assassinato, mas não há um objeto ou ação que instancie a propriedade da *incorrecção*. Quer dizer, posso acessar o assassinato através dos sentidos, mas olho para o mundo e não vejo o ‘errado’. Portanto, se o juízo ‘O assassinato é errado’ não é analítico nem empiricamente verificável, então, de acordo com o critério de verificabilidade, não tem significado. Assim, para o emotivismo, as sentenças morais em geral não tem significado.

O emotivismo moral de Ayer surge a partir dessa constatação. Os juízos morais, embora não sejam passíveis de significado no sentido em que o critério de verificabilidade propunha, possuem um *significado emotivo*. Ayer nos apresenta esta ideia em uma passagem bastante citada.

Se eu disser para alguém, ‘Você agiu incorretamente ao roubar aquele dinheiro’ não

---

<sup>42</sup> Ou, pelo menos, esta era uma tese dos emotivistas.

estou afirmando nada mais do que se eu simplesmente dissesse ‘Você roubou aquele dinheiro’. Ao adicionar que essa ação é errada, não estou fazendo nenhuma afirmação para além disso. Estou simplesmente evidenciando a minha desaprovação moral. É como se eu dissesse, ‘Você roubou aquele dinheiro’ em um tom de horror particular, ou escrevesse com a adição de alguma marca especial de exclamação. O tom, ou a marca de exclamação, não adiciona nada ao significado literal da sentença. *Serve apenas para mostrar que a expressão disso está atrelada a certos sentimentos do falante*<sup>43</sup>.

Isso significa que quando faço um juízo moral o predicado moral – errado, certo, bom, ruim... – evidencia uma *atitude* ou *sentimento* meu em relação ao ato a que o predicado está associado. Não se trata de uma expressão com valor de verdade, mas da expressão de uma emoção diante de algum objeto, pessoa ou ação. O assassinato não é *objetivamente* errado; há a necessidade de um agente valorando-o. Quando expressamos um juízo moral como o de que ‘É incorreto não cumprir as promessas’ é como se estivéssemos dizendo ‘Uuuuhhhhhh!!!’, no sentido de vaiar, para o não cumprimento de promessas. Ou, quando expressamos o juízo de que ‘É correto ser honesto’ é como se estivéssemos dizendo ‘Aaaaaêêêêê!!!’, no sentido de aprovação, para o ser honesto. É por isso que o emotivismo é uma forma de *projetivismo*. A propriedade moral não pertence ao reino dos fatos, ou a qualquer outro reino, mas apenas ao sujeito que projeta a sua emoção – de aprovação ou não aprovação – a um determinado ato ou pessoa; e essa emoção é tornada pública através dos predicados avaliativos. Essa é a ideia básica do emotivismo.

É importante distinguir entre *ter* uma emoção e *dizer* se tem. Para Ayer, fazer um juízo moral é ter uma emoção, tendo o agente a tornado pública dizendo ‘*x* é bom’, ‘*x* deve ser feito’ etc., ou não. Dizer ‘*x* acha que o assassinato é errado’ é algo bastante diferente daquilo que quer Ayer. Esta segunda abordagem é passível de verdade ou falsidade, pois *x* pode achar que o assassinato é errado ou não e nos é possível verificar isso. Portanto, ‘*x* acha que o assassinato é errado’ ou ‘Minha

---

<sup>43</sup> Cf. AYER, 1971. *Op. Cit.* p. 142. Itálico adicionado.

comunidade acha que fazer y é bom' são descrições de fatos e, portanto, não há *expressão* de emoção envolvida nestas sentenças<sup>44</sup>.

Mas o princípio de verificabilidade não é condição necessária para o emotivismo. Stevenson endossou a teoria emotivista, mas sem recorrer ao princípio de Ayer. Seu principal argumento é sobre a natureza da motivação moral. Este argumento pode ser usado para defender o não cognitivismo em geral, mas, em conjunto com a análise das sentenças morais elaborada por Stevenson, torna-se um argumento a favor do emotivismo em específico. Há duas pressuposições aqui. A primeira é a distinção humiana entre *desejos* e *crenças*. Hume em seu *Tratado* nos diz que as crenças não nos motivam a agir, apenas os desejos o fazem. Podemos *crer* verdadeiramente que nossa casa está pegando fogo ao mesmo tempo em que estamos dentro, mas se não tivermos o *desejo* de não morrer não seremos motivados a fugir das chamas. Há a famosa frase de Hume que diz “a razão é escrava das paixões”<sup>45</sup>, que quer dizer que a razão nos diz apenas quais são os fatos (crenças), não nos conduz a nenhuma ação ou volição, mas somente as paixões (desejos) o fazem. A segunda pressuposição é a de que os juízos morais necessariamente nos motivam, ao menos em parte, a agir de acordo. Em outras palavras, a crença no juízo moral *p* é uma razão para a ação de acordo com *p*<sup>46</sup>. Nesse sentido, se expresso o assentimento ao juízo de que ‘O assassinato é errado’ estou, em partes, inclinado a não cometer assassinato ou a condenar sinceramente quem o comete. Dessas duas pressuposições – crenças não nos motivam e juízos morais, ao menos em parte, nos motivam à ação – Stevenson conclui que juízos morais não são crenças. Portanto, não podem possuir valor de verdade, pois expressam um conteúdo não cognitivo do sujeito, isto é, suas emoções ou sentimentos de aprovação ou reprovação perante certos atos ou pessoas.

Stevenson divide o juízo moral em duas partes atribuindo um papel diferente a cada uma delas. Há o elemento *expressivo* e o *causativo* sempre que ocorre um juízo moral. Se digo que ‘O assassinato é errado’ estou *expressando* uma atitude minha em relação a este ato e, além disso, estou tentando *convencer* os meus ouvintes a não cometerem o assassinato ou a reprová-lo também. É como se estivesse dizendo ‘Eu

---

<sup>44</sup> Cf. AYER, 1971. *Op. Cit.* p. 140.

<sup>45</sup> HUME, D. *A Treatise of Human Nature*. Oxford: Clarendon Press, 1978, p. 415.

<sup>46</sup> Esta tese será retomada com muito mais detalhe adiante. Por hora, apenas assumamo-la para ver qual é o argumento de Stevenson.

desaprovo o assassinato; faça-o também!'. O juízo tem um elemento *descritivo* que informa sobre a emoção do indivíduo que o expressa e um elemento *imperativo*, que tem por objetivo evocar no ouvinte, ou leitor, o mesmo sentimento ou emoção<sup>47</sup>. Essa distinção, como veremos adiante, é de grande importância para entender o argumento de Hare.

Hare concordou com a ideia de Stevenson de que os juízos morais tem um significado imperativo<sup>48</sup>. Mas argumentou que o emotivismo implica num problema fundamental. Os autores que sustentavam a ideia de que os juízos morais tem um significado emotivo normalmente endossavam, também, as duas ideias seguintes. (i) Quem assumia o critério de verificabilidade como o elemento determinante para separar o que tinha significado do que não tinha se comprometia com a ideia de que os juízos morais não tinham significado (no sentido de como os juízos não morais tinham). As sentenças descritivas tinham significado porque a sua verdade podia ser determinada por meio de verificação. Mas imperativos, por exemplo, não podiam. Por isso, sustentavam que os juízos morais não podiam ser regidos por regras lógicas (ou serem assunto do qual trata a lógica), como a conjunção, disjunção, implicação, etc. No início de *Imperative Sentences*, Hare afirma o seguinte.

Os lógicos tem frequentemente tomado como verdadeiro que há uma classe de sentenças que são o assunto próprio da lógica e que eles estão em liberdade para ignorar todas as sentenças que não estão incluídas nessa classe. [...] O tipo de sentenças que são admitidas no campo da lógica são variavelmente referidas como “científicas”, “cognitivas”, “informativas”, “fatuais”, “verdadeiras ou falsas”, “teoréticas”, “referenciais”, “simbólicas”, etc.; e o tipo de sentenças que são excluídas são chamadas “emotivas”, “evocativas”, “não fatuais”, etc<sup>49</sup>.

---

<sup>47</sup> Cf. STEVENSON, 1958, *Op. Cit.* Ver também HARE, R. M. *Sorting Out Ethics*. Clarendon Press: Oxford, 1997, p. 104ss.

<sup>48</sup> Cf. HARE, R. M. *Imperative Sentences*. In: *Mind*, New Series, Vol. 58, No. 229, 1949, p. 21-39 e HARE, R. M. *The Language of Morals*. Oxford: Oxford Clarendon Press, 1952, especialmente a primeira parte.

<sup>49</sup> HARE, 1949, *Op. Cit.*, p. 21. Este tipo de tese pode ser visto em autores como R. Carnap (Ver CARNAP, R. *Meaning and Necessity: A Study in Semantics and Modal Logic*. The University of Chicago Press: Chicago, 1948,

Se as sentenças “emotivas” não são regidas por regras lógicas, então não se pode dizer que há um critério *racional* de lidar com os juízos morais. *X* poderia dizer que ‘O assassinado é errado’ e *y* que ‘O assassinado é correto’ e, dado que os operadores da lógica clássica não se aplicam aos juízos morais, não se poderia dizer que *x* e *y* tem um desacordo moral genuíno, pois não seria possível dizer que os juízos são contraditórios. Então, assumir o emotivismo deveria implicar na não racionalidade da atividade moral. (ii) Além disso – e o que é extensão da primeira ideia – se os juízos morais tratam apenas de expressão de atitudes e sentimentos de aprovação ou desaprovação, então parece que não há meios objetivos de se resolver conflitos morais. Se *x* sustenta que o assassinado é errado e *y* que não é, então parece que a única coisa que a ética tem a nos dizer sobre isso é que *x* tem o sentimento de desaprovação em relação ao assassinado e *y* de aprovação. E nada mais pode ser feito em relação a isso.

Ao se convencer desse suposto problema do emotivismo, a primeira coisa que Hare fez foi tentar mostrar que a moral é uma atividade racional e não meramente um conjunto de sentenças sem sentido. Ao fazer isso, acabou desenvolvendo uma posição que foi o ponto de partida para a recusa do naturalismo. Como dito acima, Hare aceitou de Stevenson que sentenças morais tem *significado imperativo*<sup>50</sup>. As sentenças que tem significado e são passíveis da aplicação dos princípios da lógica clássica são, segundo os emotivistas, sentenças descritivas ou, para usar o termo de Hare, indicativas<sup>51</sup>. A estratégia de Hare consiste em mostrar que os imperativos são regidos por regras lógicas, ao contrário do que sustentavam os verificacionistas, e que, se

---

p. 6), Ayer (Ver AYER, 1971, *Op. Cit.*, capítulo 6), I. A. Richards (Ver RICHARDS, I. A. *Principles of Literary Criticism*. Routledge Classics: London and New York, 2004, p. 268) e C. K. Ogden e I. A. Richards (Ver OGDEN, C. K. & RICHARDS, I. A. *Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. Harcourt, Brace & World: New York, 1946, p. 149).

<sup>50</sup> Sua posição vai além. Ele sustenta que juízos morais *implicam* em imperativos. Esta é uma tese importante. Por que sentenças morais implicam em imperativos requer um argumento. Esse argumento será apresentado no momento adequado. Por ora, suponhamos que juízos morais tem significado imperativo para ver qual o argumento de Hare para sustentar que os juízos morais podem ser regidos por regras lógicas e começarmos, assim, a ver quais são suas críticas ao naturalismo.

<sup>51</sup> HARE, 1952, *Op. Cit.*, p. 5.

os juízos morais implicam em imperativos, então os juízos morais são regidos por regras lógicas. Assim, seria possível conferir racionalidade à atividade moral. Vejamos seu argumento passo a passo.

Primeiro, a distinção entre sentenças imperativas e indicativas. Hare utiliza ‘comando’ e ‘afirmação’ respectivamente como sinônimos de ‘imperativo’ e ‘indicativo’. Ele diz que quando assentimos a uma sentença indicativa, considera-se que somos sinceros se, e somente se, *acreditamos* que é verdadeira. E quando assentimos a um imperativo, considera-se que somos sinceros se, e apenas se, *fazemos* o que o imperativo nos ordena que façamos<sup>52</sup>. Assim, a diferença entre imperativos e indicativos é que enquanto o assentimento sincero aos primeiros implica em *fazer* algo, o assentimento sincero aos segundos implica em *crer* em algo.

A sentença ‘Feche a porta’ é imperativa. A sentença ‘Você vai fechar a porta’ é indicativa. Obviamente, são sentenças diferentes. Mas Hare sugere que reescrevamos uma sentença que expresse o elemento comum entre ambas. Basta colocarmos o verbo no infinitivo e adicionarmos ‘no tempo *t*’ no final da sentença. Seria ‘Você fechar a porta no futuro imediato’. Claramente, esta última sentença não expressa o mesmo significado das duas primeiras. Ela diz apenas sobre o que são a sentença imperativa e a sentença indicativa. Está faltando algo. Ele, então, propõe que adicionemos, no final de ‘Você fechar a porta no futuro imediato’, ‘sim’ e ‘por favor’, ambas as expressões correspondendo ao elemento indicativo e imperativo das sentenças acima. Assim, para o indicativo ‘Você vai fechar a porta’ teríamos algo como ‘Você fechar a porta no futuro imediato, sim’. E para o imperativo ‘Feche a porta’, ‘Você fechar a porta no futuro imediato, por favor’<sup>53</sup>. Para o uso ordinário essas sentenças não expressariam o mesmo significado de ‘Feche a porta’ e ‘Você vai fechar a porta’. Mas, para propósitos filosóficos, Hare acredita que essa remodelação é razoável.

Agora, notemos que há algo similar entre as sentenças imperativa e indicativa remodeladas, a saber, ‘Você fechar a porta no futuro imediato’. A esta parte das sentenças Hare chama *frástica*. E a parte que é distinta, ‘sim’ e ‘por favor’, chama *nêustica*. Hare acredita que essa denominação é apropriada, pois a parte comum entre imperativo e indicativo é a parte que corresponde ao aspecto descritivo da sentença e ‘frástica’ “[...] é derivada de uma palavra grega que significa ‘apontar

---

<sup>52</sup> Cf. *Ibidem*, p. 19-20.

<sup>53</sup> Cf. *Ibidem*, p. 17-18.

ou indicar”<sup>54</sup>. A parte diferente entre as sentenças expressa o modo da sentença, imperativo ou indicativo, e ‘nêustica’ é derivada de uma palavra que significa ‘sinal de assentimento’<sup>55</sup>. Como vimos, a diferença entre indicativos e imperativos se refere ao que implica assentir a essas sentenças. Então, ‘nêustica’ capta bem esse significado.

O próximo passo do argumento é mostrar que as principais operações lógicas, como negação, conjunção, disjunção, etc., ocorrem sempre nas frásticas. Se é assim, e as frásticas de imperativos e indicativos são comuns, então as mesmas operações lógicas devem ocorrer nos imperativos também. Consideremos o operador *conjunção*. Tomemos a sentença indicativa ‘Você vai fechar a porta e permanecer na sala’ e a imperativa ‘Feche a porta e permaneça na sala’. Remodelemo-las de acordo com o método de Hare. Temos ‘Você fechar a porta e permanecer na sala no futuro imediato, sim’ e ‘Você fechar a porta e permanecer na sala no futuro imediato, por favor’. As frásticas são exatamente as mesmas no imperativo e no indicativo: ‘Você fechar a porta e permanecer na sala no futuro imediato’. Note que a *conjunção* está ocorrendo nas frásticas. Quer dizer, na parte que é comum a indicativos e imperativos. Hare sugere que isso pode ser feito com os outros operadores da lógica clássica, disjunção, negação, implicação. As operações sempre ocorrerão na parte que é similar a afirmações e comandos. Portanto, os imperativos também são passíveis de regras lógicas, Hare conclui.

Já que estamos falando de ética, vamos considerar um exemplo moral. Aproveitemos, também, para ver como o operador *negação* se comporta. Digamos que ‘O assassinato é errado’. Suponhamos que Hare está certo e que juízos morais implicam em imperativos. Então teríamos algo similar a ‘Não cometa o assassinato’. O correspondente indicativo seria ‘Você não vai cometer o assassinato’. Seguindo o método proposto por Hare teríamos: ‘Você não cometer o assassinato no futuro imediato, sim’ para o indicativo e ‘Você não cometer o assassinato no futuro imediato, por favor’ para o imperativo. Novamente, a frástica ‘Você não cometer o assassinato no futuro imediato’ é comum a ambos os modos e é onde ocorre a operação lógica. Portanto, a operação lógica ocorre nos imperativos também.

Os emotivistas mantinham que somente os indicativos podiam ser estudados pela lógica. Os imperativos não. E dado que, como Hare assume, os juízos morais implicam em imperativos, então os juízos

---

<sup>54</sup> *Idem.*

<sup>55</sup> *Idem.*

morais não podem ser regidos pela lógica, o que implica no problema da não racionalidade da atividade moral<sup>56</sup>. Hare tenta argumentar que os imperativos compartilham com os indicativos uma parte que é exatamente similar, e onde ocorrem as principais operações lógicas. E, se os emotivistas admitem que os indicativos são regidos pela lógica e se os imperativos compartilham as mesmas frásticas com os indicativos, então tem que admitir, também, que os imperativos são regidos pela lógica. Por conseguinte, tem de admitir também que há um critério de racionalidade entre os juízos morais.

É nesse sentido que Hare rompe com o emotivismo<sup>57</sup>. Mas ele continua aceitando a tese dos emotivistas de que os juízos morais, em

---

<sup>56</sup> Há um outro problema ao emotivismo moral que hoje denominamos clássico. É o problema Frege-Geach. Segundo os emotivistas, quando expresso a sentença ‘O assassinato é errado’ não estou afirmando algo com valor de verdade, mas estou expressando uma atitude de desaprovação diante do assassinato. Agora, consideremos o seguinte raciocínio: (1) Se o assassinato é errado, então influenciar o seu irmão mais novo a cometer assassinato é errado; (2) O assassinato é errado; (3) Então, influenciar o seu irmão mais novo a cometer assassinato é errado. Está claro que a função semântica de ‘O assassinato é errado’ em (1) e (2) é diferente. No primeiro caso, temos um raciocínio condicional. No segundo, temos uma expressão de desaprovação ao assassinato, o que não é o mesmo que (1). Portanto, (3) não se segue. O problema para o emotivista é, então, explicar a função semântica dessas expressões morais em contextos em que não estão sendo usadas diretamente, como no caso condicional acima ou, ainda, por que uma forma tão comum de raciocínio moral é simplesmente um equívoco. Ver GEACH, P. *Ascriptivism*. *Philosophical Review*, 69, p. 221-225.

<sup>57</sup> Aqui estamos considerando a recusa do emotivismo na origem do prescritivismo. Mas vale a pena notar que Hare tem um argumento bem mais maduro e sofisticado para recusar o emotivismo. Ele desenvolveu esse argumento em *Sorting out Ethics* (Cf. HARE, 1997, *Op. Cit.*, capítulo 6). A ideia é similar, quer dizer, o emotivismo implica em não racionalidade. Mas o argumento depende da distinção que J. L. Austin fez em *How to do Things with Words* (Ver AUSTIN, J. L. *How to do things with words*. Oxford: Clarendon Press, 1962) entre atos de fala locutórios, ilocutórios e perlocutórios. O elemento locutório é aquilo que é enunciado pelo ato de fala, ou aquilo que é dito. Um ato de fala é ilocutório quanto tem a propriedade de realizar uma ação, como em ‘Declaro-vos marido e esposa’ ou ‘Batizo-o Carlos’. Uma característica fundamental dos ilocutórios é serem dependentes de certas regras para o uso correto, quer dizer, há a necessidade de um contexto para fazerem sentido. As pessoas em questão devem saber que o ato de fala implica em tais e tais coisas, e assim por diante. Isso significa que há uma lógica nos ilocutórios. O elemento perlocutório se refere àquilo que pode ser evocado no ouvinte. ‘O

algun sentido, *dependem de nossas mentes*. Por isso, começa a desenvolver aquilo que ficou conhecido como prescritivismo universal. Essa teoria ética é composta basicamente por duas teses, a tese da prescritividade e a tese da universalizabilidade. Ambas são características intrínsecas da própria natureza da linguagem moral e, de acordo com Hare, são reveladas a partir do estudo da lógica da linguagem moral ordinária. Nesse sentido, o que Hare tenta fazer é mostrar que a linguagem moral é regida por uma lógica que lhe é inerente. ‘Lógica’ aqui certamente não significa uma abordagem de cunho formal. Como disse John Hare<sup>58</sup>, significa um estudo das regras que regem o uso correto de um termo ou da prática linguística; aquilo que Wittgenstein chamava de “gramática”<sup>59</sup>. O projeto geral e a principal motivação de Hare continua sendo o de mostrar que a moral é racional. Ele acredita que, uma vez compreendida a lógica da linguagem moral, os falantes podem usá-la de forma *consistente*. E o usar de forma consistente implica num tipo de raciocínio prático de cunho utilitarista kantiano que supostamente seria uma forma racional de resolver os conflitos morais<sup>60</sup>. Mas o que nos importa aqui é o primeiro passo de

---

ção está há três dias sem comer’ pode causar *medo, apreensão, culpa, pena*, etc. Mas a força do ato de fala não depende de um conjunto de regras. Depende muito mais do ouvinte e não há garantia de que o ato realmente vá provocar tal ou tal sentimento. Em outras palavras, significa que não há uma lógica dos atos de fala perlocutórios, pois é um fato indeterminado que tipo de *causa* o ato terá sobre o ouvinte. Grosso modo, o argumento de Hare consiste no seguinte: os emotivistas tentavam explicar o significado dos juízos morais com base no seu elemento causativo (Stevenson, por exemplo); nos termos de Austin, no elemento perlocutório; mas, a força perlocutória não pode ser regrada ou passível de uma sistematização lógica; o que vigora no campo do efeito do ato de fala no ouvinte é a pura indeterminação, irracionalidade; e se os emotivistas acham que parte do significados dos juízos morais consiste em sua força perlocutória, então eles admitem que os juízos morais implicam em irracionalidade, supunha Hare.

<sup>58</sup> HARE, E. J. *God and Morality: a Philosophical History*. Oxford, Blackwell Publishing, 2007, p. 209.

<sup>59</sup> Cf. WITTGENSTEIN, L. *Investigações Filosóficas*. Coleção Os Pensadores. São Paulo: Abril Cultural, 1998.

<sup>60</sup> Estou falando em teoria do raciocínio prático e utilitarismo kantiano somente para falar do o objetivo geral de Hare. Não nos ocuparemos dessas coisas aqui. Isso porque trata-se de teoria normativa e o nosso foco é metaético. Para mais detalhes sobre a teoria normativa de Hare, ver HARE, R. M. *Freedom and Reason*. Oxford: Oxford Clarendon Press, 1963 e HARE, R. M. *Moral Thinking, it levels, method and point*. Oxford: Oxford Clarendon Press, 1981.

sua teoria, quer dizer, a parte metaética ou o estudo da lógica da linguagem moral. Hare argumenta, que esse estudo nos revela que os termos morais não se referem a propriedades naturais, ou não podem ser definidos em termos de um conjunto de características naturais. Se isso é verdadeiro, então a própria linguagem moral nos mostra que o naturalismo é falso e, mais do que isso, que é *definitivamente* falso, pois as características lógicas são supostamente intrínsecas à linguagem moral. O prescritivismo, nesse sentido, apresenta uma série de importantes objeções ao naturalismo moral. A partir de agora, gostaria de considerar essas objeções e as principais teses do prescritivismo universal.

Podemos começar por aquilo que se costuma chamar de “Regra de Hare”. Esta “regra” se refere à suposta impossibilidade de se derivar um *dever ser* do *ser* e, obviamente, tem inspiração em David Hume. No *Tratado*, Hume chamou atenção para um modo de raciocinar que partia de premissas sobre coisas que *são* e concluía coisas sobre o que *deve ser*. Havia um passo do *é* para o *deve* que normalmente era suprimido, mas que deveria ser explicado, segundo Hume. Sua clássica formulação é a seguinte.

Não posso deixar de adicionar a esses raciocínios e observações o que pode, talvez, ser de alguma importância. Em cada sistema de moralidade que tenho encontrado até aqui, tenho sempre observado que o autor procede por algum tempo no modo ordinário de se raciocinar e estabelece a existência de Deus, ou faz observações concernentes aos estados de coisas humanos; depois, me surpreendo ao encontrar repentinamente, ao invés de proposições com as cópulas usuais *é* e *não é*, proposições que estão conectadas com *deve* e *não deve*. Esta mudança é imperceptível; mas é, contudo, plena de consequências. Isso porque com este *deve* ou *não deve* expressa uma nova relação ou afirmação, e é necessário que ela seja observada e explicada. E, ao mesmo tempo, uma razão deve ser dada, pois parece totalmente inconcebível como esta nova relação pode ser derivada de outras que são inteiramente diferentes dela<sup>61</sup>.

---

<sup>61</sup> HUME, D. *A Treatise of Human Nature*. Oxford: Clarendon Press, 1978, p. 469.

Hume apenas apontou essa dificuldade das teorias éticas. Hare, a desenvolveu mais precisamente como objeção ao naturalismo. Como vimos, ele distingue entre imperativos e indicativos. A linguagem moral, que está no nível do *dever*, implica em imperativos<sup>62</sup>. Indicativos apenas nos informam fatos sobre o mundo, ou, mais precisamente, nos dizem o que *é* o caso<sup>63</sup>. Nesse sentido, o problema é pensar em que modo deve estar a conclusão de um argumento, dadas premissas imperativas e indicativas. São pressuposições básicas da lógica que, para termos um argumento dedutivo válido, nada pode aparecer na conclusão que não seja dito, ao menos implicitamente, nas premissas. Isso significa que não se pode extrair conclusões imperativas a partir de um conjunto de sentenças indicativas apenas. Desse modo, temos a “Regra de Hare”:

*Regra de Hare:* Nenhuma conclusão imperativa pode ser validamente extraída de um conjunto de premissas que não contenha pelo menos um imperativo<sup>64</sup>.

Como sabemos, o naturalista assume que juízos morais são descrições de fatos da realidade. Então, ao dizer que ‘*x* é correto’, por exemplo, ele não está dizendo algo como ‘Faça *x*’. Está apenas enunciando indicativos, nos termos de Hare. Pare Hare, o assentimento a imperativos implicam em fazer algo enquanto que assentimento a indicativos implicam em crer. Ele inclusive transforma essa definição numa formulação tautológica dizendo que é uma tautologia não assentir a um indicativo e não crer assim como é uma tautologia não assentir a um imperativo e não fazer<sup>65</sup>. O reverso também deve ser verdadeiro, quer dizer, é uma tautologia assentir a um imperativo e fazer algo e a um indicativo e crer em algo. Assim, o assentimento a um imperativo seria condição suficiente para garantir com que o indivíduo *faça* algo.

---

<sup>62</sup> Vale lembrar que estou apenas supondo a tese de que juízos morais implicam em imperativos com a finalidade de apresentar o papel que representa na teoria de Hare e as supostas implicações disso. Não apresentei nenhuma razão de Hare ainda para justificar essa tese. Isso será feito quando estiver apresentando a tese da prescritividade. Por hora, apenas assumamos que juízos morais tem significado imperativo.

<sup>63</sup> Cf. HARE, 1952, *Op. Cit.*, p. 5.

<sup>64</sup> *Ibidem*, p. 28.

<sup>65</sup> *Ibidem*, p. 20.

A essa altura já deve estar claro o suposto erro do naturalista. Ao manter que os juízos morais são formulações indicativas somente, ele não pode argumentar, validamente, que devemos *agir* de forma tal e tal. Isso porque assentimento a indicativos implicam apenas em crença, não em ação. Além disso, ele não pode extrair nenhuma conclusão sobre como se deve agir, pois admite apenas premissas indicativas. Em outras palavras, a posição naturalista não capta a suposta característica *motivadora* dos juízos morais<sup>66</sup>. E, se seguirmos Hare, é impossível que capte, pois nenhum imperativo se segue de um conjunto de indicativos somente. Essa seria uma objeção inicial ao naturalismo moral.

Hare não foi influenciado apenas por Hume, mas por Moore também. Tanto é que uma de suas principais objeções ao naturalismo é uma reformulação do argumento da questão aberta. Sua reformulação tem a virtude de expandir a aplicabilidade do argumento não somente para os usos morais de ‘bom’, mas para todo e qualquer uso da palavra valorativa. Vejamos, então, o seu argumento.

O naturalismo pode ser atrativo, à primeira vista, porque ‘bom’ possui uma característica típica das palavras de valor denominada superveniência<sup>67</sup>. Dizer que a propriedade moral é superveniente é dizer pelo menos duas coisas: (i) que depende de algum modo dos fatos ou das propriedades naturais<sup>68</sup> e que (ii) compromete o falante, se disse de um objeto *x* que era bom por possuir as características *a*, *b* e *c*, a dizer de um objeto similar a *x*, que tem as características *a*, *b* e *c*, que também é bom<sup>69</sup>. O segundo ingrediente não poderá ser explicado até termos apresentado algumas teses prévias, portanto, fiquemos com o primeiro apenas. Hare sustenta que sempre que dizemos que algo é bom temos em mente algumas peculiaridades deste algo que *fazem* com que seja bom. Quando dizemos que ‘*x* é um bom jogador de futebol’ deve haver características fatuais de *x* que são a *razão* de o denominarmos um bom *x*, tais como ‘tem habilidade para driblar’, ‘chuta com as duas pernas’,

---

<sup>66</sup> Este ponto requer um desenvolvimento mais detalhado. Por isso, voltaremos a ele adiante.

<sup>67</sup> Cf. *Ibidem*, p. 80s.

<sup>68</sup> Cf. *Ibidem*, p. 80s, 130s.

<sup>69</sup> Cf. *Ibidem*, p. 145. “Primeiramente, tomemos a característica de ‘bom’ que foi denominada superveniência. Suponha que digamos ‘São Francisco foi um homem bom’. É logicamente impossível dizer isso e sustentar ao mesmo tempo que pode ter existido outro homem, colocado precisamente nas mesmas circunstâncias de São Francisco, que se comportou exatamente da mesma maneira, mas que diferia de São Francisco somente neste aspecto: não era um bom homem”.

‘tem velocidade  $y$ ’, ‘possui resistência física’, e assim por diante. Se alguém perguntar por que o denominamos bom, certamente iremos citar algumas dessas características *naturais*. Segundo Hare, é fácil cometer o erro de achar que essas características a partir das quais a propriedade valorativa supervêm, são, na verdade, sua definição e que juntas implicam em algo ser bom ou não. É justamente isso que faz o naturalista, ou seja, ele tenta fornecer uma definição de bom em termos de um conjunto de propriedades naturais que são a sua definição. Hare interpreta a posição do naturalista como Moore interpretava<sup>70</sup>. Quer dizer, acha que quando o naturalista diz que ‘ $x$  é bom’ e ‘ $x$  é  $N$ ’ (sendo ‘ $N$ ’ o conjunto de propriedades naturais que definem ‘bom’) ele está querendo significar exatamente o mesmo. Em outras palavras, que propriedades morais e naturais são analiticamente equivalentes<sup>71</sup>.

Assim como Moore, o ponto de Hare consiste em argumentar que ‘ $x$  é bom’ e ‘ $x$  é  $N$ ’ não são analiticamente equivalentes. Em termos simples, o seu argumento é o seguinte.

P1. Se o naturalismo é verdadeiro, então ‘ $x$  é bom’ é analiticamente equivalente a ‘ $x$  é  $N$ ’.

P2. ‘ $x$  é bom’ e ‘ $x$  é  $N$ ’ não são analiticamente equivalentes.

C. Então, o naturalismo não é verdadeiro.

O argumento é um *modus tollens*. Portanto, válido. O ponto é saber se a segunda premissa é verdadeira ou não. As razões de Hare para sustentá-la se baseiam no uso ordinário da linguagem. Ele sustenta que se o naturalismo é verdadeiro, então somos impedidos de dizer certas coisas que dizemos significativamente na linguagem comum. Suponhamos que o naturalista nos proponha que ‘ $x$  é bom jogador de futebol’ significa o mesmo que ‘ $x$  é habilidoso para driblar, chuta com as duas pernas, tem velocidade tal e possui resistência física’. Este seria o tipo de definição naturalista. E, se propriedades morais e naturais são analiticamente equivalentes, então dizer que ‘ $x$  é habilidoso para driblar, chuta com as duas pernas, tem velocidade tal e possui resistência física’ realmente é o mesmo que dizer que ‘ $x$  é bom’.

---

<sup>70</sup> Cf. *Ibidem*, p. 82, 86.

<sup>71</sup> Isso significa que o alvo de Hare é o naturalismo analítico.

Agora, notemos que o prescritivismo é uma teoria expressivista. Isso quer dizer que quando dizemos que algo é bom não estamos apenas citando características deste algo, mas estamos nos comprometendo com algo a mais. Estamos expressando o nosso endosso em relação a este algo. Nos termos de Hare, estamos *aprovando*<sup>72</sup>. Dizer ‘*x* é bom’ é aprovar *x*; e dizer ‘*y* não é bom’ é desaprovar *y*. De acordo com Hare, essa é a função da palavra ‘bom’ na nossa linguagem ordinária. Quer dizer, usamo-la para aprovar coisas.

Com essa pressuposição em mente, imaginemos que muitas vezes não queremos dizer que um jogador de futebol é bom porque ‘é habilidoso para driblar, chuta com as duas pernas, tem velocidade tal e possui resistência física’. Queremos apenas *aprovar* uma de suas características qualquer. Podemos, por exemplo, querer dizer que o ‘chutar com as duas pernas’ é bom, isto é, ‘*x* é bom porque *x* chuta com as duas pernas’. E de fato, dizemos isso significativamente na linguagem comum. Mas, se assumirmos que o ‘chutar com as duas pernas’ já faz parte da nossa definição de bom jogador de futebol, então dizer que ‘*x* é bom porque *x* chuta com as duas pernas’ é o mesmo que dizer que ‘*x* chuta com as duas pernas porque *x* chuta com as duas pernas’. Isso é uma afirmação analítica. Mas, quando aprovamos um jogador porque ele chuta com as duas pernas não estamos querendo fazer afirmações analíticas a respeito dele. Então, se assumirmos o naturalismo analítico somos obrigados a dizer certas coisas que não queremos dizer. Assim, Hare conclui que essa é uma evidência para a falsidade da posição naturalista.

O sucesso de seu argumento depende fundamentalmente de que ‘bom’ serve para aprovar ao invés de descrever um conjunto de características naturais. Falemos um pouco mais a respeito dessa suposta diferença<sup>73</sup>. Hare admite que as propriedades morais são supervenientes às propriedades naturais. Neste ponto ele acredita que o naturalismo é correto. Mas não concorda que a relação de superveniência caracteriza *exaustivamente* o significado das propriedades morais, como o

---

<sup>72</sup> Cf. *Ibidem*, 127-136 (capítulo 8).

<sup>73</sup> Na verdade, o desenvolvimento do prescritivismo é uma tentativa de caracterizar essa suposta característica de aprovar que é intrínseca a ‘bom’. Neste sentido, se pode considerar que *quase* toda característica que Hare defende ser propriedade das palavras morais, ou de ‘bom’, mais especificamente, é justificação de P2 do seu argumento geral contra o naturalismo. Ou seja, em algum sentido, é a tentativa de mostrar que ‘bom’ possui um significado diferente dos termos naturais.

naturalista defende. Dizer que ‘bom’ é definido em termos de ‘N’ é não captar uma característica fundamental dos juízos morais, isto é, a sua propriedade de aprovar ações, objetos, etc. Assim, ‘bom’ tem um significado para além daquilo que o naturalista reivindica, o significado *prescritivo*. Dado que ‘bom’ é uma propriedade superveniente às propriedades naturais, sempre que dizemos de um objeto ou ação que é ‘bom’ deve haver um conjunto de propriedades naturais que são a *razão* de o denominarmos bom. Se dissermos de um automóvel que é bom, deve haver um conjunto de fatos que o *tornam* bom, tais como, ‘estabilidade nas curvas’, ‘baixo consumo de combustível’, ‘potente’, etc. Essas características descritivas constituem o *padrão*<sup>74</sup> pelo qual se classifica os objetos como bons membros de sua classe. Dizemos que ‘ser habilidoso para driblar, chutar com as duas pernas, ter velocidade tal e possuir resistência física’ constitui o padrão pelo qual um jogador de futebol é bom.

As propriedades morais, nesse sentido, não possuem apenas significado descritivo, como o naturalista supõe, mas um significado prescritivo também. Quando dizemos que algo é bom estamos, além de especificar o padrão pelo qual o denominamos bom (aspecto descritivo), o endossando ou aprovando (aspecto prescritivo). Segundo Hare, o erro do naturalista é não reconhecer a legitimidade desta última característica. Disso se segue uma distinção importante entre *significado* e *critérios*<sup>75</sup>. Ao denominarmos as coisas boas nós as denominamos por um número indefinido de outras coisas. Dizemos que um jogador de futebol é bom porque é habilidoso para driblar, chuta com as duas pernas, tem velocidade tal e possui resistência física; que um automóvel é bom porque é estável nas curvas, econômico e potente; que um livro de filosofia é bom porque seus argumentos são inovadores, é escrito em uma linguagem bastante simples, é claro e rigoroso; e assim por diante. Quer dizer, denominamos diferentes coisas boas com base em diferentes padrões de bondade, em diferentes *critérios*. Se assentimos que para um taco de bilhar ser um bom taco de bilhar ele tem que ser feito de tal madeira, então o ser feito de tal madeira é o critério pelo qual o denominamos um bom taco de bilhar e constitui, ao mesmo tempo, o padrão para bons tacos de bilhar. Os critérios para diferentes objetos ou ações sempre serão diferentes, obviamente. Aprovamos boa comida, bom jogo, bom salto, boa casa, bom abridor de latas e bom martelo por critérios muito diferentes. No entanto, ao denominarmos os objetos bons

---

<sup>74</sup> Cf. *Ibidem*, p. 111-139, 145-147.

<sup>75</sup> Cf. *Ibidem*, 94-110 (capítulo 6).

estaremos *sempre* aprovando. O aprovar é o *significado* dos termos morais. Assim, a crítica de Hare ao naturalista é que ele só percebe os critérios ou padrões dos juízos avaliativos, mas não seu verdadeiro significado.

Para Hare, o significado avaliativo, ou prescritivo, é primário. E isso por dois motivos. Em primeiro lugar, o significado avaliativo é sempre constante<sup>76</sup>. Aprovamos bons tacos de bilhar, bons automóveis, bons jogadores de futebol, e assim por diante. Todos eles por critérios diferentes. Mas o significado avaliativo é sempre o mesmo em todas as classes de objetos. Em segundo lugar, porque o significado avaliativo pode, muitas vezes, ser usado até mesmo para modificar o significado descritivo<sup>77</sup>. À medida que o tempo passa, pode acontecer de irmos denominando bom um mesmo objeto por qualidades descritivas diferentes, de modo que quando dizemos que ‘*x* é bom’ em *t* estamos querendo dizer que é bom por possuir tais e tais características descritivas, mas quando dizemos que ‘*x* é bom’ em *t*’ pode ser que já não estamos mais nos referindo às mesmas qualidades descritivas. Isso significa que o padrão mudou, mas o significado avaliativo, além de ter sido usado para mudar o padrão, continua sendo o mesmo<sup>78</sup>.

Agora, quero lembrar de uma característica da superveniência que fora apenas mencionada acima. Afirmei que dizer que as propriedades morais são supervenientes é dizer duas coisas: que dependem, de algum modo, dos fatos ou das propriedades naturais e que compromete o falante, se disse de um objeto *x* que era bom por possuir as características *a*, *b* e *c*, a dizer de um objeto similar a *x*, que tem as características que também é bom. Esta segunda característica é a mais importante. Sempre que dizemos que um objeto ou ação é bom (boa)

---

<sup>76</sup> Cf. *Ibidem*, p. 118.

<sup>77</sup> Cf. *Ibidem*, p. 119.

<sup>78</sup> Há algumas exceções. Há casos em que usamos ‘bom’ ou alguma outra palavra avaliativa com significado descritivo primário e o avaliativo apenas secundário. Hare nos dá o exemplo de ‘caprichoso’. Quando aprovamos alguém por ser caprichoso estamos muito mais comunicando informações sobre este alguém do que o aprovando, embora o estejamos aprovando. Podemos querer dizer, por exemplo, que ele cuida muito bem dos seus livros, mantém o seu quarto limpo e organizado, etc. A explicação de Hare para esses casos em que o significado avaliativo não é primário é que o padrão de aprovação se tornou demasiadamente fixo. Assim, quando alguém aprova um objeto ou ação os critérios para a sua aprovação são tão difundidos que as pessoas tendem muito mais a tomar a aprovação como uma descrição das características do objeto do que como uma avaliação.

obviamente há um conjunto de razões que os tornam bons aos nossos olhos. Não podemos dizer que algo é bom sem que este algo tenha algumas dessas características. Então, é condição necessária para algo ser bom o ter características descritivas. Suponha que avaliamos um objeto  $x$  como bom porque possui as características naturais  $a$ ,  $b$  e  $c$ . Agora, suponha também que encontramos um objeto  $y$ , que é da mesma classe de objetos de  $x$ , é similar a  $x$  e que também tem as características  $a$ ,  $b$  e  $c$ . Poderíamos dizer, ao mesmo tempo, que  $x$  é bom, mas que  $y$  não é? A resposta é ‘não’. Mais do que isso. Segundo Hare, é um ‘não’ lógico<sup>79</sup>. Se  $x$  é denominado bom tendo como razão para isso as características  $a$ ,  $b$  e  $c$  e  $y$  é similar a  $x$ , então  $y$  também deve ter as características  $a$ ,  $b$  e  $c$ ; e, se bom é uma propriedade superveniente, então, para dizer que  $y$  não é bom,  $y$  deveria ter características descritivas distintas de  $a$ ,  $b$  e  $c$ . Isso significa que para dizer que um é bom e que outro não é, deve haver diferenças naquilo que os torna bons. Então, para dois juízos morais serem diferentes deve haver diferenças descritivas a partir das quais a propriedade moral supervem e, se não há diferenças descritivas (seja entre dois objetos ou duas ações), então o juízo moral deve ser *necessariamente* o mesmo. Esta é a tese da universalizabilidade. Nos termos de Hare

[...] todos os juízos de valor são veladamente de caráter universal, o que é o mesmo que dizer que se referem a, e exprimem a aceitação de um padrão aplicável a outras ocasiões similares. [...] Quando aprovamos um objeto, nosso juízo não é unicamente sobre aquele objeto particular, mas, inevitavelmente, sobre todos os objetos semelhantes a ele<sup>80</sup>.

Quando digo que ‘ $x$  é bom’ expresso um padrão de bondade para objetos da classe  $x$ ; se  $y$  é similar a  $x$ , então deve compartilhar desse padrão de bondade. Vejamos que é *impossível*, sem ferir o princípio de autocontradição, dizer que, embora  $x$  e  $y$  sejam similares, um é *bom* e outro não é, pois estamos afirmando, ao mesmo tempo, o seguinte: ‘ $x$  possui  $a$ ,  $b$  e  $c$ , mas  $y$ , que é similar a  $x$ , e, portanto, também possui  $a$ ,  $b$  e  $c$ , não possui  $a$ ,  $b$  e  $c$ ’. A tese da universalizabilidade, então, implica que quando faço um juízo moral me comprometo a aceitar o mesmo

---

<sup>79</sup> Cf. *Ibidem*, p. 80s.

<sup>80</sup> *Ibidem*, p. 137.

juízo moral em todas as circunstâncias *similares*, mesmo que os agentes em questão tenham seus papéis invertidos.

A tese da prescritividade pode ser explicada dizendo que os juízos morais tem a propriedade de implicar em imperativos<sup>81</sup>. Em *Moral Thinking* Hare apresenta a seguinte definição.

Dizemos algo prescritivo se, e somente se, para algum ato *A*, uma situação *S* e uma pessoa *P*, se *P* assentiu (oralmente) ao que dissemos e não fez *A* em *S*, ele logicamente estaria assentindo não sinceramente<sup>82</sup>.

Isso significa que assentir sinceramente a um juízo moral é, ao mesmo tempo, assentir ao imperativo derivado do juízo moral. Se o juízo afirma que o ‘não mentir’ é um dever, e assentimos a este juízo, então aceitamos o imperativo particular ‘Não minta!’ aplicado a nós mesmos.

Mas essa é apenas a definição da prescritividade. O que devemos procurar saber é que tipo de argumento Hare apresenta a seu favor. Isso é um pouco difícil. Mas, vejamos o que ele diz.

É então possível tratar ‘Devo fazer X’ como uma mistura confusa de três juízos.

- (1) ‘X é necessário para a conformidade ao padrão que as pessoas geralmente aceitam’ (afirmação de fato sociológico);
- (2) ‘Tenho um sentimento de que devo fazer X’ (afirmação de fato psicológico);
- (3) ‘Devo fazer X’ (juízo de valor).

[...] sempre se descobrirá, após investigação, que casos que se alega serem juízos de valor que não implicam em imperativos são casos em que o que é significado não é do tipo (3) acima, mas dos tipos (1) ou (2) ou uma mistura de ambos. Essa asserção é, certamente, impossível de provar ou mesmo de tornar plausível, a menos que saibamos quando devemos considerar um juízos como sendo do tipo (3); mas proponho superar essa dificuldade do único modo possível,

---

<sup>81</sup> Cf. *Ibidem*, p. 165-170 e HARE, 1963, *Op. Cit.*, p. 26-29.

<sup>82</sup> HARE, 1981, *Op. Cit.*, p. 21.

transformando-a numa *questão de definição*. Proponho dizer que o teste para verificar se alguém está ou não usando o juízo ‘Devo fazer X’ como um juízo de valor é, ‘Ele reconhece ou não reconhece que se assentir ao juízo ele deve também assentir ao comando “Que eu faça X”?’<sup>83</sup>.

Proponho que interpretemos essa citação como o seguinte argumento. Como vemos, Hare afirma que, se ‘Devo fazer *x*’ não implicar em um imperativo, então se trata das sentenças (1) e (2), que são afirmações de fato sociológico e psicológico, quer dizer, afirmações descritivas ou indicativas. Como vimos acima a respeito da diferença entre imperativos e indicativos, assentimento a indicativos implicam apenas em crença e não *levam à ação*. Somente os imperativos tem esta propriedade. Além disso, de acordo com a “Regra de Hare”, não se pode derivar um imperativo validamente a partir de indicativos somente. Portanto, podemos concluir que se ‘Devo fazer *x*’ se tratar de (1) e (2), então não tem a propriedade de *levar à ação*.

Ora, é uma ideia primária do prescritivismo que a moralidade trata de fornecer respostas à pergunta ‘Que devemos fazer?’<sup>84</sup>. Uma resposta a um problema moral, seja qual for, sempre será em termos de guiar a conduta em algum sentido fornecendo uma resposta do tipo ‘Façamos tal e tal coisa’. Então, Hare assume que os juízos morais tem a propriedade de guiar a conduta. Assim, ‘Devo fazer *x*’ não pode ser analisado em termos de fatos psicológicos ou sociológicos, ou, em uma palavra, em termos indicativos. Temos, então, que: se ‘Devo fazer *x*’ não implicar em imperativos, então pode ser analisado em termos de indicativos; indicativos não guiam a conduta; ‘Devo fazer *x*’ guia a conduta. Lembremos, agora, que a definição de uma sentença imperativa é que o assentimento sincero a ela implica em fazer algo, leva à ação. Assim, somente o imperativo parece satisfazer a característica dos juízos de dever de guiar a conduta. Então, Hare conclui, juízos de dever implicam em imperativos. Parece ser por isso que diz que para algo ser um juízo de dever tem de implicar num imperativo.

O argumento parece ser um tipo de *reductio ad absurdum*. Hare supõe que juízos morais não implicam em imperativos; conclui que se

---

<sup>83</sup> HARE, 1952, *Op. Cit.*, p. 167-169. Itálico adicionado.

<sup>84</sup> Cf. *Ibidem*, p. 56-78. (capítulo 4).

isso é o caso, então não satisfaz as características da moralidade e que estas só são satisfeitas se admitirmos que implicam em imperativos; assim, conclui que a suposição deve ser falsa e que juízos morais devem implicar em imperativos<sup>85</sup>.

Notemos que se é verdadeiro que juízos morais implicam em imperativos e se é verdadeiro que assentir sinceramente a um imperativo implica em fazer algo, então assentir sinceramente a um juízo moral é, ao mesmo tempo, assentir a fazer o que é prescrito pelo juízo moral. Essa é a ideia da prescritividade. E, juntamente com a tese da universalizabilidade, quer dizer que assentir sinceramente a um juízo moral é assentir, ao mesmo tempo, a fazer o que é prescrito pelo juízo moral *para todas as ocasiões similares*. Por isso, o prescritivismo é considerado uma teoria internalista. Sustentar o internalismo é sustentar que é característica intrínseca, ou, mais precisamente, *conceitual*, dos juízos morais o motivar o indivíduo a agir de acordo com os juízos morais<sup>86</sup>. Se  $x$  assente ao ‘O assassinato é errado’, então isso é condição suficiente para  $x$  estar *motivado* a não cometer o assassinato, ou a, pelo menos, expressar sinceramente que o assassinato é errado.

Agora, podemos compreender o desacordo de Hare com o naturalismo. Ao não reconhecer a característica prescritiva, o naturalista não reconhece uma característica fundamental dos juízos morais, a propriedade de motivar à ação e guiar a conduta. Ao sustentar que juízos morais são representações de fatos externos ao sujeito, o naturalista está fazendo afirmações fatuais somente, e afirmações fatuais não levam à ação. Para Hare, as propriedades morais também tem significado descritivo, mas a característica principal é o significado prescritivo, que

---

<sup>85</sup> Disse que é um pouco difícil saber exatamente sobre o argumento de Hare porque se pode dizer, como fez Warnock (Cf. WARNOCK, M. *Ethics Since 1900*, p. 128-129), que há uma ambiguidade no raciocínio de Hare. Pode-se dizer que ele favorece a sua posição ao tomar como uma *questão de definição* que para algo ser um juízo moral então deve implicar em um imperativo. Se poderia dizer talvez que ele não tem um argumento a favor da prescritividade, mas que já parte da definição de que juízos morais são prescritivos. Assim, Hare estaria favorecendo a sua posição, pois toda vez que se tentasse apresentar um suposto contraexemplo de um juízo de dever que não implica em imperativo, Hare diria que não é um juízo de dever. Mas essas são complicações que não precisamos entrar aqui. Nosso ponto é apenas notar o que se segue se os juízos morais forem realmente prescritivos.

<sup>86</sup> Cf. HARE, R. M. *Internalism and Externalism in Ethics*. In: HARE, R. M. *Objective Prescriptions and Other Essays*. Clarendon Press: Oxford, 1999, p. 96.

é primário e constante. Essa característica é revelada pela análise do uso dos termos morais. Usamos a linguagem moral para direcionar cursos de ação e esta é uma característica inerente à própria linguagem. O naturalista, para Hare, não compreendeu o funcionamento da linguagem moral.

Em desenvolvimentos posteriores ao *The Language of Morals*, Hare apresentou outro tipo de argumento contra o naturalismo. A pressuposição principal continua sendo a mesma; o naturalista reconhece apenas parte daquilo que constitui o significado dos juízos morais. Não reconhece o significado prescritivo. Mas agora Hare acusa o naturalista de ser relativista. Se isso é verdadeiro, pode até soar um pouco irônico, pois é justamente problemas como o relativismo que o naturalista quer evitar e que reivindica como sendo uma virtude da sua posição.

O argumento de Hare depende de algumas caracterizações prévias. Em *Sorting out Ethics* ele faz uma taxonomia das teorias éticas (metaéticas). Ele divide as teorias éticas em duas classes gerais sendo que cada uma tem uma subdivisão particular. Às duas classes gerais, Hare denomina teorias *descritivistas* e teorias *não descritivistas*<sup>87</sup>. A diferença entre ambas depende da ideia de condições de verdade. Normalmente se diz que para um juízo ser verdadeiro há a necessidade de que ele satisfaça algumas condições. Assim, para ‘A casa tem um buraco no teto’ ser verdadeira é preciso que haja um objeto que instancie a propriedade de ser uma casa e que, além disso, tenha a propriedade de ter um buraco no teto. Se isso é satisfeito, então a sentença será verdadeira. Senão, não. O haver uma casa e o ter o buraco no teto são as condições de verdade para a sentença acima.

Um descritivista é alguém que sustenta que o significado dos juízos morais é determinado *inteiramente* pelas condições de verdade<sup>88</sup>. Um juízo moral é verdadeiro ou falso, e há condições através das quais

---

<sup>87</sup> Cf. HARE, 1997, *Op. Cit.*, p. 47.

<sup>88</sup> Cf. *Ibidem*, p. 48. É necessário fazer uma observação sobre o ‘inteiramente’. Há casos em que podemos saber que tal sentença possui um significado diferente de tal sentença sem precisarmos nos adiantar até as condições de verdade. Sabemos que ‘Você vai fechar a porta’ tem um significado diferente de ‘Feche a porta’ simplesmente com base na sintaxe. Então, o significado não é determinado inteiramente, no sentido estrito, pelas condições de verdade. Mas, para os juízos morais, se pode deixar de lado esses pormenores e caracterizar suficientemente a diferença entre descritivismo e não descritivismo utilizando a expressão ‘inteiramente’.

podemos atestar isso. Como parte das teorias descritivistas, Hare elenca o naturalismo e o intuicionismo. Ambas as teorias mantêm que a verdade de um juízo moral depende da posse de certas propriedades que são as condições de verdade. Para um naturalista, por exemplo, as propriedades morais pertencem ao mundo natural e se dizemos que ‘Toda violência é injusta’ podemos saber sobre a verdade ou falsidade deste juízo analisando se a violência tem a propriedade da injustiça, que pode ser definida em termos naturais (causar dor, causar sofrimento, causar desconforto psicológico, etc.). Um intuicionista defende que a verdade de um juízo moral depende da posse de certas propriedades também, mas de propriedades não naturais. Essas propriedades são comumente denominadas *sui generis* e diz-se que temos acesso a elas por alguma forma de intuição. Segundo Hare, a diferença entre intuicionismo e naturalismo é que o naturalista pensa que as condições de verdade dos enunciados morais podem ser especificadas sem se fazer uso dos termos morais enquanto que o intuicionista não descarta os termos morais para especificar as condições de verdade<sup>89</sup>.

O não descritivista, por outro lado, é alguém que sustenta que o significado dos juízos morais não é determinado inteiramente pelas condições de verdade. Ele reconhece que há condições de verdade, mas sustenta que há algo a mais. O emotivismo e o racionalismo fazem parte das teorias não descritivistas. Apesar de Hare não ser um emotivista, ele acredita que ao terem dito que o significado dos juízos morais não depende inteiramente das condições de verdade, os emotivistas tiveram um grande acerto. Como sabemos, o emotivista mantêm que quando se afirma que determinado fato ou pessoa tem uma propriedade moral não se está apenas observando um fato a respeito daquela pessoa ou ação, mas se está projetando as emoções de aprovação ou desaprovação. Assim, o significado dos juízos morais não depende apenas de características externas ao sujeito. É por isso que um não descritivista não se compromete com a verdade ou falsidade dos juízos morais. O racionalismo nada mais é do que o prescritivismo universal. Para um racionalista o significado dos juízos morais é determinado por dois componentes, o significado descritivo e o prescritivo. Assim, os juízos morais dependem em certo sentido das condições de verdade (um juízo moral sempre especifica um padrão de avaliação, como vimos, que é descritivo), mas tem um significado para além disso. Para usar um exemplo do próprio Hare, ‘x é uma boa pessoa’ pode significar que ‘x é honesta’, ‘x não trapaceia em jogos de cartas’, etc., que são propriedades

---

<sup>89</sup> Cf. *Ibidem*, p. 64s.

descritivas, mas, para além disso, o enunciado tem um significado prescritivo que é primário<sup>90</sup>. Em poucas palavras, essa é a taxonomia elaborada por Hare.

Tendo compreendido isso, fica fácil entender a sua objeção ao naturalismo. Dado que o naturalista sustenta que o significado dos juízos morais é determinado apenas pelas condições de verdade, então saber que um juízo é verdadeiro é saber se o termo moral ('certo', 'errado', 'bom') está sendo corretamente aplicado à propriedade em questão. O ponto é saber como se descobre a que tipo de propriedades o termo moral é corretamente aplicado. Hare sugere que, assim como nos casos do significado de termos não morais, deve-se examinar o uso linguístico dos falantes nativos de um língua, por exemplo<sup>91</sup>. O naturalista poderia dizer que um juízo moral é verdadeiro, quer dizer, o predicado moral 'bom' está sendo corretamente aplicado, se os falantes nativos do português o aplicam à propriedade que está agora sendo aplicado. Além disso, o naturalista pode dizer que se acordamos que o significado de determinada propriedade moral é definido por tais e tais propriedades naturais, então as ações que tiverem tais e tais propriedades morais serão boas e as que não tiverem não serão.

O problema é que não é um ponto de muito acordo o saber a que tipo de propriedades naturais as propriedades morais se aplicam<sup>92</sup>. Até mesmo os falantes da mesma língua aplicam os termos morais a propriedades muito diferentes. Os falantes do português de uma região  $x$  podem acreditar que 'errado' não se aplica ao aborto. Os falantes do português de  $y$  podem acreditar que se aplica. E pode ser que, mesmo nos limites de  $x$  ou  $y$  somente, há pessoas que acreditam que 'errado' se aplica a comer carne animal e há pessoas que acreditam que não se aplica. Agora, imaginemos esse tipo de desacordo numa escala maior. Imaginemos os termos morais em geral e a grande variedade cultural que há. É certo que haverá muitos desacordos sobre a correta aplicação dos termos morais e não encontraremos uma regra geral que determine qual é a correta de forma definitiva<sup>93</sup>. Isso significa que haverá desacordo sobre que tipos de propriedades naturais são a definição das propriedades morais. Para considerarmos um exemplo não moral, há muitos *padrões* diferentes para ' $x$  é um bom carro'. Pode-se dizer que esse enunciado significa ' $x$  é veloz e potente', que ' $x$  é leve e

---

<sup>90</sup> Hare, nesse sentido, é um racionalista não descritivista.

<sup>91</sup> Cf. *Ibidem*, p. 66.

<sup>92</sup> Cf. *Ibidem*, p. 69.

<sup>93</sup> Cf. *Ibidem*, p. 67-68.

econômico', que 'x é antigo e com tal design', e assim por diante. O ponto é que não há um critério objetivo para determinarmos qual é o padrão correto para o uso dos termos morais. Se o naturalismo se compromete com isso, como Hare afirma que se compromete, então é certo que implica em relativismo. Diferentes culturas irão reivindicar padrões diferentes para a aplicação dos predicados morais. E, dado que tudo o que determina o significado dos juízos morais são as condições de verdade, não teremos um meio objetivo para tentarmos resolver o problema de quem está certo. Esse é o problema com o naturalismo.

Somente um não descritivista dá conta desse problema, para Hare. Isso porque, ele sustenta que para além das condições de verdade há um outro elemento que constitui o significado dos juízos morais. E este elemento supostamente está presente em qualquer juízo moral, mesmo que sejam incompatíveis. Estamos falando da prescritividade. A maneira mais adequada de fugir ao relativismo é sustentar o racionalismo. Pois, mesmo que as culturas apresentem condições de verdade distintas, elas concordam sobre algo fundamental. Ou seja, concordam que os juízos morais tem significado prescritivo, ou avaliativo. Assim, se uma cultura disser que 'O assassinato na guerra é justo' e outra disser que 'O assassinato na guerra não é justo', por mais que estejam em desacordo diante do padrão de aplicação do termo 'justo', elas concordam sobre uma coisa, que o juízo moral tem um significado prescritivo. Este é o ponto comum. Tendo estabelecido que há um ponto comum, o problema consiste agora em saber quem está certo. Mas isso é papel da teoria normativa que supostamente é derivada do prescritivismo universal, coisa que não trataremos aqui. O que é importante, é observar que no caso do racionalismo, supostamente há um modo de resolver o problema do desacordo dado que há uma base comum. Mas no caso do naturalismo, parece não haver como resolver o problema, pois o significado dos juízos morais é exaustivamente caracterizado pelas condições de verdade.

Para concluir, relembremos as principais coisas que vimos nessa seção. Comecei esboçando algumas características principais do emotivismo moral. Apresentei o critério de verificabilidade e disse em que sentido ele determina que sentenças morais não tem significado. A partir disso, apresentei a ideia de significado emotivo em Ayer e Stevenson. Depois, afirmei que o emotivismo supostamente implica no problema da não racionalidade da atividade moral e disse que foi por isso que Hare, apesar de concordar em algumas coisas, rompeu com essa teoria. Vimos a distinção de Hare entre sentenças imperativas e indicativas. Seu argumento a favor da racionalidade da ética, que

consiste em distinguir entre duas partes constituintes de significado de uma sentença, a frástica e a nêustica, e mostrar que tanto imperativos quanto indicativos compartilham das mesmas frásticas para, daí, mostrar que as relações lógicas ocorrem sempre nas frásticas; portanto, em imperativos também, ao contrário do que sustentavam os emotivistas. Depois disso, apresentei algumas das ideias de Hare de *The Language of Morals*. A sua reformulação do argumento da questão aberta e, a partir disso, sua análise da linguagem moral a partir do uso da linguagem ordinária. Disse que, para Hare, a linguagem moral tem algumas características que a linguagem não moral não tem, sendo a principal delas a prescritividade. A prescritividade é a ideia de que juízos morais implicam em imperativos e, dado que assentimento sincero a imperativos implica em algum tipo de ação, assentimento sincero aos juízos morais implica em agir de acordo com os juízos morais. Essa é, dentre outras, uma das principais objeções de Hare ao naturalismo, pois o naturalista supostamente não reconhece a característica das prescritividade dos juízos morais.

A partir de agora, gostaria que considerássemos outra objeção ao naturalismo moral, a ideia de que não *há* propriedades morais.

#### 4. Mackie e a teoria do erro

Para começar, imagine que você seja um cristão. Você segue a maioria dos valores comumente endossados por tal religião, faz orações, seja para agradecer a Deus por certos acontecimentos que acredita terem sido influenciados por ele seja para pedir que certas coisas aconteçam ou deixem de acontecer, vai à igreja toda semana e, à medida do possível, tenta divulgar a religião para as outras pessoas. Você acredita em milagres, acredita que a oração pode alterar o curso natural dos eventos, acredita que o mundo foi projetado de tal e tal forma, acredita que seu corpo possui uma alma e acredita que a vida na Terra é apenas “uma parte” e que depois da morte continuará existindo em algum outro lugar. Você tem todas essas crenças, e isso sem contar as inúmeras crenças morais mais específicas que *guiam a conduta* de um adepto do Cristianismo. Agora, note que todas essas crenças dependem de uma crença mais geral que é, por assim dizer, a crença básica. Você deve acreditar, naturalmente, que há um Deus e que este Deus possui certas propriedades, tais como a onipotência, a onipresença e a onisciência. As outras crenças não fazem sentido se não se tiver esta crença. Há uma coisa importante aqui. Se se conseguir provar que Deus de fato não existe, como o ateísmo tem tentado fazer, então isso tem implicações

para com todas as outras crenças que dependem dessa crença mais geral. Se Deus não existe, então não faz sentido crer que a oração mudará a ordem natural dos acontecimentos, que milagres são possíveis etc. Se Deus não existe, então não há no mundo algo que torne suas crenças verdadeiras. Se Deus não existe, então o ter essas crenças é uma espécie de *erro*.

O teórico do erro está tentando fazer algo muito similar ao ateuísta, mas com a moralidade. Obviamente, há um número indeterminado de crenças morais. ‘O assassinato é errado’, ‘A tortura não é uma prática correta’, ‘Usar animais como fonte de alimento não é certo’, ‘A melhor forma de governo é uma democracia’, ‘Promover a igualdade entre as pessoas é bom’, etc., são alguns exemplos. O modo como as pessoas vivem é diretamente influenciado por crenças desse tipo. Mas há algo mais fundamental que parece ser pressuposto para se garantir a verdade dessas sentenças. A existência de propriedades morais como ‘bom’, ‘correto’, ‘justo’ e ‘errado’. Quando dizemos que ‘Torturar alguém é errado’ estamos querendo dizer que tal prática tem tal propriedade. Se não há coisas como a ‘incorreção’, então parece que o juízo em questão implica em problemas; é *falso*, para ser mais preciso, pois não há algo que corresponda à incorreção. Quer dizer, as crenças morais em algum sentido dependem de algo mais geral e fundamental, que são as propriedades morais. E se for consistentemente mostrado que tais propriedades não existem, então parece que estamos com problemas. Esse é o ponto do teórico do erro. Ele sustenta que, apesar de atribuímos propriedades morais a objetos, ações, pessoas, etc., e falarmos como se essas coisas realmente tivessem essas propriedades, no fundo não há coisas como *bondade*, *incorreção*, e *justiça*. Por isso, as sentenças morais são sistemática e uniformemente falsas. A moralidade é uma espécie de *erro*<sup>94</sup>.

---

<sup>94</sup> Um modo errado de interpretar a ideia do teórico do erro é pensar que as crenças morais derivam de juízos morais mais gerais, como ‘Causar mais dor do que prazer é errado’, ‘O assassinato é errado’ e que, se for possível mostrar que esses juízos mais gerais não são consistentes ou são falsos, então os juízos morais derivados deles não se sustentam. Isso é um erro, pois a moralidade ainda poderia continuar subsistindo e as sentenças morais não seriam falsas mesmo que fosse mostrado que é correto causar mais dor do que prazer ou que o assassinato é a coisa certa a se fazer em qualquer circunstância (por mais estranho que possa parecer). Isso porque as propriedades morais não seriam atingidas. A ideia do teórico do erro é que a consistência dos juízos morais depende da existência das propriedades morais. E, se tais propriedades não

Um dos representantes mais influentes da teoria do erro é J. L. Mackie com o seu *Ethics: Inventing Right and Wrong*<sup>95</sup>. Sua teoria do erro é composta basicamente por duas partes, uma *conceitual* e uma *ontológica*. Na primeira, ele estabelece qual é o conceito de propriedade moral usualmente aceito, seja no âmbito filosófico ou no uso ordinário. Grosso modo, Mackie mantém que trata-se de propriedades prescritivas e categoricamente objetivas. Na segunda parte, ele argumenta que simplesmente não existem propriedades desse tipo. Há pelo menos dois argumentos principais para sustentar isso. O argumento da relatividade e o argumento da estranheza. O argumento da estranheza tem duas partes, uma metafísica e uma epistemológica – e, supostamente, uma terceira, a da superveniência<sup>96</sup>. A ideia básica é que no mundo não há fatos e propriedades morais e, por isso, as sentenças morais não têm como ser verdadeiras. Assim, o discurso moral é sistemática e uniformemente *falso*. Note que se a posição de Mackie nos tem algo a dizer – e parece que tem, senão não teria chamado tanta atenção dos eticistas – então o naturalista tem um sério problema para resolver. O naturalista sustenta justamente o que Mackie nega, a existência de fatos morais. Por isso, a teoria do erro é uma importante objeção ao naturalismo moral.

Mas dizer todas essas coisas não ajuda muito, a menos que se explique o que tudo isso significa. Desse modo, nessa seção irei apresentar a teoria do erro de Mackie explicando qual a sua interpretação das propriedades morais e quais os seus argumentos para sustentar que elas não existem. Começarei pela parte conceitual explicando respectivamente o que Mackie entende por propriedades prescritivas e categoricamente objetivas. Tendo estabelecido seu alvo, poderei passar para os seus argumentos. Primeiro, o argumento da relatividade. Depois, o seu argumento mais famoso, o argumento da estranheza. No final, direi algo sobre a implicação da tese de Mackie para a moralidade comum.

---

existem, como ele tenta mostrar, então as sentenças morais são uniformemente falsas. Este é o erro, este é o problema para a moralidade.

<sup>95</sup> MACKIE, J. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

<sup>96</sup> Esse ponto requer um esclarecimento. O próprio Mackie afirma que o argumento da estranheza é composto por *duas* partes, a parte metafísica e a parte epistemológica. No entanto, ao apresentar o argumento, acrescenta um ingrediente a mais, a relação de superveniência entre propriedades morais e naturais. Não sei até que ponto se poderia tratar isso como uma terceira parte do argumento da estranheza. Mas o importante é que, de alguma forma, é razão para se aceitar o argumento.

Primeiro, uma palavra sobre rótulos. Mackie afirma que podemos interpretar a sua posição como sendo a de um *cético moral* ou de um *subjetivista*<sup>97</sup>. Alguém pode dizer que é um cético a respeito da moralidade *x*; da moralidade cristã ou budista, por exemplo. Não é neste sentido que Mackie diz ser cético. Estas são questões de primeira ordem, sobre *moralidades*. Mackie é cético sobre questões de segunda ordem, sobre *a* moralidade. Sustentar isso significa não acreditar que existam propriedades morais *e*, por conseguinte, que as sentenças que atribuem certas propriedades morais a certas ações são sempre falsas. Cabe dizer que questões de primeira e segunda ordem “não são apenas distintas, mas completamente independentes”<sup>98</sup>. Assim, posso sustentar que os juízos morais são todos *falsos* e que a moralidade implica num *erro*, mas não preciso autorizar, ao mesmo tempo, a violência sexual, a injustiça e o assassinato. Ser um cético moral não significa não agir moralmente. Da mesma forma, alguém pode dizer que é um subjetivista com relação a assuntos de primeira ordem. Pode dizer que o correto é agir da forma que as pessoas “acham” ser o correto ou que o importante é agir segundo a “consciência”, seja lá o que isso signifique. Não é isso que Mackie está dizendo por ‘subjetivista’. Além disso, é possível confundir a posição de Mackie mesmo com assuntos de segunda ordem ao interpretar ‘subjetivista’ como alguém que sustenta que os juízos morais como ‘*X* é errado’ significa que o falante desaprova *x*, ou a comunidade tal desaprova *x* – de fato, geralmente é isso que ‘subjetivista’ significa –. Ou, ainda, que os juízos morais representam a demonstração dos sentimentos ou atitudes de alguém ou de um grupo de pessoas. Não é isso que Mackie está dizendo. Seu ponto não é sobre o que os juízos morais *são*, mas sobre o que *não são*<sup>99</sup>. Ele aceita ser chamado de subjetivista porque sustenta que os juízos morais não são objetivos, como veremos. Se não são objetivos, então devem ser subjetivos em algum sentido<sup>100</sup>.

A parte conceitual da teoria de Mackie é a ideia de que as propriedades morais são prescritivas e categoricamente objetivas. Tivemos uma boa noção do que significa dizer que a moralidade é prescritiva quando tratamos das ideias de Hare. Dizer que juízos morais são prescritivos é dizer que não se pode assentir sinceramente a um juízo moral e se recusar a agir de acordo. Em outras palavras, isso

---

<sup>97</sup> MACKIE, 1977, *Op. Cit.*, p. 15-18.

<sup>98</sup> *Ibidem*, p. 16.

<sup>99</sup> *Ibidem*, p. 17.

<sup>100</sup> *Ibidem*, p. 18.

significa que as propriedades morais são conceitualmente *motivadoras* ou que *guiam a nossa conduta*. Mackie afirma que essa característica dos juízos morais é endossada por boa parte da tradição filosófica.

[...] a principal tradição da filosofia moral Europeia, de Platão em diante, tem combinado a visão de que os valores morais são objetivos com o reconhecimento de que os juízos morais são parcialmente *prescritivos* ou *diretivos* ou *guias-de-ação*. [...] Na teoria platônica das Formas, e em particular a Forma do Bem, as realidades são externas e extra mentais. Elas são o elemento estrutural central da constituição do mundo. Mas, é mantido também que o conhecê-las ou ‘vê-las’ não irá apenas dizer aos homens o que fazer, mas *irá assegurar que eles façam* isso, apensar de qualquer inclinação contrária. [...] Estando conscientes das Formas do Bem, da Justiça e da Beleza e do resto eles vão, apenas por este conhecimento, sem nenhuma motivação além, ser *impelidos a perseguir e a promover* esses ideais. Similarmente, Kant acredita que a razão pura pode por si mesma ser *prática* [...]. Além disso, Sidgwick argumenta que se há uma ciência da ética – e ele assume que pode haver e define a ética como ‘*a ciência da conduta*’ – o que deve ser ‘deve em outro sentido ter existência objetiva’<sup>101</sup>.

A esta lista, podemos adicionar Hume quando, no *Tratado*, argumenta que a razão é inerte, não tem influência sobre as ações, mas apenas as paixões o tem<sup>102</sup>. Podemos adicionar Ayer e Stevenson quando argumentam que os juízos morais são a *expressão* de sentimentos e atitudes<sup>103</sup>. Além disso, o discurso ordinário geralmente também endossa a ideia de que juízos morais são prescritivos. Ficamos indignados quando alguém não fez tal coisa se sabia que era ‘bom’ e acreditamos ser razão suficiente para alguém fazer *x* o fato de *x* ser correto. Se nos pedem para justificarmos porque fizemos *a* e não *b* dizemos que *a* era a coisa mais certa a fazer, e assim por diante.

---

<sup>101</sup> *Ibidem*, p. 23-24. Itálico adicionado.

<sup>102</sup> Cf. HUME, 1978, *Op. Cit.*, p. 457.

<sup>103</sup> Cf. AYER, 1971, *Op. Cit.*, capítulo 6 e STEVENSON, 1958, *Op. Cit.*

Além de serem prescritivos, Mackie afirma que o conceito usual de juízo moral é de que são categóricos. Ele está pensando a partir da distinção kantiana entre imperativos categóricos e imperativos hipotéticos. Na *Fundamentação da Metafísica dos Costumes*, Kant sustenta que “os imperativos mandam ou *hipotética* ou *categoricamente*”. Os imperativos hipotéticos “representam a necessidade prática de uma ação como meio para conseguir uma outra coisa que se quer”. Já o imperativo categórico é “aquele que representaria uma ação como objetivamente necessária por si mesma, sem referência a um outro fim”<sup>104</sup>. Agir com base no imperativo hipotético, por conseguinte, não é agir moralmente, para Kant, pois a motivação não é essencialmente de cumprir o dever, mas de satisfazer algum outro interesse do agente. Posso ajudar arduamente a combater a pobreza mundial participando de ações não governamentais de arrecadação de bens, incentivando as pessoas a doar parte de seus bens para os mais necessitados, etc., com a motivação de que, se continuar a fazer isso, poderei gozar de um posto social que me propicia a satisfação de vários tipos de interesses sociais e econômicos, por exemplo. Este é um tipo e ação baseada no imperativo hipotético, não tem valor moral. O agir com base no imperativo categórico, por outro lado, é completamente dissociado de qualquer interesse subjetivo da parte do agente. Os juízos devem ser incondicionais e, portanto, independentes de qualquer interesse dos agentes. Não posso impor que o assassinato seja correto simplesmente porque *acho* legal tirar a vida das outras pessoas. A ideia de que os juízos morais são categóricos encontra forte apoio na moralidade comum. Temos uma tendência maior a acreditar que, dentre duas ações, uma é mais louvável se foi individualmente desinteressada e por consideração com a outra parte enquanto a outra teve como motivo apenas a satisfação de um interesse do agente. Acreditamos que as ações de Oscar Schindler, que abriu mão de muitos bens e arriscou sua própria vida para salvar judeus durante a Segunda Guerra, são muito mais estimáveis do que as ações de alguém que arriscou sua própria vida pelo bem de outros, que de fato incrementou o bem estar de outros, mas que tinha apenas um interesse egoísta maior cujos meios para satisfazê-lo incluíam ajudar os outros.

A terceira característica dos juízos morais, de acordo com Mackie, é serem objetivos. Aqui, ele novamente cita Platão, Aristóteles, Kant, Sidgwick, Hutcheson, Price, etc., como representantes da tradição

---

<sup>104</sup> KANT, I. *Fundamentação da Metafísica dos Costumes*. Tradução de Guido Antônio de Almeida. São Paulo: Discurso Editorial: Barcarolla, 2009, p. 189.

filosófica que sustenta que juízos morais tem essa característica. Mas, como apontou Miller, é um pouco difícil saber *exatamente* o que Mackie quer dizer por ‘objetivos’, pois ele diz várias coisas diferentes.

Dizer que uma exigência é objetiva é dizer que ela pode ser objeto de *conhecimento* (24, 31, 33), que pode ser *verdadeiro ou falso* (26, 33), que pode ser *percebido* (31, 33), que pode ser *reconhecido* (42), que é *antecedente e independente* das nossas preferências e escolhas (32, 34, 43), *que é parte da constituição do mundo* (12), *que suporta e valida* algumas das nossas preferências e escolhas (22), que é capaz de ser *simplesmente verdadeiro* (30) ou *válido como uma questão de lógica* (30), que *não é constituído pela nossa escolha ou decisão de pensar de um certo modo* (30), que é *extra-mental* (23), que é algo de que podemos estar *conscientes* (38), que é algo passível de *introspecção* (39), que é algo que pode *figurar como uma premissa em uma hipótese explicativa ou inferência* (39), e assim por diante<sup>105</sup>.

Mas mesmo que não apresente uma definição muito precisa de ‘objetivos’ é perfeitamente possível termos uma ideia básica do que Mackie quer dizer. Só pelo fato de citar Platão e Kant já podemos supor que um juízo moral é objetivo se é independente das opiniões, crenças, preferências dos indivíduos a que se aplica. Como sabemos, as Formas são independentes daquilo que pensamos sobre elas, no caso de Platão, e uma ação moral deve ser completamente desprovida de inclinações, no caso de Kant. E, a partir de expressões como ‘antecedente e independente das nossas escolhas e preferências’, ‘que é extra-mental’, ‘que pode ser reconhecido’, ‘que pode ser verdadeiro ou falso’, etc., podemos ter a confirmação disso. O senso comum também sugere que a objetividade é uma característica dos juízos morais. Quando dizemos que algo é errado não estamos meramente querendo dizer que *nos parece* errado e que para outro alguém poderia ser correto, mas que aquele algo realmente tem a propriedade de ser errado. E quando dizemos que alguém tem o dever de fazer algo, não estamos meramente reivindicando que seria vantajoso para ele ou para nós mesmos, mas que

---

<sup>105</sup> MILLER, 2011, *Op. Cit.*, p. 116.

há razões independentes disso, razões supostamente independentes dos desejos.

O naturalismo compartilha principalmente com a ideia de que juízos morais são categóricos e objetivos. Na verdade, a objetividade tem sido uma das principais atrações do naturalismo. Se os juízos morais são, de algum modo, independentes dos sujeitos, então pode haver respostas corretas aos problemas morais, pois algo não pode ser correto para um e incorreto para outrem. O que está em questão não é como os indivíduos se posicionam em relação a tais e tais atos terem tais e tais propriedades morais, mas com tais e tais atos *realmente* terem tais e tais propriedades morais. Ou seja, para o naturalista, determinada ação ou pessoa tem ou não tem a propriedade moral de ser *justa*, *boa*, ou *incorreta* e isso é parte do mundo natural. Ao sujeito basta reconhecer esses fatos.

Essa foi a primeira parte da teoria de Mackie, a parte conceitual. Como vimos, ele sustenta que os juízos morais são frequentemente vistos como prescritivos e categoricamente objetivos. A segunda parte da teoria consiste em argumentar que não existe nenhum tipo de coisa que instancie a propriedade de ser prescritivo, categórico e objetivo. Se não existem fatos e propriedades morais, então não há nada a que corresponda às nossas afirmações sobre o certo e o errado e, portanto, não há nada que torne verdadeiras as sentenças morais. Por isso, o discurso moral é sistemática e uniformemente falso, um erro. Mackie apresenta dois argumentos para mostrar que não existe algo no mundo com as qualidades atribuídas às propriedades morais. O primeiro ficou conhecido como o argumento da relatividade.

A premissa central é o *fato* de que há desacordo moral. Note que este pressuposto não é nem argumentável, pois é uma verdade antropológica e é certo que muito poucas pessoas discutiriam seriamente se realmente há desacordo moral. Pode haver desacordos morais dentro de uma mesma família que tem, supostamente, valores razoavelmente homogêneos. Uma comunidade pode ter muitas coisas em comum, mas certamente haverá inúmeros desacordos. Isso se intensifica quando pensamos em comunidades distintas ou países distintos. O fato é que as pessoas têm códigos morais um pouco diferentes e isso implica que terão respostas diferentes para os mesmos assuntos morais. E muitas vezes essas respostas são radicalmente diferentes. Ora, isso deve poder dizer algo a respeito das questões de segunda ordem. Se as propriedades morais fossem instanciadas, tal como as propriedades naturais são, então, será que haveria tantos desacordos entre as pessoas? O que nos faz pensar que não, é que nas ciências como a biologia, química ou

física, que supostamente tratam de apreender verdades objetivas, há muito menos desacordos. E os desacordos que existem normalmente são resultado de hipóteses e inferências baseadas em evidências inadequadas que, com investigação mais acurada, podem ser corrigidos<sup>106</sup>. Mas como corrigir as divergências morais? A que tipo de fatos podemos recorrer?

Segundo Mackie, os valores parecem muito mais ser o resultado de um modo de vida que as pessoas adotam do que o modo de vida adotado ser regulado pelo endosso ou reconhecimento de certos valores. Ele nos dá o exemplo da monogamia. As pessoas aprovam a monogamia porque participam de um modelo monogâmico de vida; e não participam de um modelo monogâmico de vida porque aprovam a monogamia. Em outras palavras, isso quer dizer que as pessoas não reconhecem certos valores objetivos para, daí, viver de acordo com esses valores, mas que adotam os valores na medida em que tem tal e tal estilo de vida. Mackie acredita que isso explica melhor a existência de desacordos morais do que a falha em reconhecer certas propriedades morais objetivas. O desacordo é muito mais o resultado dos diferentes modos de vida do que da percepção errônea das propriedades morais. Ele diz:

Em resumo, o argumento da relatividade tem alguma força simplesmente porque as variações atuais nos códigos morais são mais prontamente explicadas pela hipótese de que elas refletem modos de vida do que pela hipótese de que elas expressam percepções, a maioria delas seriamente inadequadas e distorcidas, de valores objetivos<sup>107</sup>.

No entanto, ele mesmo reconhece uma suposta fraqueza deste argumento. Se poderia dizer que os valores objetivos que constituem a moralidade são mais gerais e que não há muito desacordo a respeito desses valores gerais. As pessoas normalmente concordam que é bom respeitar as outras pessoas, que mentir é incorreto, que o assassinato é condenável, que devemos cuidar das crianças e dos mais velhos. Além disso, se poderia dizer que há grande acordo sobre o fato de que devemos promover a felicidade ou, ao menos, minimizar a infelicidade e que este é o princípio objetivo que regula a vida das pessoas. Mas Mackie sustenta que essa objeção, se apresenta alguma dificuldade ao seu argumento, apresenta apenas parcialmente. Isso porque não mostra

---

<sup>106</sup> MACKIE, 1977, *Op. Cit.*, 36.

<sup>107</sup> *Ibidem*, p. 37.

que os princípios morais são objetivos. Assume apenas que há um acordo sobre os princípios gerais e, mesmo que estes princípios sejam objetivos, algo que a objeção não mostra, ainda assim se pode dizer que os princípios mais específicos e que regulam a vida das pessoas são objeto de grande desacordo. Por isso, Mackie acredita que o seu argumento da relatividade continua sendo influente diante de qualquer tipo de contraexemplo que apela para algum ponto de acordo entre as moralidades.

Com o argumento da relatividade, Mackie supõe que é mais razoável pensar nas propriedades morais como resultado daquilo que as pessoas aprovam do que como o reconhecimento de certos fatos *naturais* e objetivos. Esse argumento estabelece que temos boas razões para pensar que as propriedades morais não são objetivas, dados os desacordos.

Mas esse é o ponto menos importante. O principal argumento de Mackie é o argumento da estranheza. Este argumento objetiva mostrar que, no mundo, não há propriedades morais e, se houvesse, elas teriam que ser alguma entidade muito estranha e diferente de tudo o que conhecemos até hoje.

O argumento é composto por uma parte metafísica e uma epistemológica, e é expresso por esta tão citada passagem:

Se houvessem valores objetivos, então eles seriam entidades ou qualidades ou relações de um tipo muito *estranho*, completamente diferente de tudo no universo. Correspondentemente, se estivéssemos conscientes delas, teria que ser através de alguma faculdade especial de percepção moral ou intuição, completamente diferente dos métodos ordinários de conhecer todas as outras coisas<sup>108</sup>.

Mas o que seria tão estranho sobre as propriedades morais? Dado que Mackie assume que os valores são comumente entendidos como propriedades prescritivas e categoricamente objetivas, então deveria haver, no mundo, ações que tivessem propriedades como “deve ser promovida” ou “não deve ser feita” como partes de sua própria constituição. Quer dizer, seria muito estranho imaginar que a ação de torturar alguém tivesse “nela” a propriedade de que não deve ser feita ou

---

<sup>108</sup> *Ibidem*, p. 38. Itálico adicionado.

a propriedade da *incorreção*. Se olharmos para alguém torturando uma pessoa veremos, no máximo, pessoas, vestidas de tal modo, sendo as roupas de tais cores, com tais e tais características físicas, podemos ver os objetos que o torturador estiver usando (se estiver usando algum), podemos ver sangue, interpretar as reações do torturado como sendo expressões de sofrimento, e assim por diante. Mas em nenhum momento veremos a propriedade da *incorreção* como parte constitutiva dessa ação. E seria no mínimo estranho uma ação que tivesse tal propriedade. Como diz Mackie, seria diferente de todas as outras coisas do mundo. Por isso, a estranheza está também em como reconheceríamos tais propriedades. Provavelmente não poderia ser através de nenhuma de nossas fontes ordinárias de conhecimento. Não vemos, ouvimos, tocamos, cheiramos ou sentimos a *incorreção*, a *bondade* ou a *justiça*. Além disso, parece que não reconhecemos as propriedades morais por nenhum tipo de inferência *a priori* ou puramente dedutiva. Assim, o conhecimento dessas propriedades teria que ser através de alguma outra fonte, completamente diferente e peculiar.

Além da parte metafísica e epistemológica há um argumento a mais para sustentar a estranheza das propriedades morais; diz respeito à superveniência. Mackie parece tratar esta parte como um argumento independente, já que diz que o argumento da estranheza tem duas partes apenas. No entanto, sugere que é uma outra forma de mostrar a estranheza das propriedades morais. É irrelevante se o argumento tem duas ou três partes. O importante é o que essas ideias implicam. Sobre a superveniência, Mackie, então, diz o seguinte.

Qual é a conexão entre o fato natural de que uma ação é um exemplo de crueldade deliberada – quer dizer, causar dor apenas por diversão – e o fato moral de que é errada? Não pode ser uma implicação, uma necessidade lógica ou semântica. Além disso, não é meramente que as duas características ocorrem juntas. A incorreção deve, de algum modo, ser ‘consequente’ ou ‘superveniente’; é errado porque é um exemplo de crueldade deliberada. Mas o que *no mundo* é significado por este ‘porque’?

Um naturalista normalmente assume que a propriedade moral é resultado do conjunto de propriedades naturais. As propriedades naturais de algum modo implicam as propriedades morais. Ou, ainda, as propriedades morais são de uma ordem superior às naturais, que são a

sua instanciação. O ponto de Mackie é sobre que tipo de implicação é essa entre as propriedades morais e naturais. Se não há nada no mundo que corresponda a uma propriedade prescritiva e categoricamente objetiva, então o que justifica a obtenção dessa propriedade misteriosa a partir de um conjunto de propriedades naturais? O problema consiste em saber por que algo metafisicamente estranho se segue de coisas que podemos ver e tocar. Parece que o naturalista está dando um “passo no escuro”.

Com esses argumentos, o argumento da relatividade e o argumento da estranheza – que é composto por um argumento metafísico, um epistemológico e um argumento sobre a superveniência – Mackie acredita que podemos concluir que não há nenhum tipo de entidade que instancia as propriedades de ser prescritiva e categoricamente objetiva. Note que o argumento da estranheza apenas estabelece que, se houvessem propriedades desse tipo, então elas seriam de uma natureza muito estranha e peculiar. Quer dizer, não estabelece que tais propriedades *não* existam. Isto é o que supostamente se segue do argumento da estranheza e da relatividade. Se olharmos para o mundo não encontraremos tais propriedades metafisicamente estranhas. Portanto, elas não existem, sustenta Mackie. Se assim é, então não há nada que corresponda às nossas sentenças morais. Quando dizemos que um governo autoritário é injusto ou que comer carne de animais não humanos é uma prática incorreta, não há, no mundo, propriedades como a justiça e a incorreção. Isso significa que nossas sentenças morais não podem ser verdadeiras. Assim, discurso moral é sempre *falso*. Em uma palavra, é um *erro*.

Que tipo de implicação isso teria para a moralidade comum? Se a posição de Mackie realmente nos convence e acreditamos que a moralidade é um erro e nossas crenças morais são unanimemente falsas, então será que deveríamos sair por aí roubando, mentindo e assassinando pessoas? A resposta é ‘não’. Questões metaéticas são normativamente neutras. E isso significa dizer que do fato de assumirmos que a moralidade “não existe”, por assim dizer, não se seguem conclusões sobre como devemos nos comportar. Os filósofos que defendem a teoria do erro são também conhecidos por *ficcionalistas*. Eles sustentam que a moralidade é uma espécie de ficção, mas não defendem que não deveríamos ser morais, pois se trata de uma ficção bastante útil. Podemos pensar em várias ficções que são bastante úteis; Papai Noel é um bom exemplo. Desde cedo ensinamos às crianças que se elas se comportarem bem, se forem educadas, fazerem o dever de casa e se alimentarem bem, por exemplo, então serão bem

recompensadas com um bom presente pelo Papai Noel, que supostamente saberá sobre tudo o que fizeram ou deixaram de fazer. Obviamente, não há uma coisa tal que corresponda a este objeto fictício, mas as pessoas tendem a acreditar que é uma ferramenta bem útil para ensinar a “moralidade básica” para as crianças. O mesmo com a moralidade. Mesmo que não existam objetos objetivos e categoricamente prescritivos, como o naturalista parece sustentar que existem, nos é bastante útil que as pessoas continuem agindo moralmente.

Porque a teoria do erro é uma objeção ao naturalismo? Ora, o naturalista está sustentando que fatos morais existem objetivamente e são parte do mundo natural. Se Mackie está correto, então fatos morais não são naturais, mas apenas propriedades ficcionais que atribuímos a certos objetos e ações. Por isso, não se pode, ao contrário do que pensa o naturalista, dizer que certos juízos morais são verdadeiros, pois não há objetos que correspondam à instanciação de propriedades prescritivas e categoricamente objetivas.

## **OBSERVAÇÕES FINAIS**

Para concluir, é útil relembrar as principais coisas tratadas nesse primeiro capítulo. Na primeira seção, apresentei uma definição do realismo moral a partir de Brink e disse que há duas formas de defendê-lo. Uma delas é assumir que há fatos e propriedades morais independentes de nossas mentes e que são fatos e propriedades não naturais. A outra forma é dizer que há fatos e propriedades morais independentes de nossas mentes e que são naturais. A partir disso, fiz várias distinções, tais como entre o naturalismo metafísico, metodológico e moral. Ao tratar do problema da definição do que é para uma propriedade o ser natural, afirmei que há um tratamento padrão, que consiste em dizer que algo é natural se é passível de estudo pelas ciências naturais. A seguir, apresentei o critério de Copp, que diz que uma propriedade é natural se, e somente se, qualquer proposição sintética sobre a sua instanciação que pode ser conhecida, pode ser conhecida apenas empiricamente. Concluí fazendo uma distinção entre o naturalismo analítico e o naturalismo sintético.

A partir da segunda seção, comecei a apresentar algumas críticas ao naturalismo moral. Apresentei o argumento da questão aberta de Moore. Moore acreditava que o naturalismo moral é falso porque era supostamente vulnerável a uma falácia, que chamou de falácia

naturalista. Essa falácia era cometida pelos eticistas que identificavam as propriedades morais com algum tipo de propriedade natural. Em termos simples, o seu argumento era o de que para qualquer definição da propriedade moral *bom* é sempre possível aplicar o predicado moral à própria definição ou, ao menos, perguntar de a propriedade natural em questão é ou não *boa*. E, para Moore, isso é razão suficiente para mostrar que não há uma identidade analítica entre propriedades morais e naturais. Portanto, o naturalismo deveria ser falso.

Na terceira seção, considerei os argumentos de Hare contra o naturalismo. Comecei apresentando algumas ideias gerais do emotivismo para ver em que sentido Hare rompe com este modelo não cognitivista e começa a desenvolver o seu prescritivismo. Depois disso, expus sua reformulação do argumento da questão aberta de Moore. Tentei explicar as várias características das palavras de valor que as tornam diferentes das palavras que referem propriedades naturais na visão de Hare. Dei atenção à tese da prescritividade e à tese da universalizabilidade, que são o cerne da teoria de Hare. Uma das principais objeções de Hare ao naturalismo é de que o naturalista não reconhece a prescritividade das propriedades morais ao afirmar que estas são naturais. Por isso, não dá conta da suposta característica motivacional dos juízos morais. Para finalizar, apresentei o seu argumento mais recente contra o naturalismo, a ideia de que o naturalismo implica em relativismo, pois os *padrões* de avaliação das ações supostamente sempre variam entre as culturas.

E na última seção apresentei a teoria do erro de Mackie como objeção ao naturalismo moral. Essa teoria é composta por duas partes, uma conceitual e uma ontológica. Ao iniciar pela parte conceitual, disse que Mackie entende os juízos morais como objetivos e categoricamente prescritivos. Expliquei o que significa cada uma dessas noções e parti para a segunda parte. Nesta parte, explicitiei os argumentos de Mackie a favor da ideia de que propriedades morais objetivas e categoricamente prescritivas não existem e, portanto, os juízos morais são uniformemente falsos. Basicamente, são dois argumentos. O argumento da relatividade e o argumento da estranheza. Sua conclusão é de que, dado que supostamente não existem fatos morais, então o discurso moral é falso e a moralidade é uma espécie de erro.

## II – NATURALISMO NÃO REDUCIONISTA

“The chief problem [of naturalism], of course, is to explain what is might mean to claim that moral properties are *natural* properties. I think that once this is properly explained, naturalism will seem enormously attractive”. (COPP, 2003).

### INTRODUÇÃO

David Copp tem razão. A razoabilidade do naturalismo depende do seu sucesso explicativo das propriedades morais em termos de propriedades naturais. Dadas as objeções acima consideradas, parece que o naturalismo analítico não atingiu seu objetivo suficientemente. Seu problema parece estar em sustentar que propriedades morais e naturais são analiticamente equivalentes. Qualquer argumento que mostre que em alguma medida as propriedades morais possuem algum predicado que propriedades naturais não possuem já é condição suficiente para trazer problemas. Parece que o naturalismo analítico não tem uma boa explicação do que significa dizer que propriedades morais são naturais. Moore, Hare e Mackie, dentre outros, tentaram mostrar que o empreendimento do naturalista está fadado ao fracasso e que o naturalismo é *definitivamente*<sup>109</sup> falso. O argumento da questão aberta exerceu profunda influência nos desenvolvimentos de teorias metaéticas posteriores. Isso fez com que duas alternativas fossem principalmente seguidas: (i) aceitar o realismo moral não naturalista, como fez Moore, ou (ii) rejeitar o realismo optando por algum tipo de irrealismo como o emotivismo, o subjetivismo ou o prescritivismo. Em ambas as alternativas, o naturalismo fora deixado de lado.

Mas os tempos mudam. Vários avanços recentes na metafísica, epistemologia, filosofia da linguagem e filosofia da mente têm propiciado o renascimento de uma nova versão do naturalismo. Esse naturalismo renovado não requer o reducionismo das propriedades morais às propriedades naturais. Essa talvez seja a sua principal vantagem e, por isso, tem atraído o apoio de vários filósofos. Richard Boyd, David Brink e Nicholas Sturgeon têm sido seus protagonistas. Esses autores acreditam ter finalmente respondido o problema que

---

<sup>109</sup> HARE, R. M. *The Language of Morals*, New York: Oxford University Press, 1952, p. 92-93.

Copp, na citação acima, considera ser o principal desafio do naturalismo. A motivação central para resgatar o naturalismo e dar-lhe uma nova roupagem é propiciar uma defesa do realismo moral. A grande tendência a endossar essa posição é manifesta: o novo realismo moral naturalista, se verdadeiro, traz objetividade à ética, nos certifica de que existem respostas corretas aos problemas morais e que nos é possível descobrir quais são essas respostas. Além disso, o naturalismo não reducionista fornece várias respostas aos clássicos problemas que são atribuídos ao naturalismo.

Neste capítulo, irei apresentar as principais ideias do naturalismo não reducionista, mostrando em que medida não é vulnerável a algumas objeções frequentes ao naturalismo e como se diferencia do naturalismo reducionista. Vou dividir essa tarefa em três partes. Na primeira, e mais introdutória, quero considerar alguns avanços em várias áreas da filosofia que tornaram possível o seu desenvolvimento. Gostaria de mostrar como os naturalistas têm se apropriado dessas ideias e as aplicado na ética. Aqui veremos principalmente argumentos negativos a favor do naturalismo. Farei alguns comentários mais específicos mostrando como o naturalismo evita críticas como o argumento da questão aberta e a reformulação hareana deste argumento.

Na segunda parte, irei me concentrar de maneira mais específica na teoria desenvolvida por Richard Boyd. Vou dedicar uma seção inteira a este assunto porque, em primeiro lugar, Boyd desenvolve uma teoria semântica externalista que é o pilar do naturalismo e, em segundo, porque talvez a mais importante objeção que se tem hoje ao novo naturalismo se dirige justamente à sua pressuposição semântica. Isso é algo que teremos que discutir futuramente e, portanto, é bom ter claro o alvo a que a objeção se dirige. Estarei dando atenção principal aos elementos semânticos da teoria de Boyd, mas é importante notar que ele não trata apenas do problema semântico das propriedades morais. Ficarão evidentes tanto aspectos metafísicos quanto epistemológicos de sua teoria. No final desta seção, vou apontar também como esta teoria lida com dois dos principais problemas atribuídos ao naturalismo.

Na terceira parte, quero apresentar um modo de defender o naturalismo baseado na inferência a favor da melhor explicação. Esta tem sido uma das estratégias argumentativas centrais do novo naturalismo. Aqui ficará mais evidente ainda o papel desempenhado pelo realismo moral, pois a estratégia argumentativa é similar à dos realistas científicos. Vou começar pela objeção ao naturalismo não reducionista estabelecida por Gilbert Harman. Depois, mostrarei como os naturalistas lidam com este problema e como isso constitui uma

defesa de que há fatos morais e que estes são naturais. Este percurso nos dará uma noção razoável do novo naturalismo, em que medida se distingue do naturalismo reducionista e como lida com algumas objeções. O leitor perceberá que, embora a principal razão para o estabelecimento do naturalismo seja fornecer uma defesa do realismo moral, não darei muita atenção a esta tese mais geral. Isso porque, em primeiro lugar, quero restringir a discussão de modo mais específico ao naturalismo moral somente e, em segundo, porque não quero me comprometer com discussões sobre as versões não naturalistas do realismo moral.

### 1. Os pressupostos

Nesta seção, gostaria de apresentar alguns dos desenvolvimentos que tornaram possível o avanço do novo naturalismo. Vou seguir a seguinte ordem: (i) primeiro, vou falar sobre a recusa da tese de que se dois termos são sinônimos então devem referir o mesmo objeto; (ii) segundo, sobre a tese de que alguns termos singulares e tipos naturais são designadores rígidos; (iii) terceiro, sobre a tese de que os estados mentais são propriedades funcionais e, (iv) quarto, sobre a teoria semântica externalista. No decorrer desta apresentação, ficará evidente uma quinta tese, a saber, (v) o naturalismo moral não requer nenhum tipo de reducionismo.

O argumento da questão aberta, assim como as críticas de Hare ao naturalismo, assumem como teoria semântica aquilo que é comumente denominado de descritivismo. Esta teoria, que é principalmente representada pelos trabalhos de Frege e Russel, sustenta que o sentido de um nome é dado por um conjunto de descrições definidas associadas ao nome<sup>110</sup>. O significado de ‘Aristóteles’, por exemplo, é conferido através das inúmeras designações pelas quais nos referimos a este indivíduo, como ‘O professor de Alexandre’, ‘O criador da lógica formal’, ‘O autor da *Ética Nicômaco*’, ‘O maior aluno de Platão’, ‘O último grande filósofo da Antiguidade’, etc. Uma das principais pressuposições tanto de Moore quanto de Hare é o papel desempenhado pela sinonímia. Moore argumentou que se dois termos são sinônimos, então um falante competente da língua deve associar o mesmo conjunto de descrições definidas a ambos os termos e, por

---

<sup>110</sup> Cf. DEVITT, M. & STERELNY, K. *Language and Reality: An Introduction to the Philosophy of Language*. Second Edition, Blackwell Publishers: United States, MIT Press, 1999, p. 45.

consequente, os termos devem fixar a mesma propriedade, e um falante competente não deve poder duvidar se o significado é o mesmo. Dado que os termos que expressam propriedades morais e os termos que expressam propriedades não morais não são sinônimos, como o argumento da questão aberta supostamente mostra, pois os falantes competentes associam um conjunto de descrições diferentes a uns e a outros, então estas propriedades devem ser diferentes. Da mesma forma, Hare defendeu que *bom* e um conjunto de termos que expressam propriedades naturais *N* não podem ser sinônimos ou substituídos numa sentença sem mudança de significado, pois os falantes competentes usam *bom* no discurso comum para *aprovar* enquanto que os termos que expressam propriedades naturais apenas *descrevem* fatos. Mas com os desenvolvimentos na metafísica e filosofia da linguagem, protagonizados por Putnam<sup>111</sup> e Kripke<sup>112</sup>, descobriu-se que dois termos podem fixar a mesma referência mesmo que tenham significados diferentes. Isso fez com que se rejeitasse uma das principais pressuposições do argumento da questão aberta e da reformulação de Hare, a saber, a sinonímia como critério de identidade entre duas propriedades. Este é o primeiro pressuposto a favor da reascensão do naturalismo que gostaria de abordar.

Considere o seguinte exemplo.

Água é H<sub>2</sub>O.

O significado dos termos ‘água’ e ‘H<sub>2</sub>O’ são diferentes. Ao primeiro, associamos a propriedade de ser molhado, incolor, insípido, usado para beber, para regar as plantas, que corre nos rios e preenche os lagos e oceanos etc. O segundo possui um significado científico preciso usado para representar uma composição química entre dois elementos. Usamos ‘água’ no discurso cotidiano para dizer ‘Quero água’, ‘Há água escorrendo na rua’, ‘A piscina está com pouca água’ ou ‘A água do mar é salgada’. Mas não usamos ‘H<sub>2</sub>O’ no discurso comum para falar ‘O H<sub>2</sub>O do mar é salgado’ ou ‘Quero H<sub>2</sub>O para matar minha sede’. No

---

<sup>111</sup> Cf. PUTNAM, H. The Meaning of “Meaning”. In *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University Press, 1975, pp. 215-271.

<sup>112</sup> Cf. KRIPKE, S. Identity and Necessity. In SCHWARTZ, S. P. *Naming, Necessity and Natural Kinds*. Ithaca: Cornell University Press, 1977, pp. 67-101. E Cf. KRIPKE, S. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

entanto, hoje sabemos que ‘Água’ é ‘H<sub>2</sub>O’, quer dizer, o objeto referido pelos termos é exatamente o mesmo. Mas nem sempre foi assim.

Se olharmos para o passado retrocedendo algumas centenas de anos, veremos que, embora associássemos as mesmas propriedades ao termo ‘água’ que associamos hoje, tais como, insípida, incolor, que corre nos rios, etc., nós não sabíamos que se tratava de um material de tal e tal composição química denominado H<sub>2</sub>O. Isto é, não sabíamos que ‘água’ é ‘H<sub>2</sub>O’ porque não compreendíamos que o mesmo objeto que corre nos rios e é utilizado para matar a sede possui uma fusão entre dois elementos, hidrogênio e oxigênio. Somente com os desenvolvimentos da ciência moderna é que pudemos descobrir que ‘água’ e ‘H<sub>2</sub>O’ são o mesmo objeto. Assim, mesmo que o significado associado aos termos ‘água’ e ‘H<sub>2</sub>O’ seja diferente, é uma verdade *necessária*<sup>113</sup> que ‘água’ é ‘H<sub>2</sub>O’. Portanto, ‘água’ e ‘H<sub>2</sub>O’ possuem uma identidade, não analítica, mas *sintética*.

Se quisermos mais exemplos desse tipo de identidade sintética, basta pensar em temperatura = energia molecular cinética, luz = radiação eletromagnética<sup>114</sup>. Estes exemplos ilustram o que na literatura filosófica é usualmente chamado de verdades necessárias *a posteriori*. É uma verdade necessária que água é H<sub>2</sub>O. É *a posteriori* porque não descobrimos isto apenas analisando os termos em questão, mas através de investigação empírica.

Se dois termos podem não ser sinônimos e mesmo assim referir um mesmo objeto, então parece que uma das principais pressuposições do argumento da questão aberta é falsa.

O naturalista pode admitir que não há nem sinonímia nem implicação de significado entre termos morais e não morais, naturais por exemplo, e ainda assim manter que os fatos e propriedades morais são idênticos a, ou constituídos por, fatos e propriedades científicos naturais e sociais. A ideia do naturalista de identidade ou constituição pode

---

<sup>113</sup> O porquê de ser uma verdade necessária requer algumas qualificações, que serão dadas abaixo.

<sup>114</sup> BRINK, D. O. *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press, 1989, p. 157.

ser construída como expressando necessidades morais sintéticas [*a posteriori*]<sup>115</sup>.

Assim, nós temos dois conceitos que expressam diferentes significados e podem referir a mesma propriedade. O argumento sugere que se poderia *pensar* que os fatos e propriedades morais e naturais são diferentes, mas que isso não significa que *são* realmente diferentes. O naturalista pode dizer que há uma diferença entre, por um lado, palavras, termos e conceitos, e, por outro, realidade e matéria e argumentar que dois conceitos ou termos podem ter significados muito diferentes, mas referir um mesmo objeto<sup>116</sup>.

Se quisermos, podemos considerar o argumento de Hare à luz da ideia de que dois termos podem ter significados diferentes e mesmo assim referir a mesma propriedade. Como vimos, ele argumenta que ‘*x é bom*’ e ‘*x é N*’ são diferentes, pois dizer que possuem o mesmo significado nos impossibilita de dizer certas coisas que estamos acostumados a dizer no discurso comum. Ou seja, se assumo que a definição de *bom é N* e que o que constitui *N* é o conjunto de propriedades descritivas *abc*, então não posso dizer que *a* ou *b* ou *c* são bons (algo que dizemos *significativamente* no discurso comum), pois estaria apenas enunciando algo analiticamente verdadeiro tal como ‘*b é b*’.

Vejam, agora, como esse argumento funcionaria para ‘água/H<sub>2</sub>O’. Naturalmente, teremos que supor que ainda não *sabemos* que ‘água’ é ‘H<sub>2</sub>O’ e que há apenas uma tese que reivindica isso (tal como o naturalista reivindica que *bom é N*). Vamos supor então que ‘água’ possui um conjunto de propriedades descritivas *abc*, que pode ser algo como, a propriedade de ser insípida, incolor e preencher os rios e mares. Imaginemos que eu diga ‘*b é H<sub>2</sub>O*’. Seguindo a tese que reivindica que ‘água’ é ‘H<sub>2</sub>O’, então eu estaria dizendo algo analiticamente verdadeiro, a saber, que ‘*b é b*’. Isso porque *b* é uma instância de ‘água’ e ‘água’ é ‘H<sub>2</sub>O’, portanto *b é*, também, uma instância de ‘H<sub>2</sub>O’. Mas, se atentarmos para o *uso* de nossa linguagem no discurso ordinário, veremos que quando dizemos que ‘*b é H<sub>2</sub>O*’ não estamos querendo dizer ‘*b é b*’. Portanto, a tese que diz que ‘água’ é

---

<sup>115</sup> *Ibidem*, p. 166. Brink não usa o termo ‘*a posteriori*’ nesta passagem. Mas adiante (*Ibidem*, p. 175) afirma que a relação entre propriedades morais e naturais é uma relação necessária *a posteriori*.

<sup>116</sup> Cf. KIRCHIN, S. *Metaethics*. Palgrave Macmillan, 2012, p. 46.

‘H<sub>2</sub>O’ é falsa, pois estaria forçando-nos a admitir a analiticidade de certas coisas que dizemos no discurso comum sem a pretensão de que sejam afirmações analíticas. Ou, pelo menos é isso que o raciocínio empregado por Hare nos força a concluir.

Notemos agora como a descoberta de que ‘água’ é ‘H<sub>2</sub>O’ torna falsa a premissa que apela para a linguagem ordinária a fim de mostrar que quando dizemos ‘*b* é H<sub>2</sub>O’ não estamos querendo dizer ‘*b* é *b*’. Se esta premissa é falsa, o argumento é certamente inválido, pois a conclusão não se segue das premissas. Vimos que se digo que ‘*b* é H<sub>2</sub>O’, pressupondo que ‘água’ é ‘H<sub>2</sub>O’, estou dizendo o mesmo que ‘*b* é *b*’. Ora, a descoberta científica *mostra* que ‘água’ é ‘H<sub>2</sub>O’. Então, quando digo que ‘*b* é H<sub>2</sub>O’ estou *realmente* dizendo o mesmo que ‘*b* é *b*’, pois *b* é uma propriedade de ‘água’ e, portanto, é uma propriedade de ‘H<sub>2</sub>O’. Se a linguagem ordinária nos diz que quando usamos ‘*b* é H<sub>2</sub>O’ não estamos querendo dizer meramente ‘*b* é *b*’, isso nos mostra que o conceito de ‘água’ e ‘H<sub>2</sub>O’ possuem significados diferentes somente. Mas isso não quer dizer que ‘água’ e ‘H<sub>2</sub>O’ sejam realmente diferentes, quer dizer, podem referir o mesmo objeto. Do fato de usarmos a linguagem de modo tal que não parece que ‘água’ é ‘H<sub>2</sub>O’ não significa que ‘água’ não seja ‘H<sub>2</sub>O’.

Isso traz consequências desastrosas para a reformulação do argumento da questão aberta de Hare. Suponha que futuramente descobramos que o naturalismo está certo e que não há uma diferença entre propriedades morais e naturais; *bom* é *N* e, portanto, dizer ‘*x* é *bom*’ é dizer o mesmo que ‘*x* é *N*’. Suponha que *N* é constituído pelo conjunto de propriedades descritivas *abc* e que eu diga que ‘*b* é *bom*’. Dado que *bom* é *N* e que *b* é uma instância de *N*, então *b* é também uma instância de *bom*. Portanto, quando digo que ‘*b* é *bom*’ estou realmente dizendo algo analiticamente verdadeiro, isto é, que ‘*b* é *b*’. Novamente, se a linguagem ordinária sugere que dizer “*b* é *bom*” e ‘*b* é *b*’ são coisas diferentes, como Hare argumenta, isso não implica que *bom* e *N* são propriedades diferentes. Significa apenas que *bom* e *N* possuem significado diferente na nossa prática linguística. Mas, como vimos, o naturalista aceita isso. Seu ponto é que propriedades morais e não morais são idênticas. Se admitirmos que isso é verdadeiro então temos que admitir, também, que o argumento de Hare não é uma refutação *definitiva* do naturalismo.

Um segundo desenvolvimento que contribuiu para a articulação do novo naturalismo foi a ideia de que nomes e termos para classes naturais são designadores rígidos. Esta ideia está intimamente relacionada com a anterior. Primeiro, porque se trata de uma crítica a

mais à teoria semântica descritivista e, segundo, porque nos ajuda a entender porque certas descobertas são verdades necessárias *a posteriori*. O conceito de designador rígido foi apresentado por Kripke como parte de um de seus argumentos modais para refutar as teorias descritivistas do significado. Acredito que a melhor forma de entender o que é um designador rígido é no contexto de sua crítica ao descritivismo. Os descritivistas sustentam que o significado de um nome é fixado através de um número de descrições definidas que são atribuídas a um objeto. Segundo Kripke<sup>117</sup>, uma das teses pressupostas pelo descritivismo é que se tanto o termo *t* quanto a descrição definida, ou o conjunto das descrições definidas *D*, fixam o mesmo objeto *o*, então a substituição de *D* por *t* deve preservar o significado da proposição expressa<sup>118</sup>. Dizer que ‘Aristóteles era magro’ e ‘O professor de Alexandre era magro’ é expressar a mesma proposição.

Agora, consideremos esta verdade necessária: ‘Se Aristóteles existiu, então Aristóteles era Aristóteles’. É uma verdade necessária porque é impossível que um objeto seja distinto dele mesmo. Se *t* pode ser substituído por *D* e preservar o significado da proposição, então o seguinte também deve ser uma verdade necessária: ‘Se Aristóteles existiu então Aristóteles era o professor de Alexandre’. Mas, será que esta última afirmação é uma verdade necessária? Parece que não. Dizer que ‘Se Aristóteles existiu então Aristóteles era o professor de Alexandre’ é dizer que é impossível que exista um Aristóteles tal que não seja professor de Alexandre. Ora, é certamente concebível um mundo possível em que Aristóteles existiu, mas não foi professor de Alexandre. Pode não ter sido professor, nem filósofo, muito menos se interessado por filosofia, por exemplo. Kripke conclui que o conjunto de descrições *D* não pode fixar o significado de *t* e, por isso, deve haver algo errado com o descritivismo<sup>119</sup>.

---

<sup>117</sup> Cf. KRIPKE, 1980, *Op. Cit.*, p. 71ss.

<sup>118</sup> Cf. SOAMES, S. *Philosophical Analysis in the Twentieth Century*. Volume 2, Princeton University Press, 2003, p. 337. Soames elenca duas teses que constituiriam a posição descritivista. A primeira é a tese em questão, a saber, “se *D* dá o significado de *n* então a substituição de um pelo outro preserva tanto o significado assim como a proposição expressa”. E a segunda diz que “o referente de um nome *n* é semanticamente fixado por uma descrição, um conjunto de descrições ou um agregado de descrições *D* que o falante associa a *n* naquele instante”. *Idem*.

<sup>119</sup> Neste caso, a primeira tese é falsa.

Mas isso revela algo importante para nossa discussão. É possível um mundo em que Aristóteles não tenha sido Aristóteles? Como dissemos acima, não, pois é uma verdade lógica que um objeto é sempre igual a ele mesmo. É possível conceber mundos em que Aristóteles tenha sido fundador da lógica formal, carpinteiro, atacante titular do Barcelona ou um músico. Mas não é possível conceber um mundo em que Aristóteles não tenha sido Aristóteles. É claro que não se trata meramente do nome ‘Aristóteles’, pois poderia ter tido outro, mas *do próprio* Aristóteles. Se começarmos a imaginar um mundo em que Aristóteles é um artesão e não filósofo, é *Aristóteles* que é artesão. Se imaginarmos um mundo em que é tal e tal, é *ele mesmo* que é tal e tal. Devemos notar que num mundo em que Aristóteles é artesão, as descrições definidas que são associadas a ele obviamente mudam. Ao invés de ‘O último grande filósofo da Antiguidade’ tem-se que dizer algo do tipo ‘O homem que faz caixas laqueadas’. Mas embora as descrições mudem, o referente do nome ‘Aristóteles’ não muda. Por isso, Kripke denominou *designador rígido* o termo que não muda o seu objeto referente de mundo para mundo. Este conceito contrasta com *designador flácido*, ou o termo que muda o seu referente entre os mundos possíveis<sup>120</sup>. ‘Aristóteles’ é um designador rígido. ‘O fundador da lógica modal’ é um designador flácido, pois podemos conceber um mundo em que Sócrates fundou a lógica modal.

Scott Soames apresenta a seguinte definição:

Um termo singular *t* é um designador rígido de um objeto *o* se designa *o* no que diz respeito a todos os estados possíveis do mundo (em que *o* existe); e, além do mais, *t* nunca designa algo além de *o* (no que diz respeito a qualquer estado possível do mundo)<sup>121</sup>.

Dito de outro modo, um termo *t* designa rigidamente um objeto *o* se é impossível conceber um mundo em que *o* existe e *t* não designa *o*. Além disso, se *t* é rígido então há somente um objeto que é designado por *t*. Um designador flácido, por outro lado, não designa o mesmo objeto *o* em todos os mundos possíveis e pode designar vários objetos como em ‘O homem que casou com Marta’, por exemplo. Kripke fornece um

<sup>120</sup> Para um comentário mais preciso sobre esta distinção, Cf. LYCAN, W. G. *Philosophy of Language: a contemporary introduction*. Second Edition, Routledge, 2000, p. 47.

<sup>121</sup> SOAMES, 2003, *Op. Cit.*, p. 340-341.

critério para distinguir termos singulares<sup>122</sup> rígidos de flácidos. Para todo termo singular, basta colocá-lo na seguinte fórmula: ‘ $x$  poderia não ter sido  $x$ ’. Se obtivermos uma frase verdadeira, então o designador não é rígido. Considere ‘Aristóteles foi o fundador da lógica formal’. Utilizando o critério teremos ‘Aristóteles poderia não ter sido o fundador da lógica formal’. Esta sentença é verdadeira, pois como já vimos não é uma necessidade sobre Aristóteles que tenha fundado a lógica formal. Assim, ‘O fundador da lógica formal’ é um designador flácido. Mas se considerarmos ‘Aristóteles poderia não ter sido Aristóteles’ veremos que se trata de uma afirmação falsa. Aristóteles poderia não ter sido várias coisas, e poderia ter sido outras, mas não poderia não ter sido *o próprio* Aristóteles. Quer dizer, poderia não ter tido as propriedades comumente associadas a ele, mas não poderia não ser ele. Portanto, é um designador rígido<sup>123</sup>.

De acordo com Horgan e Timmons, duas importantes consequências resultam da tese de Kripke<sup>124</sup>. A primeira é que sentenças que expressam identidade entre duas propriedades e que são compostas por designadores rígidos, como em ‘Água é  $H_2O$ ’, são verdades necessárias sem que sejam analíticas. O naturalista pode agora sustentar que fatos e propriedades morais são naturais sem precisar dizer que *bom* é analiticamente equivalente a *N*. Em outras palavras, isso significa que o naturalista não precisa ser reducionista. O argumento da questão aberta e os argumentos de Hare contra o naturalismo, se atingem, atingem apenas o naturalista que alega que os fatos e propriedades morais são analiticamente equivalentes aos fatos e propriedades naturais. O naturalista, agora, está livre desse tipo de objeção. A segunda consequência, que se segue da primeira, é que essas sentenças que são constituídas por designadores rígidos e expressam identidade entre duas propriedades constituem definições, não analíticas, mas sintéticas. Então, o naturalista pode agora sustentar que propriedades morais expressam propriedades naturais sem a necessidade de ter

---

<sup>122</sup> Termos singulares englobam tanto nomes próprios quanto descrições definidas.

<sup>123</sup> Agora, podemos entender porque é uma verdade necessária que ‘água’ é  $H_2O$ . Há um mundo possível em que algo seja água e não seja  $H_2O$ ? Parece que não, pois a substância que compõe a água é  $H_2O$ . Então ‘água’ é ‘ $H_2O$ ’ em todos os mundos possíveis, portanto, ‘água’ é um designador rígido.

<sup>124</sup> HORGAN, T. & TIMMONS, M. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, Volume 16, 1991, p. 447-465, p. 451.

sempre claro a que tipo de propriedades naturais as propriedades morais correspondem. Dado que a definição é sintética, então podemos ir descobrindo, à medida que o nosso conhecimento sobre o mundo evolui, a que tipo de propriedades naturais exatamente as propriedades morais correspondem. Novamente, em outros termos, isso significa dizer que o naturalista não precisa ser reducionista.

Um terceiro avanço que influenciou o desenvolvimento do naturalismo não reducionista foi a defesa, protagonizada por Putnam, de que os estados mentais são propriedades funcionais. O conjunto de teses a favor dessa ideia é usualmente denominado de funcionalismo. Aqui, gostaria apenas de fazer uma caracterização breve desta teoria para podermos ver como Brink e Boyd a adaptam para a filosofia moral. Há várias subdivisões dentro da teoria funcionalista que uma análise mais precisa certamente iria requerer, como o funcionalismo estado-máquina, o psicofuncionalismo e o funcionalismo analítico, mas não irei me ocupar destas especificidades aqui<sup>125</sup>.

Em filosofia da mente, o funcionalismo é uma teoria sobre a natureza dos estados mentais. Ela surgiu como uma alternativa a teorias como o dualismo cartesiano entre mente e corpo, ao behaviorismo e à teoria da identidade entre mente e cérebro. De acordo com Janet Levin, o funcionalismo

[...] é a doutrina de que o que faz algo ser um pensamento, desejo, dor (ou qualquer outro tipo de estado mental) depende, não de sua constituição interna, mas apenas de sua função, ou do papel que ela desempenha, no sistema cognitivo em que é parte. Mais precisamente, teorias funcionalistas tomam a identidade de um estado mental como sendo determinada por suas relações causais com estímulos sensoriais, outros estados mentais e o comportamento<sup>126</sup>.

---

<sup>125</sup> Para mais detalhes, ver: LEVIN, J. Functionalism, *Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.). Disponível em: <<http://plato.stanford.edu/archives/fall2013/entries/functionalism/>>. Acesso em: 17 de julho de 2016. E POLGER, T. W. Functionalism. *Internet Encyclopedia of Philosophy*. Disponível em: <<http://www.iep.utm.edu/functionism/>>. Acesso em: 17 de julho de 2016.

<sup>126</sup> LEVIN, 2013, *Op. Cit.*

Grande parte dos argumentos a favor do funcionalismo são argumentos negativos que sustentam que qualquer outra teoria sobre os estados mentais é falsa. Neste sentido, o funcionalismo é distinto do dualismo mente-corpo que sustenta que os estados mentais *são* um tipo especial de substância, a substância pensante (ou, *res cogitans*, como disse Descartes). Os funcionalistas não afirmam que um estado mental é constituído por tal e tal substância, mas que o que faz com que algo seja um estado mental é a sua propriedade de este *fazer* algo, desempenhar alguma *função*. É diferente também do behaviorismo que sustenta que não há diferença entre dois estados mentais a menos que haja alguma diferença notável no comportamento que é associado a cada um destes estados mentais. Há uma série de contraexemplos à teoria behaviorista que mostram que é possível imaginar dois indivíduos reagindo de maneira diferente aos mesmos estados mentais, como o exemplo do “ator perfeito” ou do “super-espartano”. Esses contraexemplos mostram que, se o behaviorismo é verdadeiro, então não é possível para dois indivíduos se comportarem de maneiras diferentes e terem o mesmo estado mental; mas um “ator perfeito” pode ter um comportamento exatamente similar ao meu frente ao estímulo *x* e ter um estado mental distinto, assim como um super-espartano pode ser educado a nunca expressar o comportamento usual de dor e ter o estado mental da dor<sup>127</sup>; assim, o behaviorismo deve ser falso, argumentam os funcionalistas. Além disso, o funcionalismo também rejeita a teoria da identidade entre cérebro e estados mentais. Esta teoria afirma que os estados mentais são estados cerebrais, ou, nas palavras de um de seus principais defensores, “sensações são processos cerebrais<sup>128</sup>”. Se isso é verdadeiro, então para cada estado mental deve haver um estado cerebral correspondente. Mas, como notou Putnam<sup>129</sup>, isso é falso, pois, para além dos cérebros dos mamíferos, o cérebro dos répteis, dos moluscos, etc., também são propensos a estados mentais semelhantes, como dor, por exemplo. No entanto, o cérebro de uma andorinha é significativamente diferente do que o cérebro de um humano, e isso quer dizer que os estados cerebrais também devem ser diferentes mesmo que tenham estados mentais

---

<sup>127</sup> PUTNAM, H. Brains and Behaviour. In.: *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University press, 1975, p. 325-341.

<sup>128</sup> SMART, J. J. C. Sensations and Brain Processes. *The Philosophical Review*, Vol. 68, No. 2. 1959.

<sup>129</sup> PUTNAM, H. The Nature of Mental States. In.: *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University press, 1975, p. 436.

similares, como no caso da dor. Então, a teoria da identidade entre mente e cérebro também deve ser falsa.

Os funcionalistas normalmente caracterizam um estado mental utilizando-se do conceito de *realização múltipla* ou *constituição*. Dizer que um estado mental não possui uma constituição interna, mas que é funcional, significa dizer que é realizado por uma série de estados físicos. Considere, por exemplo, o estado mental *dor*. Este estado mental é causado por algum tipo de dano corporal, produz o sentimento de querer se livrar deste estado, de que algo está errado com o corpo, produz sofrimento, produz angústia, e assim por diante. Todas essas propriedades constituem o estado mental *dor* ou, em outras palavras, são instâncias da *dor*, de modo que uma caracterização deste estado mental envolve todas estas propriedades funcionais. Costuma-se comparar a caracterização funcional dos estados mentais com objetos que usamos cotidianamente. Considere o exemplo da chave. Uma chave é uma entidade funcional. Ela não é meramente uma composição de determinada matéria (também é), mas é principalmente caracterizada pelas suas qualidades funcionais como abrir fechaduras, trancar fechaduras, assegurar coisas importantes num cofre, etc. A natureza dos estados mentais é similar, dizem os funcionalistas. Um estado mental causa tais e tais sentimentos, tais e tais reações e implica em tais e tais ações. Todas essas propriedades realizam multiplamente o estado mental. Em outras palavras, o estado mental é superveniente a essas propriedades.

Os naturalistas têm se apropriado dessa ideia e a aplicado na ética. Brink sugeriu que as propriedades morais são funcionais:

[O] realista moral pode alegar que as propriedades morais são propriedades funcionais. Ele pode alegar que o que é essencial às propriedades morais é o papel causal que elas desempenham nas atividades características dos organismos humanos. Em particular, o realista pode dizer que as propriedades morais são aquelas que incidem sobre a manutenção e a prosperidade dos organismos humanos. Presumivelmente, manutenção e prosperidade consistem em condições necessárias para a sobrevivência, outras necessidades associadas ao bem estar básico, desejos de vários tipos e capacidades humanas características. Pessoas, ações, políticas, estados de coisas, etc., irão suportar propriedades morais

sobre o bom na medida em que contribuem para a satisfação dessas necessidades, desejos e capacidades. Pessoas, ações, políticas, estados de coisas, etc., irão suportar propriedades morais sobre mal na medida em que elas falham para promover ou interferem na satisfação dessas necessidades, desejos e capacidades. Os estados físicos que contribuem ou interferem na satisfação dessas necessidades, desejos e capacidades, são os estados físicos sobre os quais, de acordo com esta teoria funcionalista, as propriedades morais em última instância supervêm<sup>130</sup>.

Ao manter que as propriedades morais são propriedades funcionais, o naturalista não precisa mais sustentar que há uma relação de identidade analítica entre propriedades morais e propriedades naturais, embora propriedades morais sejam naturais. Como Brink aponta, há um número indeterminado de propriedades naturais, como aquelas que contribuem para a sobrevivência, satisfação das necessidades básicas, bem estar básico, capacitações, etc., que, todas elas constituem a propriedade moral *bom*. Ao dizer que as propriedades morais são naturais, o naturalista tem em mente uma distinção entre o ‘é’ de *identidade* e o ‘é’ de *constituição*<sup>131</sup>. Neste sentido, o naturalista pode dizer que fatos e propriedades morais são constituídos por fatos e propriedades naturais, assim como o funcionalista diz que estados mentais são constituídos por propriedades funcionais. Isso não requer mais que o naturalista endosse a ideia de identidade analítica entre propriedades morais e naturais. Além disso, ele pode sustentar que existe um número indefinido de propriedades naturais que realizam multiplamente as propriedades morais. O naturalista não precisa fazer uma redução a uma terminologia natural, seja lá o que isto signifique<sup>132</sup>, pois esta redução seria incapaz de captar todas as instâncias naturais em que a propriedade moral é realizada. Por isso, mantém o vocabulário moral sustentando que os termos morais referem propriedades naturais<sup>133</sup>.

---

<sup>130</sup> BRINK, D. Moral Realism and Skeptical Arguments form Disagreement and Queerness. *Australian Journal of Philosophy*, 62:2, 1984, p. 111-125, p. 121, 122.

<sup>131</sup> Cf. BRINK, 1989, *Op. Cit.*, p. 157.

<sup>132</sup> Mais sobre este ponto abaixo.

<sup>133</sup> Cf. STURGEON, N. Ethical Naturalism. In COPP, D. (ed) *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, 2006, p. 98-99.

Consideremos um exemplo simples. Há inúmeras coisas que denominamos boas. Ajudar alguém que esteja passando por algum tipo de necessidade (inúmeros tipos de necessidade), não mentir na maioria das vezes, mentir se isso for necessário para um bem maior, cumprir as promessas, não ser agressivo, evitar a guerra, não roubar, ajudar alguém a suprir suas necessidades básicas, ajudar uma senhora com suas sacolas de compras, etc. A lista é muito grande. Essas propriedades que são boas não pertencem a nenhum domínio que não possa ser identificado empiricamente, portanto são naturais. Mas não há uma propriedade natural que seja comum a todas essas instâncias do *bom*. Então, não podemos reduzir o *bom* a uma propriedade natural comum. No entanto, *bom* parece ser multiplamente *realizado* ou *constituído* por esse conjunto de propriedades naturais que, embora sejam bastante diferentes umas de outras, são instâncias de uma mesma propriedade, quer dizer, da propriedade de ser *bom*. Portanto, Brink acredita que é muito mais razoável aceitar que propriedades morais são constituídas ou multiplamente realizadas por propriedades naturais do que que são idênticas. Aceitar que são idênticas é aceitar que há algo similar em todas as instâncias das coisas que avaliamos como boas. Mas o fato de propriedades morais poderem ser realizadas por um número indefinido de propriedades naturais é uma razão para recusarmos a identidade entre propriedades morais e naturais e aceitarmos a constituição<sup>134</sup>.

A determinação de quais propriedades naturais constituem as propriedades morais é papel das teorias normativas. O utilitarismo ou o kantismo, por exemplo, irão decidir, cada um a seu critério, quais são as propriedades naturais a partir das quais as propriedades morais supervêm<sup>135</sup>.

Não é difícil ver, mais uma vez, porque o argumento da questão aberta, ou a reformulação de Hare, não se aplica a esta forma de naturalismo. Recordemos que Moore argumentava que se é possível para um falante competente perguntar significativamente se ‘É *x*, que é *N*, *bom*?’, então é uma questão aberta se *bom* é *N* e, portanto, as propriedades morais e naturais não são analiticamente equivalentes. Agora, o novo naturalista concorda que ‘É *x*, que é *N*, *bom*?’ é uma questão aberta e não tem nenhuma pretensão de que seja fechada. Ele afirma que *bom* é realizado por um indefinido número de propriedades naturais, digamos  $N_1, N_2...N_n$ . Suponhamos que ele diga que ‘ $N_1$  é *bom*’. Ora, *bom* pode instanciar  $N_1, N_2...N_n$ , mas nenhuma instância individual

<sup>134</sup> Cf. BRINK, 1989, *Op. Cit.*, p. 157-159, 177.

<sup>135</sup> Este ponto será considerado de forma mais precisa adiante.

dessas propriedades, somente  $N_1$  ou somente  $N_2$ , é suficiente para realizar o *bom* totalmente. Portanto, quando se diz que ' $N_1$  é *bom*' não há a pretensão de que  $N_1$  e *bom* sejam sinônimos ou analiticamente equivalentes, pois *bom* é instanciado em  $N_1, N_2...N_n$  e  $N_1$ , somente, não instancia também  $N_2...N_n$ . Isso significa que para o não reducionista faz sempre sentido perguntar de um  $x$ , que é  $N$  se é *bom*, pois  $N$  é apenas uma instância que realiza o *bom*.

O último avanço que contribuiu para o renascimento do naturalismo que quero considerar foi a rejeição da teoria semântica tradicional descritivista que deu lugar às teorias externalistas, ou teorias causais da referência. Vimos, acima, alguns argumentos contra a teoria descritivista, como a ideia de que dois termos podem referir o mesmo objeto sem que sejam sinônimos e a ideia de que nomes e tipos naturais são designadores rígidos. Dado que uma parte importante deste trabalho será dedicada à discussão desta pressuposição semântica do naturalismo, gostaria de dedicar uma seção inteira para apresentar a aplicação desta ideia na ética. Aqui, vou considerar apenas um argumento negativo a favor da teoria semântica externalista. É importante estarmos cientes deste argumento, pois nas partes seguintes veremos que ele é pressuposição fundamental para discutirmos sobre a plausibilidade desse novo naturalismo.

O argumento a favor da teoria causal da referência que quero apresentar é a ideia de Putnam de que o significado dos termos não é inteiramente determinado pelas nossas mentes, mas algo externo é necessário. Assumamos que um termo singular possui uma extensão e uma intensão. A extensão de um termo é o conjunto de objetos que são fixados pelo termo. Assim, a extensão de 'leão' é todos os objetos que, no mundo, instanciam a propriedade de ser leão, ou seja, todos os leões. A intensão de um termo é o conceito ou o conjunto de descrições que os falantes associam ao termo. Por exemplo, os termos compostos 'Criatura com coração' e 'Criatura com rim' possuem intensões diferentes, pois são expressões linguísticas diferentes. Mas, pressupondo que todos os indivíduos com coração possuem rim, essas intensões distintas possuem a mesma extensão ou fixam o mesmo conjunto de objetos. Neste sentido, a teoria descritivista é a tese de que a extensão de um termo é totalmente determinada pela intensão. Passemos ao argumento de Putnam.

Imaginemos um mundo possível e o chamemos Terra Gêmea (TG). O nome é devido ao fato de que TG é muito similar ao nosso mundo (Terra ou T). Exceto por uma peculiaridade que destacaremos adiante, TG é *exatamente* similar a T. Há pessoas, casas, edifícios,

carros, política, economia, agricultura, etc., exatamente como há em T. Para cada coisa existe em T, existe também uma cópia em TG. Portanto, TG é igual a T, exceto por uma coisa. A composição química do líquido que os habitantes de TG chama ‘água’ não é  $H_2O$ , como em T, mas uma composição completamente diferente, denominada XYZ. Dado que a diferença é somente esta, o estado fenomenológico da água é o mesmo em T e TG. Os habitantes de TG reconhecem a água como incolor, insípida, que serve para matar a sede, que corre nos rios, que preenche os oceanos e lagos. Exatamente como os habitantes de T o fazem. Em TG as pessoas falam português e até mesmo o termo usado para denominar esse líquido é ‘água’.

Suponhamos que um habitante de T viaje para TG. À primeira vista, ele se entenderia muito bem com seus anfitriões a respeito de ‘água’, pois usam este termo para se referir ao objeto incolor, insípido, que serve para matar a sede, etc. O habitante de T naturalmente formaria a crença de que quando os habitantes de TG falam ‘água’ eles estão se referindo àquela substância composta por  $H_2O$ . Agora, imaginemos que um habitante de TG viaje para T. Ele formaria a crença de que quando os habitantes de T usam ‘água’ eles estão se referindo à substância XYZ. Temos duas opções agora. Podemos supor que quando os habitantes de T e os habitantes de TG usam ‘água’, eles estão se referindo ao mesmo objeto. Quer dizer, podemos supor que ‘água tal como usada pelos habitantes de T’ (‘T-água’) e ‘água tal como usada pelos habitantes de TG’ (‘TG-água’), possuem a mesma extensão. Então, quando um habitante de TG diz ‘água’ ele se refere tanto a  $H_2O$  quanto a XYZ. Do mesmo modo, quando um habitante de T usa ‘água’ ele se refere a  $H_2O$  e XYZ. A outra opção é dizer que quando eles usam ‘água’ estão se referindo a objetos diferentes. ‘T-água’ se refere somente a  $H_2O$  e ‘TG-água’ somente a XYZ. Isso significa que ‘T-água’ e ‘TG-água’ possuem a mesma intensão, mas extensões diferentes.

Obviamente, a teoria descritivista compromete-se com a primeira dessas opções. Pois, uma vez que a extensão de um termo, seja qual for, é fixada através do conjunto de descrições e, tanto os habitantes de T quanto os habitantes de TG usam exatamente as mesmas descrições para se referir a um objeto que tem exatamente as mesmas características fenomênicas, então ‘T-água’ e ‘TG-água’ tem de fixar o mesmo objeto.

Putnam defende que somente a segunda opção é verdadeira, pois nesse caso ‘água’ tem dois significados bem diferentes. E sua principal razão é a nossa intuição ao considerarmos a referência de ‘T-água’ e ‘TG-água’. A nossa intuição é que o significado do termo ‘água’ tal como usado pelos habitantes de T, ‘T-água’, não é água em TG. O

significado do termo ‘água’ tal como usado pelos habitantes de TG, ‘TG-água’, não é água em T<sup>136</sup>. Quando um habitante de T diz, apontando para copo, por exemplo, ‘Isso é água’, ele está dizendo ‘Isso é H<sub>2</sub>O’. E quando um habitante de TG diz, na mesma circunstância, ‘Isso é água’ ele está dizendo ‘Isso é XYZ’. Isso porque a extensão de ‘T-água’ é o conjunto de todos os objetos instanciados por H<sub>2</sub>O e a extensão de ‘TG-água’ é o conjunto de todos os objetos instanciados por XYZ<sup>137</sup>.

Isso significa que ‘T-água’ e ‘TG-água’ são designadores rígidos. Quer dizer, instanciam exatamente o mesmo objeto em todos os mundos possíveis em que existem. ‘T-água’ designa rigidamente H<sub>2</sub>O enquanto que ‘TG-água’ designa rigidamente XYZ. Caso contrário, os termos não difeririam em significado e poderíamos dizer que quando um habitante de T ou TG diz ‘água’ ele está se referindo tanto a H<sub>2</sub>O quanto a XYZ.

O descritivista teria que aceitar a conclusão de que ‘T-água’ e ‘TG-água’ possuem a mesma extensão, pois tanto os habitantes de T quanto os habitantes de TG tem as mesmas intensões sobre ‘água’. Mas isso é falso, como argumenta Putnam, pois o significado de um termo não é determinado unicamente pela intensão. Um indivíduo A e um indivíduo B podem ter um mesmo conjunto de descrições definidas sobre um termo F e A se referir a um objeto e B a outro objeto distinto. Portanto, conclui Putnam, o significado não está na cabeça<sup>138</sup>. Este argumento mostra que algo externo ao falante é necessário para a fixação do significado, algo que não depende da mente do falante e do conjunto de propriedades que o falante associa ao predicado. O resultado do argumento da terra gêmea é que o internalismo semântico é falso e que, obviamente, somente uma teoria externalista do significado é adequada.

Esta é somente uma caracterização negativa a favor da teoria semântica externalista. Mas a ideia básica aqui é de que pelo menos para alguns termos a referência não é inteiramente determinada pelos falantes, mas por uma relação causal com o mundo externo. O naturalista moral têm se apropriado desta ideia para sustentar que as propriedades morais são naturais, quer dizer, que a referência dos termos morais não é um tipo de avaliação ou aprovação que depende de algum código ou padrão moral, mas que é uma propriedade que pertence ao mundo e que é externa às opiniões, crenças, sistemas conceituais,

---

<sup>136</sup> PUTNAM, 1975, *Op. Cit.*, p. 224.

<sup>137</sup> *Idem.*

<sup>138</sup> *Ibidem*, p. 227.

etc., dos agentes morais. Neste sentido, ele pode argumentar que há fatos morais independentes de nossas mentes, que a ética pode proporcionar respostas objetivas aos problemas morais e que nos é possível conhecer essas respostas, desde que empreguemos a teoria normativa adequada. Iremos discutir mais adequadamente as implicações desta ideia na seção seguinte.

A ideia que queria considerar nesta seção é que o naturalismo não requer nenhuma forma de reducionismo. Como argumenta Sturgeon, a posição que afirma que para se ser um naturalista há que ser reducionista reivindicando que as propriedades morais podem ser reduzidas ao vocabulário da física já parte de um ponto não esclarecido. Isso porque normalmente se atribui ao naturalista o papel de reduzir as propriedades morais a uma terminologia fatural, descritiva<sup>139</sup> ou que denote propriedades naturais, mas a menos que se pressuponha que propriedades morais não são naturais e que, portanto, o naturalismo é falso, este tipo de afirmação não faz sentido, pois o naturalista está reivindicando justamente que as propriedades morais são fatuais, descritivas e que denotam propriedades naturais. Não há que se fazer redução alguma se se pressupor, ao menos para o bem do argumento, que o naturalismo é verdadeiro. Para finalizar esta seção, consideremos mais um argumento negativo a favor da não necessidade do reducionismo para uma teoria naturalista.

Sturgeon supõe duas alternativas que aquele que exige o reducionismo tem em mente. A primeira é uma (i) redução da ética aos termos da física; a segunda, (ii) uma redução da ética aos termos da teoria social, psicologia e biologia, que supostamente são ciências naturais<sup>140</sup>. As duas alternativas são falsas, ele argumenta. A primeira é falsa porque existem muito mais propriedades físicas do que o vocabulário da física consegue representar<sup>141</sup>. Uma tentativa de reduzir as propriedades morais à terminologia da física somente aumentaria o número de predicados necessários. Além disso, se o fisicalismo é verdadeiro, então se segue que todas as propriedades tratadas pela biologia, psicologia, sociologia, que são ciências naturais, são propriedades físicas. No entanto, não precisamos, e nem temos como, fazer uma redução completa dessas propriedades ao vocabulário da

---

<sup>139</sup> Cf. HARE, 1952, *Op. Cit.*, p. 82ss.

<sup>140</sup> Cf. STURGEON, N. *Moral Explanations*. In David Copp and David Zimmerman (eds.), *Morality, Reason and Truth*. Totowa, NJ: Rowman & Littlefield, 1985a, p. 59ss.

<sup>141</sup> *Ibidem*, p. 60.

física. A biologia, psicologia e teoria social certamente lidam com propriedades naturais e não é um ponto passível de discussão se a plausibilidade dessas disciplinas depende do fato de poderem ser redutíveis aos termos da física e aqueles que defendem que o naturalismo exige o reducionismo em nenhum momento exigem algum tipo de reducionismo dessas disciplinas. No entanto, eles querem que o naturalista apresente uma redução da ética a algum vocabulário não moral. De acordo com Sturgeon, não é de todo coerente exigir somente da ética este tipo de teste redutivo e não de outras disciplinas que também lidam com propriedades naturais<sup>142</sup>.

O que tentei mostrar até aqui foram alguns avanços em algumas áreas da filosofia que serviram como pressupostos para a defesa de uma versão renovada do naturalismo. Essas ideias contribuem conjuntamente para a tese de que uma defesa do naturalismo não requer o reducionismo e, por isso, não está sujeita às objeções clássicas como o argumento da questão aberta e as várias versões alternativas que tentam mostrar alguma dissimilaridade semântica ou metafísica entre propriedades morais e não morais.

Passemos a um ponto mais específico. Quero agora dar atenção especial à teoria semântica do naturalismo não reducionista. Até aqui vimos apenas um argumento negativo a favor de que a referência dos termos não é inteiramente determinada pelas descrições definidas associadas pelos falantes. Putnam argumentou que é necessário algo a mais, que é externo. Como afirma David Copp, Putnam sugeriu esta estratégia, mas quem a desenvolveu mais detalhadamente na ética foi Richard Boyd<sup>143</sup>. Por isso, agora, irei apresentar a versão de Boyd da teoria semântica do naturalismo.

## 2. Semântica moral

### 2.1. A teoria de Boyd

John Locke distinguiu entre essência nominal e essência real dos objetos<sup>144</sup>. À essência nominal corresponde todo o conjunto de

---

<sup>142</sup> *Idem.*

<sup>143</sup> Cf. COPP, D. Milk, Honey, and the Good Life on Moral Twin Earth. *Synthese*, Volume 124, No. 1, 2000, p. 113-137, p. 113.

<sup>144</sup> LOCKE, J. Ensaio Acerca do Entendimento Humano. In LOCKE, J. *Carta Acerca da Tolerância; Segundo Tratado Sobre o Governo; Ensaio Acerca do*

qualidades que atribuímos a um objeto. Ele nos dá o exemplo do ouro dizendo que a essência nominal do ouro é de “um corpo amarelo, de um certo peso, maleável, fusível e fixo”<sup>145</sup>. Poderíamos acrescentar outras qualidades tais como, substância brilhante, rara, densa, etc. Por outro lado, a essência real de um objeto é a propriedade material da qual o objeto é constituído. No caso de ouro, sua essência real é ter o número atômico 79. Ouro, então, tem um conjunto de propriedades através das quais é superficialmente identificado, mas possui uma propriedade real que *faz* com que seja ouro. Locke pensava que o significado de um termo como ouro é estabelecido através da essência nominal. Neste sentido, para algo ser considerado ouro devia apresentar o conjunto das qualidades comumente associadas tais como substância amarela, brilhante, maleável, rara, etc. No entanto, como sabemos, a posição de Locke não estava correta. Se o estivesse, então para que um objeto fosse considerado ouro seria condição suficiente o ter as características comumente associadas de substância amarela, brilhante, etc. Mas isso não é condição suficiente, pois podem existir objetos que satisfazem todas essas características nominais e não são ouro. A pirita de ferro ou, como também é conhecida, “ouro dos tolos”, nominalmente instancia todas as qualidades do ouro. Mas não é ouro, pois não é o *elemento químico de número atômico 79*. Como Kripke e Putnam argumentaram, isso significa que a referência de um termo não pode ser fixada através de um conjunto de qualidades frequentemente associadas ao termo, mas depende de algo que é externo àquilo que pensamos sobre o objeto. Em outras palavras, depende de como o mundo é. Por isso, na visão desses dois autores, *x* é ouro se e somente se *x* é constituído por aquela substância que é causalmente responsável pela nossa percepção de co-instâncias de amarelo, brilhante, maleável, etc.<sup>146</sup>.

Podemos dizer que depois de Kripke e Putnam esse tipo de definição tornou-se tradicional para classes naturais. Uma característica comum é que essas definições naturais estabelecem as condições necessárias e suficientes para algo ser considerado um membro de uma classe. Assim, para algo ser considerado ouro é uma condição necessária que seja composto pelo tipo de *substância química de número atômico 79*; e se algo é composto pela *substância química de número atômico 79*, então isso é condição suficiente para que este algo seja ouro. Mas

---

*Entendimento Humano*. Tradução de Anoar Aiex e E. Jacy Monteiro, 2. Ed., São Paulo: Abril Cultural, 1978, Livro III, Capítulo VI, §1-3, p. 240.

<sup>145</sup> *Ibidem*, §2.

<sup>146</sup> Cf. MILLER, 2011, *Op. Cit.*, p. 163.

algumas teorias semânticas recentes tem explorado a possibilidade de definições para tipos naturais que não estabelecem as condições necessárias e suficientes para um indivíduo ser considerado um membro de uma classe<sup>147</sup>. Em *How to be a Moral Realist*, Richard Boyd argumenta que existem vários termos cuja definição depende de um conjunto de propriedades, diferentes entre si, que, agrupadas por mecanismos que as unificam, constituem a definição de alguns tipos naturais.

[...] alguns termos têm definições que são determinadas por uma coleção de propriedades tal que a posse de um número adequado dessas propriedades é suficiente para recair sob a extensão do termo. [...] É caracteristicamente insistido, portanto, que nossos conceitos de tais tipos são “abertos” e há uma indeterminação na extensão *legitimamente* associada com as definições sobre o agregado de propriedades ou atribuição de critérios. A “imprecisão” ou “vagueza” de tais definições é vista como uma característica perfeitamente apropriada do uso linguístico ordinário [...]<sup>148</sup>.

Assim, um termo depende, não de *uma* propriedade que é a sua extensão, mas de um *conjunto* de propriedades que *constituem* a sua extensão e, por conseguinte, formam a sua definição. Essas propriedades são unificadas por relações nomológicas que determinam que tais e tais indivíduos fazem parte do agregado. As relações nomológicas são descobertas através de investigação científica. Portanto, não temos uma definição precisa de quais propriedades constituem o agregado. Através das descobertas empíricas o agregado vai se constituindo e assim formamos a nossa definição natural. Esse tipo de definição não apresenta condições suficientes para a extensão de um termo, pois sempre podemos descobrir novas propriedades que são nomologicamente unificadas ao agregado. Boyd denomina essas

---

<sup>147</sup> Cf. BOYD, R. *How to be a Moral Realist*. In SAYRE-MCCORD, G. (ed) *Essays on Moral Realism*, Ithaca: Cornell University Press, 1988, p. 322 e BOYD, R. *Metaphor and Theory Change*. In ORTONY, A. (ed), *Metaphor and Thought*, Cambridge: Cambridge University Press, 1979, p. 484.

<sup>148</sup> BOYD, 1988, *Op. Cit.*, p. 322.

definições naturais de *agregado de propriedades homeostáticas*. Tentemos compreender o que tudo isso significa por partes.

Vamos considerar os dois exemplos seguintes. O primeiro é sobre o termo ‘metal’. Seja no nosso discurso comum ou no científico usamos o termo ‘metal’ para expressar uma variedade de propriedades. Há um agregado de propriedades que são características dos indivíduos que compõe a extensão desse termo tais como condutibilidade, ductilidade, maleabilidade, brilho, etc. Frequentemente, dizemos que algo é metal quando apresenta essas propriedades. Além disso, essas propriedades são co-instanciadas por diferentes indivíduos. A prata, o ouro, o alumínio, o ferro, o chumbo, etc., são todos instâncias de ‘metal’. Além disso, deve haver alguns mecanismos que *unificam* todas essas propriedades sob o tipo ‘metal’. Como sugerido por Rubin<sup>149</sup>, um bom candidato pra isso é a propriedade de *ser composto por átomos com caráter metálico*. Assim, se algo apresenta a propriedade de ser composto por átomos com caráter metálico, ou tendência a perder elétrons, este algo é metal seja prata, ouro, ferro ou qualquer outra instância. A minha definição de ‘metal’ não é precisa ou exaustiva, pois sempre pode acontecer de descobrirmos um novo material que apresente a propriedade de ser composto por átomos com caráter metálico e, portanto, fazer parte do aglomerado de indivíduos que constituem a definição de ‘metal’.

O segundo exemplo que gostaria que considerássemos é sobre o termo ‘tigre’. De acordo com Boyd, há um agregado de propriedades que definem ‘tigre’. Constituição morfológica, constituição física e constituição comportamental podem ser três delas<sup>150</sup>. Para cada uma dessas propriedades temos um conjunto específico de outras propriedades. Por exemplo, para a constituição morfológica e física podemos ter a propriedade de ter quatro patas, uma cauda, duas orelhas, dentes com tamanho tal, um coração, dois rins, estômago, etc. Para a constituição comportamental, a propriedade de ser caçador, agressivo, viver em bandos, etc. Hoje sabemos que existe uma variedade de indivíduos diferentes que, todos eles, co-instanciam essas propriedades. Temos, por exemplo, o tigre de bengala, o tigre siberiano, o tigre de Sumatra, o tigre de Bali e vários outros. E para que possamos reunir todos eles sob a classe ‘tigre’ deve haver mecanismos que unificam todos esses indivíduos. Boyd propõe que *intercâmbio genético entre*

---

<sup>149</sup> Cf. RUBIN, M. Synthetic Ethical Naturalism. University of Massachusetts, *Dissertations*, Paper 24, 2009, p. 129.

<sup>150</sup> Cf. BOYD, 1988, *Op. Cit.*, p. 324.

*certas populações e isolamento reprodutivo de outras* pode ser o mecanismo unificador para o tipo tigre. Isso nos permite englobar todos os indivíduos que apresentam esta propriedade sobre a espécie tigre.

Nesse sentido, podemos dizer que há um grupo de propriedades F, ‘metal’ e ‘tigre’ para considerarmos os nossos exemplos, que são agregados de forma contingente e que co-ocorrem em vários casos<sup>151</sup>. Boyd utiliza o termo ‘homeostase’ para se referir a esse grupo de indivíduos que compõem a definição de um termo. Ele diz que o uso é apenas metafórico<sup>152</sup>, todavia não apresenta uma definição mais precisa. Mas, à luz das características de definições a partir de agregados de propriedades apresentadas abaixo, podemos concluir que sua intenção é de expressar a ideia de um conjunto de propriedades que compartilham suas qualidades constitutivas com um todo maior e por isso essas propriedades são unificadas em grupos.

Consideremos, agora, duas características relevantes sobre agregados de propriedades homeostáticas elencadas por Boyd. Quando temos um conjunto de propriedades homeostáticas F,

Ou a presença de algumas das propriedades em F tendem (sob condições apropriadas) a favorecer a presença de outras, ou há mecanismos e processos subjacentes que tendem a manter a presença das propriedades em F, ou ambos<sup>153</sup>.

Podemos dizer que essas são condições para termos um agregado de propriedades homeostáticas. É possível tornar essas condições mais explícitas dizendo que F é um agregado de propriedades homeostáticas se e somente se:

- (i) a presença de algumas das propriedades em F (sob condições apropriadas) tende a favorecer a presença de outras;
- (ii) há mecanismos ou processos subjacentes que tendem a manter a presença das propriedades em F;
- (iii) ambos, (i) e (ii)<sup>154</sup>.

---

<sup>151</sup> *Ibidem*, p. 323.

<sup>152</sup> *Idem*.

<sup>153</sup> *Idem*.

<sup>154</sup> Apesar de manter algumas diferenças, a ideia de transformar a citação de Boyd em três princípios mais explícitos é baseada em Rubin. Cf. RUBIN, 2009, *Op. Cit.*, p. 125.

Lembremos dos exemplos acima. No caso do agregado de propriedades homeostáticas ‘tigre’ a presença de determinadas propriedades comportamentais como ser caçador, agressivo, caçar em bandos, ser veloz, etc., tende a favorecer a presença de propriedades físicas, como dentes grandes, garras afiadas, faro acurado, etc. A ductilidade claramente favorece a presença da maleabilidade, no caso de ‘metal’. Além disso, há mecanismos que mantêm a presença das propriedades unificadas em F, tais como *intercâmbio genético entre certas populações e isolamento reprodutivo de outras*. Desse modo, um agregado de propriedades homeostáticas sempre pode abarcar mais indivíduos na medida em que as propriedades presentes tem uma tendência a atrair outras propriedades. Isso não quer dizer que os indivíduos que compõem o agregado tenham que ser idênticos, basta que compartilhem um conjunto relevante de propriedades.

Como vimos, há um termo *t* que é aplicado ao agregado homeostático (‘tigre’, ‘metal’). Ora, *t* não pode ter uma definição analítica<sup>155</sup>. Isso porque é o agregado de propriedades unificado pelo mecanismo comum a todas as propriedades que forma a definição *natural* de *t*. A extensão do agregado de propriedades homeostáticas é sempre determinada *a posteriori*, portanto é sempre uma questão em aberto. Na medida em que descobrimos novos indivíduos que compartilham as propriedades e os mecanismos que unificam o agregado de propriedades, expandimos a extensão do termo para esses indivíduos, mas não há a exigência de que todos os membros do agregado apresentem uma uniformidade no sentido estrito. É por isso que agregados homeostáticos imperfeitos são possíveis, pois um indivíduo pode fazer parte do agregado F e mesmo assim não apresentar todas as propriedades de F. *x* pode ser um metal, por exemplo, apresentar maleabilidade e não apresentar brilho. *y* pode ser um tigre e ter a constituição física comum a todos os indivíduos do agregado e não partilhar de algumas propriedades da constituição comportamental, como viver em bandos.

Dado que há mecanismos que unificam o agregado de propriedades homeostáticas, os membros que constituem o agregado apresentam um grau de uniformidade. Se apresentam uniformidade, então podemos dizer que é confiável fazermos induções e explicações a partir do agregado<sup>156</sup>. No exemplo do ‘tigre’, o mecanismo que unifica

---

<sup>155</sup> Cf. BOYD, 1988, *Op. Cit.*, p. 323.

<sup>156</sup> Cf. *Ibidem*, p. 322.

todos os indivíduos que co-instanciam as propriedades que constituem este agregado é o *intercâmbio genético entre certas populações e isolamento reprodutivo de outras*. Todos os indivíduos observados apresentam esta característica comum que os unifica. Então é razoável supor que os não observados também irão apresentar esta propriedade unificadora e que farão parte do agregado homeostático em questão. Nesse sentido, agregados de propriedades homeostáticas são explicativos, pois a condição para podermos dizer que determinado indivíduo é um membro do agregado  $x$  é este indivíduo apresentar a propriedade que unifica o agregado; e é o próprio agregado que estipula quais indivíduos são abarcados pela característica unificadora. Assim, dizemos que os agregados são explicativos porque nos informam que tais e tais indivíduos são abarcados pelo mecanismo unificador e que, portanto, fazem parte de tal espécie ou tal definição.

Uma vez que o agregado é constituído *a posteriori*, as propriedades unificadas pelo mecanismo homeostático são tipos naturais. Como sugere Boyd, “a importância causal do agregado de propriedades homeostáticas  $F$  associado ao mecanismo homeostático subjacente relevante é tal que o tipo ou propriedade denotada por  $t$  é um tipo natural”<sup>157</sup>. Dado que o agregado homeostático é constituído por propriedades naturais, sua definição é sempre passível de revisão. Na medida em que aparecem novas evidências ou novos desenvolvimentos científicos, nossa capacidade para estudar as propriedades naturais é aperfeiçoada. Isso significa que sempre poderemos descobrir novos indivíduos que co-instanciam o mecanismo homeostático e, portanto, também fazem parte da definição do agregado. Para retornar aos exemplos, é sempre possível descobrir uma nova variedade da espécie ‘tigre’ ou um novo elemento que co-instancia a propriedade de ser ‘metal’.

Como temos visto, uma pressuposição evidente para agregados de propriedades homeostáticas é o abandono da teoria semântica descritivista. Boyd concorda com Putnam e Kripke. O que o que fixa a referência não são os estados mentais ou as descrições que os indivíduos associam ao termo, mas as propriedades naturais que co-instanciam os mecanismos que unificam o agregado homeostático. Portanto, a referência de um termo  $t$  é estabelecida por conexões causais que determinam que tais e tais objetos fazem parte da extensão de  $t$ . Que o predicado  $t$  possa ser aplicado ao objeto  $x$  é determinado por uma cadeia

---

<sup>157</sup> *Ibidem*, p. 323.

causal independente do indivíduo que usa o termo<sup>158</sup>. Consideremos a formulação de Boyd:

*Grosso modo*, e para casos não degenerados, um termo *t* refere um tipo (propriedade, relação, etc) *k* quando há mecanismos causais cuja tendência é acarretar, através do tempo, que o que é predicado do termo *t* será aproximadamente verdadeiro de *k* [...]. Tais mecanismos incluirão tipicamente a existência de procedimentos que são aproximadamente acurados para reconhecer membros ou instâncias de *k* (ao menos para casos simples) e que governam relevantemente o uso de *t*, a transmissão social de certas crenças relativamente aproximadas sobre *k* formuladas como alegações sobre *t* [...], um padrão de deferência para os competentes sobre *k* no que diz respeito ao uso de *t*, etc. [...]<sup>159</sup>.

Terence Horgan e Mark Timmons usam a expressão ‘tese sobre a regulação causal’ para se referir à teoria da referência endossada por Boyd<sup>160</sup>. Assim, podemos dizer que *k regula causalmente* o uso de *t*, em que *k* é a propriedade natural e *t* é o predicado usado para se referir a *k*. Portanto, o que regula o nosso uso de ‘metal’ nos autorizando a aplicar este termo a tais e tais objetos são as propriedades naturais que, por apresentarem os mecanismos causais que os agregam num todo homeostático, determinam o que é e o que não é extensão deste termo. Para citar um exemplo mais simples, o que regula o meu uso de ‘água’ para aplicá-lo ao objeto *x* não é a minha intenção de aplicar o termo a *x*, mas é o fato de *x* apresentar a fórmula H<sub>2</sub>O. Mas para agregados de propriedades homeostáticas, como afirma Boyd, aquilo que é predicado por *t* é apenas aproximadamente verdadeiro de *k*. Isso porque a definição do agregado é sempre passível de revisão e não apresenta condições suficientes para que um indivíduo faça parte do agregado. Podemos descobrir, por exemplo, que o material *x*, que considerávamos como uma co-extensão de ‘metal’, a despeito de ser formado por uma

---

<sup>158</sup> *Ibidem*, p. 321.

<sup>159</sup> *Idem*.

<sup>160</sup> HORGAN, T. & TIMMONS, M. Troubles for New Wave Moral Semantics: The Open Question Argument Revived. *Philosophical Papers*, Volume 21, No. 3, 1992b, p. 153-175, p. 158-159.

propriedade que é muito similar à propriedade que unifica os diferentes metais sobre o agregado de ‘metal’, na verdade não é extensão de ‘metal’, pois é formado por uma propriedade bastante peculiar até então desconhecida. Por isso, temos de admitir a possibilidade de que as propriedades que são predicadas por *t* são aproximadas de *k*.

Essa é a teoria semântica externalista apresentada por Boyd. Vejamos agora sua aplicação na ética.

## 2.2. Aplicação às propriedades morais

Boyd defende que ‘*bom*’<sup>161</sup> é um agregado de propriedades homeostáticas<sup>162</sup>. Na sua visão, ‘*bom*’ é constituído por um conjunto de propriedades que são agrupadas por aquilo que ele denomina de mecanismos unificadores. Se isso é verdadeiro, então não temos uma definição rigorosa que estabeleça precisamente a extensão desse termo. Estamos a descobrir, empiricamente, quais propriedades são instâncias de ‘*bom*’. Boyd sugere que ‘*bom*’ moral corresponde às “coisas que satisfazem necessidades humanas importantes”<sup>163</sup>. Obviamente, há um número indeterminado e crescente de coisas que satisfazem as necessidades humanas. Boyd nos fornece uma pequena lista:

[...] algumas dessas necessidades são físicas ou médicas. Outras são psicológicas ou sociais; essas (provavelmente) incluem a necessidade por amor e amizade, a necessidade de envolver-se em esforços cooperativos, a necessidade de exercer o controle sobre a própria vida, a necessidade para apreciação e expressão intelectual e artística, a necessidade de lazer, etc<sup>164</sup>.

Podemos imaginar inúmeras instâncias que satisfazem esses grupos de necessidades humanas, tais como ser autônomo, ser fisicamente saudável, ser mentalmente saudável, ter amigos, criar e apreciar algum tipo de arte e assim por diante. À medida que evoluímos, seja cultura, econômica, política ou numericamente, novas necessidades surgem. O aumento populacional e a crescente exploração dos bens naturais

---

<sup>161</sup> Estou usando o predicado ‘*bom*’ para se referir as propriedades morais em geral.

<sup>162</sup> Cf. BOYD, 1988, *Op. Cit.*, p. 331.

<sup>163</sup> *Ibidem*, p. 329.

<sup>164</sup> *Idem*.

levaram os governos a investir mais na criação de instituições voltadas à preservação do meio ambiente, por exemplo. Segundo Boyd, é assim que vamos descobrindo novas instâncias do ‘*bom*’. O fato é que existem muitas – ou, como sugere Brink, talvez infinitas – instâncias em que o ‘*bom*’ é realizado. E a questão sobre quais ‘*bens*’<sup>165</sup> há, é uma questão empírica difícil<sup>166</sup>. Tal como a definição de tipos naturais é imprecisa, não exaustiva, passível de revisão, etc., a definição do agregado de propriedades que constituem o ‘*bom*’ é aperfeiçoada ou complementada à medida que descobrimos novos ‘*bens*’.

Parece que ‘*bom*’ satisfaz as duas condições para ser um agregado de propriedades homeostáticas.

[...] esses bens humanos (ou melhor, as instâncias que os satisfazem) são homeostaticamente agregados. Em parte, eles são agregados porque esses próprios bens – quando presentes em equilíbrio ou moderação – *se suportam mutuamente*. Além disso, há *mecanismos* psicológicos e sociais que quando, e na medida em que, estão presentes, contribuem para a homeostase. [Esses mecanismos] provavelmente incluem o cultivo de atitudes de respeito mútuo, democracia política, relações sociais igualitárias, diversos rituais, costumes e regras de cortesia, fácil acesso à educação e informação, etc<sup>167</sup>.

Dizer que o agregado de propriedades homeostáticas se suporta mutuamente, significa que a presença de algumas instâncias do ‘*bom*’ tende a favorecer a presença de outras. A realização de alguns ‘*bens*’, como acesso fácil ao cuidado médico, por exemplo, suporta vários outros, tais como ser física ou mentalmente saudável, ter autonomia sobre a própria vida e assim por diante. O mesmo é válido para qualquer outra instância de ‘*bom*’. É importante observar que Boyd não está pensando no suporte mútuo das instâncias de ‘*bom*’ para um único indivíduo. As propriedades que compõem o agregado homeostático são instâncias do ‘*bom*’ entre as pessoas ou entre comunidade em geral<sup>168</sup>. Se se tratasse apenas da vida de um único indivíduo, então a presença de

---

<sup>165</sup> Uso ‘*bens*’ como plural de ‘*bom*’.

<sup>166</sup> Cf. *Idem*.

<sup>167</sup> *Idem*. Itálico meu.

<sup>168</sup> Cf. *Ibidem*, p. 353.

algumas propriedades não tenderia a favorecer a presença de outras (ou tenderia num contexto limitado, somente para o indivíduo em questão, o que não faz parte da ideia de Boyd). Então, satisfação de necessidades básicas, acesso à educação e distribuição econômica igualitária, por exemplo, tendem a atrair outros ‘bens’ para a sociedade como um todo.

Nos exemplos de ‘metal’ e ‘tigre’, havia mecanismos bem definidos para unificar os indivíduos em agregados homeostáticos. O *intercâmbio genético* entre *certas populações e isolamento reprodutivo de outras* unificava todos os objetos sob o agregado ‘tigre’. Como Boyd reconhece, no caso de ‘bom’ há diversos mecanismos homeostáticos que unificam grupos de instâncias do ‘bom’. Quer dizer, diferentemente de um mecanismo unificador para todo o agregado, o ‘bom’ possui vários agregados internos. Ele cita mecanismos como *cultivo de atitudes de respeito, relações sociais igualitárias, regras de cortesia*, etc. Para cada um desses mecanismos, podemos imaginar inúmeras ações que os instanciam. Consideremos, por exemplo, o mecanismo de relações sociais igualitárias. Há um conjunto indeterminado de ações que são unificadas por este mecanismo, como o tratamento imparcial de um grupo de pessoas numa seletiva de emprego, a não discriminação por raça ou sexo, o dar a oportunidade de todas as pessoas envolvidas num processo de tomada de decisão argumentarem a seu favor, etc. Regras de cortesia podem agrupar ações de ser sincero, dar atenção aos outros, ser comprometido com o trabalho, etc. É importante salientar que esses mecanismos unificadores não agregam regras ou normas de conduta, mas sim instâncias de ‘bom’, quer dizer, ações em que ‘bom’ é *realizado*. Enquanto no caso de ‘tigre’ o mecanismo homeostático unifica somente os indivíduos da espécie tigre, no caso de ‘bom’ há mecanismos sobre relações pessoais, relações sociais, comportamento dos indivíduos e assim por diante. Como o próprio Boyd concorda, essa é uma diferença do ‘bom’ para com os agregados da biologia ou química<sup>169</sup>.

Considerando que o agregado de propriedades homeostáticas para ‘bom’ aparentemente não apresenta apenas *um* mecanismo unificador de todas as suas instâncias, mas vários, é bastante razoável se perguntar o que faz com que estes diferentes mecanismos sejam agrupados sob o agregado ‘bom’. A sugestão de Brink é interessante. Como vimos, Brink concorda com Boyd que há um agregado de propriedades que constituem o ‘bom’. Por isso, diz que o ‘bom’ é superveniente a este

---

<sup>169</sup> Cf. *Ibidem*, p. 353-354.

agregado de propriedades. A nossa pergunta, então, é sobre o que reúne todas essas propriedades sob o agregado ‘*bom*’. Brink diz o seguinte.

Quais propriedades morais supervêm fortemente sobre quais propriedades naturais é determinado diferentemente por diferentes teorias morais [...]. O utilitarismo hedonista, por exemplo, pode ser construído como fazendo uma reivindicação naturalista; ele alega que correção<sup>170</sup> = a maximização do prazer. [...] A determinação de quais propriedades e fatos naturais constituem quais propriedades e fatos morais é uma questão de teoria moral substantiva<sup>171</sup>.

Isso significa que a determinação do mecanismo que agrega todas as propriedades que constituem ‘*bom*’ é dependente da teoria normativa em questão. A definição da propriedade moral nos é dada sinteticamente e o papel da teoria normativa é similar ao papel desempenhado pela química ou biologia, por exemplo, em fornecer definições a partir das propriedades reais dos objetos. Os químicos foram responsáveis pelo descobrimento de que água é H<sub>2</sub>O. As melhores teorias normativas tentam nos fornecer as melhores definições para o agregado de propriedades que constituem ‘*bom*’. Certamente, diferentes teorias normativas concordam sobre quais propriedades constituem o ‘*bom*’ no âmbito de uma variedade superficial e comum de instâncias. Por exemplo, o utilitarismo hedonista poderia concordar com uma teoria normativa kantiana que *a*, *b* e *c* são instâncias do ‘*bom*’. Mas em casos mais específicos, poderíamos obter respostas diferentes. Uma teoria de viés kantiano poderia sustentar que o ato de mentir não é uma instância do ‘*bom*’ enquanto que uma teoria utilitarista poderia defender que é (desde que contribua para a maximização do prazer ou minimização da dor)<sup>172</sup>. Assim, podemos dizer que o mecanismo que unifica as

---

<sup>170</sup> ‘Correção’ não se refere ao ato de corrigir ou àquilo que é correto, mas à propriedade moral. Em inglês o termo é ‘*rightness*’.

<sup>171</sup> BRINK, 1989, *Op. Cit.*, p. 175-178.

<sup>172</sup> Consideremos que haja desacordo entre ambas as teorias normativas em relação à mentira apenas para o bem do exemplo. Se o utilitarismo e o kantismo realmente apresentam respostas diferentes ou contrárias às mesmas ações é uma discussão profunda da qual não gostaria de me ater aqui, pois não é este o ponto.

propriedades que constituem o bom pode ser, por exemplo, *a maximização do prazer*.

Se poderia perguntar: por que as teorias normativas muitas vezes diferem sobre quais objetos instanciam o agregado de ‘*bom*’? A resposta é que é uma questão empírica difícil saber quais propriedades pertencem ao agregado e quais não pertencem. É por isso que a nossa definição é constituída *a posteriori*. O agregado de propriedades homeostáticas é “aberto”, por assim dizer, pois sempre é possível incluir novos objetos que fazem parte dos mecanismos que unificam o agregado. É por isso que um naturalista poderia muito bem dizer ‘E daí?’ para alguém que lhe apresentasse o argumento da questão aberta de Moore.

Boyd defende que ‘*bom*’ é *definido* pelo agregado de propriedades e pelos mecanismos que unificam estas propriedades<sup>173</sup>. Se é definido pelo agregado, então é bom tudo aquilo que pertence ao agregado. Seguindo Rubin, podemos dizer que

BOM: *x* é bom se, e somente se, *x* instancia o agregado de propriedades homeostáticas de ‘*bom*’<sup>174</sup>.

‘*Bom*’, então, abarca um aglomerado inumerável de propriedades que são a sua instância. O que garante que a escravidão não seja um ato ‘*bom*’, por exemplo, é que esta propriedade não pertence ao agregado homeostático de ‘*bom*’, e isso nos é informado pelas melhores teorias normativas.

Se são as teorias normativas que determinam qual é a extensão de ‘*bom*’, então é uma questão independente de nossas mentes a que tipo de objetos, ações, etc., esta propriedade moral se aplica, argumentam os naturalistas. As ações que compartilharem dos mecanismos unificadores do agregado de ‘*bom*’ serão instâncias do ‘*bom*’. Estas instâncias são a referência da propriedade moral. Como sugere Boyd, a “referência é em si uma noção epistêmica”<sup>175</sup>, quer dizer, vai sendo fixada à medida em que as melhores teorias nos proporcionam uma explicação aproximada de quais propriedades a constituem. Este é um desenvolvimento da teoria semântica externalista para a moralidade.

Agora, para finalizar esta seção, gostaria de dizer em que sentido a teoria de Boyd responde duas famosas objeções ao naturalismo. Já vimos essas objeções na primeira parte deste trabalho e, para não

---

<sup>173</sup> BOYD, 1988, *Op. Cit.*, p. 329.

<sup>174</sup> Cf. RUBIN, 2009, *Op. Cit.*, p. 134.

<sup>175</sup> BOYD, 1988, *Op. Cit.*, p. 321.

ficarmos apenas com a apresentação da teoria de Boyd, vejamos algumas de suas supostas vantagens. A primeira é a ideia de que a moralidade é construída através de deliberação racional, acordos, escolhas das melhores alternativas, etc., isto é, a tese sustentada pelos construtivistas. Segundo esta abordagem, não há fatos morais a serem descobertos, mas construções conjuntas sobre o que é o certo e o errado. Boyd afirma que para descobrirmos novos fatos morais certo desenvolvimento da moralidade é necessário. O construtivista pode argumentar, por exemplo, que do fato de a moralidade estar num estado  $x$  de desenvolvimento propiciar a descoberta dos fatos morais  $a$ ,  $b$  e  $c$  implica que as propriedades morais são construções sociais. Ele poderia dizer que essa dependência da moralidade atual para o desenvolvimento posterior mostra justamente que as propriedades morais são construídas socialmente. Isso porque, ele poderia argumentar,  $x$ ,  $y$  e  $z$  (fatos morais quaisquer) só foram aceitos porque os princípios  $p$ ,  $q$  e  $r$  foram anteriormente aceitos. Não tivessem as pessoas endossado  $x$ ,  $y$  e  $z$ , mas  $k$ ,  $l$  e  $m$ , então  $p$ ,  $q$  e  $r$  não seriam parte daquilo que consideramos ‘*bom*’. A principal razão para a adoção de uma posição construtivista é a ideia de que nas ciências é a observação que desempenha o papel do raciocínio científico enquanto que na moral é a intuição. Segundo este modelo argumentativo, nós não observamos propriedades morais, de modo que não podemos dizer que se  $x$  é ‘*bom*’ ou não olhando para o mundo externo, mas apenas testando os juízos morais através de nossas intuições morais sobre o que é ‘*bom*’.

Para vermos a resposta do naturalismo, gostaria de falar um pouco mais sobre o que significa dizer que a definição de ‘*bom*’ é sintética. Como temos visto, ao argumentar que a descoberta das instâncias de ‘*bom*’ depende da teoria normativa em questão, Boyd está fazendo uma analogia com o raciocínio científico. Segundo ele, é preciso levar a sério esse tipo de analogia. Certamente, o conhecimento científico é cumulativo. Isso significa que as descobertas feitas num tempo  $t$  dependem de avanços e descobertas feitas em um tempo  $t'$ , que é anterior a  $t$ . Se olharmos para a história da ciência vemos que os avanços dependem de condições sociais adequadas, desenvolvimentos tecnológicos, desenvolvimentos científicos anteriores, etc. Considere, por exemplo, o desenvolvimento da física e química no século XVIII. Só foram possíveis devido a fatores como, o trabalho nestas ciências tornou-se economicamente incentivado de modo que as pessoas podiam se dedicar exclusivamente às suas pesquisas, desenvolvimento suficiente da tecnologia para propiciar os instrumentos necessários para as descobertas, o próprio conjunto das descobertas anteriores que

constituem o estado atual do conhecimento, etc<sup>176</sup>. Em uma palavra, o estado da ciência em determinada época contribui para novas descobertas.

O raciocínio moral é similar a isso. De acordo com Boyd, a descoberta das instâncias de ‘*bom*’ depende do estado atual do desenvolvimento moral. Como aponta Boyd, não teríamos compreendido (quer dizer, descoberto) que a democracia política faz parte do agregado de propriedades homeostáticas de ‘*bom*’ se algumas condições empíricas não fossem satisfeitas, tais como as condições que possibilitaram o desenvolvimento das primeiras e limitadas formas de democracia. A descoberta de uma instância de ‘*bom*’ depende tanto do estado moral atual quanto de fatores externos, como a organização social que proporciona a descoberta dessas propriedades. Assim como na ciência, à medida que algumas descobertas morais são feitas, outras são possíveis. Quando se descobriu, por exemplo, que a igualdade é uma instância de ‘*bom*’, e se começou a dar consideração aos interesses de todos os indivíduos, se começou a perceber o quão errada era a prática da escravidão. Assim, descobriu-se outra instância da moralidade, uma instância da ‘*incorreção*’, neste caso. Pode-se sugerir que isso, de alguma forma, teve influências para a descoberta de que se deve dar peso igual ao interesse de todas as pessoas no que diz respeito à decisões políticas, contribuindo assim para a descoberta de que a democracia é uma instância de ‘*bom*’. O ponto que quero ressaltar não é que as descobertas dos fatos morais se deram nesta ordem, mas a analogia do naturalismo com a ciência. Quer dizer, a moralidade é cumulativa, evolui e as descobertas morais de uma época dependem do estado de desenvolvimento moral atual dessa época. Ou, ao menos é nesse sentido que argumenta Boyd.

Nesse sentido, o naturalista não precisa ceder ao construtivista e dizer que são as nossas intuições sobre a moralidade que determinam o que é bom. O naturalista pode argumentar que, na moralidade, o que desempenha o papel que é desempenhado pela observação na ciência é, também, a observação. Se ‘*bom*’ é um conjunto de propriedades homeostáticas (naturais), então para sabermos se um juízo moral é verdadeiro ou falso basta olharmos para o mundo e vemos se aquilo a que se está atribuindo tal propriedade moral realmente tem tal propriedade. Se sim, então o juízo é verdadeiro; se não, não. Isso, obviamente, irá depender da teoria normativa que assumirmos. Considere que assumo uma teoria utilitarista, por exemplo. Basta

---

<sup>176</sup> *Ibidem*, p. 330.

investigarmos, quer dizer, observarmos, o mundo e constatarmos quais ações tem a propriedade de maximizar a felicidade e quais não tem. As ações que tiverem esta propriedade serão instâncias do ‘*bom*’.

A segunda objeção sustenta que não há fatos morais, pois se existissem não haveria tantos desacordos em relação aos mais diversos assuntos morais. Se assumirmos a teoria de Boyd, podemos responder a esse problema de diversas formas. Em primeiro lugar, podemos dizer que o agregado de propriedades homeostáticas falha na bivalência. Quer dizer, não é uma questão certa que a atribuição de uma propriedade moral a uma ação, pessoa, instituição, etc., será claramente verdadeira ou falsa e as pessoas poderão perceber isso no mesmo instante. Na verdade, o agregado de propriedades homeostáticas até mesmo requer que haja respostas divergentes ou, inclusive, conflitantes. Isso porque, como vimos, o agregado é “vago”, “impreciso” e sua extensão é “indeterminada”. É uma questão empírica difícil e que exige acúmulo de conhecimento moral o descobrir quais propriedades pertencem ao agregado de ‘*bom*’. Se é assim, então certamente haverá pessoas reivindicando posições divergentes ou conflitantes sobre a extensão do agregado<sup>177</sup>.

Em segundo lugar, sabemos que as pessoas não adotam unanimemente a mesma teoria normativa para identificar os objetos que constituem o agregado de propriedades homeostáticas. Teorias diferentes podem oferecer respostas diferentes para a questão de se um determinado objeto é ou não uma instância do ‘*bom*’.

Além disso, como temos visto, o naturalista assume que o conhecimento moral é cumulativo. Diferentes culturas podem ter diferentes graus de conhecimento moral e de um sistema de crenças morais aproximadamente verdadeiro. Isso gera divergências no desenvolvimento moral das culturas. Vamos considerar um exemplo. Imaginemos três culturas diferentes, A, B e C. Assumamos, para o propósito do argumento, que é um fato moral que o aborto é permissível, em qualquer circunstância, para os seis primeiros meses de gestação. A, B e C têm diferentes graus de aproximação em relação a esta verdade moral. De acordo com A, o aborto é permissível, exceto em casos que não houver uma causa relevante que o justifique. Assim, o aborto é permissível nos casos em que houve violência sexual, em que as condições econômicas dos pais não propiciariam uma educação minimamente adequada ao futuro filho, etc. Os membros de A sustentam que o aborto não seria permissível em casos em que contasse

---

<sup>177</sup> Cf. *Ibidem*, p. 326, 327, 338, 339.

somente a arbitrariedade dos pais sem nenhuma causa adicional relevante. A cultura B, por outro lado, sustenta que o aborto é permissível em qualquer circunstância, havendo motivo relevante ou não. Basta haver consentimento dos pais e deve-se poder abortar uma vida. A única restrição é que isso vale somente para os quatro primeiros meses de gestação. Além disso, a cultura C sustenta que o aborto não é permissível *definitivamente*. Ora, parece que, nesse caso, as culturas A, B e C têm diferentes graus de aproximação da verdade moral. A cultura B certamente mantém uma visão mais próxima do fato moral de que o aborto é permissível. A cultura A encontra-se mais distante da verdade. Enquanto isso, nem sequer podemos falar em aproximação da verdade em relação à cultura C. Essa divergência inicial irá determinar desacordos morais, em primeiro lugar, e, em segundo, o desenvolvimento da moralidade.

Vimos que existem algumas definições para tipos naturais que não apresentam condições necessárias e suficientes para um objeto ser pertencente a sua classe, mas são constituídas por um conjunto de propriedades que juntas constituem a definição do termo. Tais definições são sempre “abertas”, no sentido de que é possível descobrirmos novos indivíduos que a constituem. Boyd argumenta que as propriedades morais possuem definições deste tipo. Se isso é verdadeiro, então as propriedades morais são multiplamente realizadas por um conjunto indefinido de propriedades naturais. Descobrimos quais propriedades naturais constituem as propriedades morais com base nas teorias normativas. Neste sentido, nossa melhor teoria normativa dará a mais aproximada resposta sobre quais são as propriedades naturais sobre as quais as propriedades morais supervêm. Essas definições homeostáticas são independentes de nossas mentes, pois o que determina quais propriedades constituem o agregado moral é o fato de as propriedades naturais serem agrupadas ou não pelos mecanismos unificadores. Esta teoria é, também, uma valiosa ferramenta para o naturalista responder algumas objeções frequentemente endereçadas à sua posição.

### 3. Explicações morais

Nesta seção, gostaria de voltar atenção para um problema mais específico a fim de apresentar outra estratégia usada pelo novo naturalismo para sustentar que há fatos morais e que estes pertencem ao mundo natural. A discussão que quero considerar teve origem com um problema apresentado por Gilbert Harman em seu *The Nature of*

*Morality*<sup>178</sup>. Harman faz a seguinte pergunta: os princípios morais podem ser testados e confirmados no sentido em que os princípios científicos podem? Ora, se Boyd está correto e os fatos morais são agregados de propriedades homeostáticas, então é certo que os princípios morais podem ser testados assim como os princípios científicos e não há necessidade de considerarmos este problema aqui, se poderia dizer. Mas, a discussão do problema de Harman deu origem a uma forma diferente de defender o realismo naturalista de maneira análoga à defesa do realismo científico baseando-se na inferência a favor da melhor explicação. Essa estratégia difere da adotada por Boyd. Nesse sentido, a fim de apresentarmos a defesa dos naturalistas, vale a pena a considerarmos aqui. Vou começar reconstruindo o problema estabelecido por Harman. Por um momento, irei desconsiderar as ideias dos naturalistas estabelecidas até aqui para o propósito de melhor compreendermos o ponto de Harman. Depois disso, pretendo apresentar as respostas a este problema, sugeridas por Sturgeon e Brink, e que caracterizam a posição naturalista.

### 3.1. “O problema com a ética”

Em geral, os princípios e teorias científicas podem ser testados e confirmados – ou recusados – analisando como o mundo realmente é. Na verdade, para o estabelecimento de um princípio ou teoria científica há um rigoroso processo de testes e experiências para determinar se aquilo que a teoria reivindica é verdadeiro. Se a teoria nos informa que “uma força é requerida para mudar o movimento de um corpo”<sup>179</sup> então, para ser verdadeira, isso deve acontecer no mundo e, neste caso, com regularidade universal. Além disso, se quisermos saber dentre duas teorias qual explica melhor a realidade, nós *olhamos* para o mundo e a que tiver um maior grau de sucesso explicativo será a melhor. Assim, parece que a *observação* desempenha um papel fundamental na ciência como critério para determinar a plausibilidade das teorias. Agora, considere o seguinte:

Podem os princípios morais serem testados no mesmo sentido, no mundo? Você pode observar

---

<sup>178</sup> HARMAN, G. *The Nature of Morality*. New York: Oxford University Press, 1977.

<sup>179</sup> LEVINSON, I. J. *Introduction to Mechanics*. Second Edition: Prentice-Hall, Englewood Cliffs, 1968, p. 14. Esta é a primeira lei de Newton.

alguém fazer algo, mas você pode alguma vez perceber a correção ou incorreção do que ele faz? Se você virar uma esquina e observar um grupo de baderneiros colocarem gasolina em um gato e incendiá-lo, você não precisa *concluir* que o que eles estão fazendo é errado; você não precisa descobrir nada exterior; você pode *ver* que isso é errado. Mas é a sua reação evidência à atual incorreção do que você vê ou é simplesmente a reflexão do seu “senso” moral, um “senso” que você adquiriu talvez como resultado da sua educação moral?<sup>180</sup>

Esta é uma pergunta difícil de responder. Parece que tudo o que podemos observar no caso descrito por Harman são um grupo de garotos, um gato, um material inflamável, e o ato de incendiar o animal; poderíamos citar a maneira como os garotos estão vestidos, a forma como se comportam, etc. Mas por mais que observemos atentamente o que eles estão fazendo parece que não conseguimos identificar a *incorreção* presente ali. No entanto, sabemos que o que eles estão fazendo é obviamente errado. Mas, assumindo que a incorreção não está *ali*, como podemos saber que o que eles estão fazendo é obviamente errado? Como podemos saber que o juízo moral ‘É errado incendiar gatos’ é verdadeiro? Deve-se poder deriva-lo de algum princípio mais geral como ‘É sempre errado infringir dor deliberadamente e cruelmente’, se poderia dizer. Mas como podemos garantir que este princípio é verdadeiro?

Pode-se dizer que os princípios morais são passíveis de teste e confirmação. Suponha que você esteja considerando o princípio de que ‘É sempre correto optar pela preservação da vida de cinco pessoas e pela morte de uma se você estiver numa situação em que você tem de escolher se salva uma e deixa cinco morrerem ou salva cinco e deixa uma morrer’<sup>181</sup>. Suponha agora que um médico tem de escolher se deixa um paciente morrer e salva a vida de cinco outros ou se dedica toda a sua atenção para o bem estar de um único paciente e deixa os outros cinco morrerem<sup>182</sup>. Parece muito mais razoável salvar a vida do maior

---

<sup>180</sup> HARMAN, *Op. Cit.*, p. 4.

<sup>181</sup> Cf. *Ibidem*, p.3.

<sup>182</sup> Para o bem do exemplo, assumamos que as únicas alternativas são cuidar dos cinco ou de um, não podendo cuidar de todos ao mesmo tempo e que se ele

número de pessoas e deixar apenas um morrer do que deixar que cinco morram enquanto um sobreviva. Neste sentido, parece que o princípio moral passou no teste e é verdadeiro, quer dizer, realmente é mais correto salvar a vida do maior número. Mas, é este princípio testado no mesmo sentido em que um princípio científico pode ser testado? Nesse caso fizemos uma experiência de pensamento apenas. Um cientista obviamente também faz experiências de pensamento, mas parece que pode ir muito além disso. Ele pode contrastar o seu princípio com experimentos reais, com como o mundo é. Parece que na moralidade nós não podemos ir tão longe. Mas quando fazemos reivindicações morais não queremos meramente expressar nossas opiniões, queremos mais do que isso. Queremos mostrar que *X* é realmente bom, que *Y* é a coisa certa a fazer, que *Z* é errado e não deve ser incentivado, e assim por diante. Quer dizer, nós normalmente tentamos convencer os outros de que determinadas ações, pessoas ou coisas tem a propriedade da *incorreção*, da *bondade*, da *justiça*, etc.

Se poderia dizer que não há observação na ética porque quando dizemos que uma ação é certa ou errada não vemos essas propriedades na ação, mas já pressupomos uma teoria sobre o *bom* de modo que a avaliação que fazemos será determinada de acordo com a nossa concepção moral prévia enquanto que na ciência as observações não são enviesadas pela teoria que assumimos. Mas, como Harman nota, “não há observações puras”<sup>183</sup>. Qualquer observação pressupõe teoria. Imagine que você está apenas observando o grupo de jovens baderneiros incendiando um gato, sem nenhum comprometimento moral de saber se é errado ou não aquilo que estão fazendo. Você verá apenas luz atingindo sua retina. Para dar sentido a isso, precisará de algum conhecimento sobre várias coisas como objetos, pessoas e aquilo que elas estão fazendo. Terá que saber sobre gatos, que tem a capacidade de sentir dor, que gasolina é inflamável, etc. Agora considere o aspecto moral. Você também precisa de algumas pressuposições teóricas prévias para poder inferir que o que os jovens estão fazendo é errado. Precisa saber que ações como provocar dor deliberadamente em pessoas ou animais são erradas, que atos de crueldade devem ser reprimidos, e assim por diante. Caso contrário, não terá razões para sustentar que é errado.

---

optar por cuidar de um os cinco irão morrer e se ele resolver cuidar dos cinco então estes irão sobreviver e um único indivíduo irá morrer.

<sup>183</sup> *Ibidem*, p. 4.

Além disso, normalmente apelamos para a observação a fim de dar suporte ao nosso ponto de vista moral. Suponha que, ao virar a esquina, você vê o grupo de jovens incendiando um gato, mas que, dado que não costuma dar muita atenção para o que as pessoas estão fazendo ao seu redor, você não vê a gasolina ou o fogo e acaba imaginando que eles na verdade estão brincando com o animal de modo tal que não estão causando nenhum sofrimento. Alguém pode tentar te convencer de que os jovens estavam fazendo algo errado argumentando que você não *observou bem* o que eles realmente estavam fazendo ou que se você tivesse *prestado atenção aos fatos* veria que a ação era errada. Isso significa que na ética, como na ciência, a observação parece ser importante para determinar se uma ação é justa ou injusta, se é correta ou reprovável. No entanto, parece estranho dizer para alguém apontando com o dedo ‘Olhe, bem ali, ali está a *incorreção* daquele ato!’. Isso nos faz suspeitar que o papel desempenhado pela observação na ética e na ciência é diferente. É justamente esta a tese de Harman. Ele diz que

[a] observação desempenha um papel na ciência que não parece desempenhar na ética. A diferença é que você precisa fazer suposições sobre certos fatos físicos para explicar a ocorrência das observações que suportam uma teoria científica, mas você não parece precisar fazer suposições sobre nenhum fato moral para explicar a ocorrência das chamadas observações morais que tenho falado. No caso moral, parece que você precisa fazer suposições apenas sobre a *psicologia* ou *sensibilidade moral* da pessoa que faz a observação moral. No caso científico, a teoria é testada contra o mundo<sup>184</sup>.

Para assumirmos uma posição científica, precisamos pressupor que a teoria ou os princípios que suportam a nossa posição são verdadeiros. Isso significa que foram contrastados com o mundo ou que foram adequadamente testados e confirmados, de modo que quando alguém perguntar pela validade de um princípio implícito na nossa argumentação podemos argumentar que há experiências que provam que o princípio é verdadeiro. Mas, no caso da moralidade, não parece que precisamos pressupor que os princípios morais que estão implícitos ou os juízos morais que fazemos são observações cuidadosas de fatos

---

<sup>184</sup> *Ibidem*, p. 6. Itálico meu.

morais ou que são proposições verdadeiras sobre o mundo. Parece muito mais que as posições morais que endossamos dependem daquilo que aprendemos na família que fomos educados, na sociedade em que fazemos parte, dos sentimentos morais que nutrimos por nós mesmos e pelos outros, dos valores que consideramos corretos e assim por diante. Para fazermos juízos morais não parece que precisamos de uma capacidade para captar certos fatos morais que são inerentes às ações ou pessoas.

Mas note que Harman usa a expressão ‘*explicar*’. Ele diz que na ciência, para explicar determinadas observações fatuais que dão suporte a uma teoria científica, é necessário fazermos suposições sobre certos fatos físicos enquanto que, na moral, para explicarmos porque fazemos tais e tais suposições morais (como a de que atear fogo em gatos é errado) não precisamos recorrer a fatos morais. Isso significa que os fatos físicos desempenham um papel explicativo nas observações que fazemos, mas que os fatos morais não desempenham um papel explicativo para as nossas avaliações, pois, como Harman afirma, a *melhor explicação* para o endossarmos a crença moral  $x$  se deve muito mais à nossa sensibilidade moral do que à existência de fatos morais. Este é o cerne do argumento de Harman, que passarei a explicar adiante. Mas, como vimos, acabei usando novas expressões como ‘papel explicativo’ e ‘melhor explicação’. Para compreendermos o ponto do argumento de Harman é necessário entender primeiro o que isso significa.

Quando Harman fala sobre objetos científicos ele não está pensando somente em coisas como, animais, seres humanos, plantas, organismos, montanhas, tempestades e rios e mares; quer dizer, não está pensando apenas em coisas que podemos observar seja a olho nu ou com a ajuda de algum dispositivo. Ele está pensando, principalmente, naquelas entidades que são inobserváveis. Mas como, se poderia dizer, ele está pensando que a observação desempenha um papel importante para analisar o comportamento de entidades que não podem ser observadas? É aí que entra em jogo uma estratégia argumentativa normalmente usada por aqueles filósofos denominados realistas em ciência. Eles defendem a tese de que as entidades científicas que não podemos observar, tais como átomos, elétrons, prótons, quarks, etc., são entidades reais e que existem independentemente de pensarmos sobre elas ou não. Para sustentar essa posição, argumentam que essas entidades desempenham um papel na *melhor explicação* do mundo e que, por isso, devem existir. O tipo de raciocínio empregado é denominado *inferência a favor da melhor explicação*. Consideremos,

primeiro, um exemplo cotidiano para tentarmos compreender o que isso significa.

Suponha que você acorda e vai até a cozinha<sup>185</sup>. Na noite anterior, você havia deixado uma fatia de queijo em cima da mesa. Chegando lá agora de manhã você percebe que não há fatia de queijo alguma, mas apenas algumas migalhas espalhadas. Você começa a pensar sobre o que pode ter acontecido enquanto estava dormindo, sobre o que *explica* o fato de o queijo ter desaparecido. No mesmo instante você lembra que ouviu alguns ruídos na cozinha durante a noite. Com isso, você conclui que havia um rato na cozinha e que ele comeu a fatia de queijo. Esta inferência certamente não é dedutiva, pois o queijo poderia ter desaparecido e você poderia ter ouvido os ruídos na cozinha e, mesmo assim, o queijo não ter sido comido por um rato. Mas é bastante razoável supor que um rato comeu o seu queijo. Você analisa outras hipóteses possíveis como, alguém ter entrado em sua casa e roubado o queijo deixando algumas migalhas para fazer você pensar que foi um rato, o gato do vizinho, etc. Mas no mesmo instante você conclui que estas alternativas não explicam *tão bem* o evento quanto a suposição de que um rato comeu seu queijo, pois seria muito difícil alguém ter entrado na sua casa durante a noite (muito menos para roubar apenas uma fatia de queijo) e que o seu vizinho na verdade não tem nenhum gato. Portanto, você conclui que realmente há um rato e que foi ele que comeu o seu queijo. Você concluiu isso a partir de uma inferência baseada na melhor explicação dos fatos.

Agora consideremos como os realistas científicos utilizam este tipo de raciocínio para sustentar que entidades como prótons ou elétrons são reais. Primeiro, vejamos um exemplo de como um físico procede no reconhecimento de partículas subatômicas. Em 1897 Charles Wilson descobriu um método de identificação dessas partículas e denominou-o Câmara de Nuvem (também conhecida como Câmara de Wilson). Ela

[...] consiste em uma câmara fechada com uma grande janela de vidro, de modo que seu volume pode ser iluminado e o que se passa em seu interior pode ser visto ou fotografado. Uma das paredes da câmara é um êmbolo móvel. A câmara contém vapor saturado, proveniente de uma

---

<sup>185</sup> Este exemplo é devido a Samir Okasha, com algumas modificações. Cf. OKASHA, S. *Philosophy of Science: a very short introduction*. Oxford: Oxford University Press, 2002.

quantidade em excesso de algum liquido [...]. Se o volume da câmara for subitamente aumentado pelo deslocamento do êmbolo, a expressão adiabática produzirá resfriamento que torna o vapor supersaturado e instável com tendência a condensar-se. Se houver íons na câmara, a condensação, de preferência, ocorrerá sobre eles. [...] *A câmara de Wilson possui elétrons que permitem que se mantenha um campo elétrico fraco de varredura. Se uma partícula de alta energia penetrar na câmara exatamente antes do momento da expansão do êmbolo torná-la sensível, não haverá tempo suficiente para os íons serem varridos [...] e as gotículas que se formam tornarão visíveis as posições dos íons. [...] Quando um elétron de alta energia atravessa essa câmara, o aspecto apresentado é o de uma linha reta fina semelhante a um rosário, ou “rastros” ou “trajetória”, que mostra muito bem para onde ele se dirige*<sup>186</sup>.

O que nos interessa aqui é a postulação, aparentemente óbvia no modo como o físico descreve, de que *existem* elétrons<sup>187</sup> a partir do fato de que há “rastros” ou “trajetórias” de vapor. O físico não visualizou os prótons, apenas os rastros. Mas ele conclui que o que causou os rastros foram os prótons. Ao notarem esse fenômeno, os realistas científicos inferem, a partir da observação dos rastros de vapor feita pelos físicos, que os prótons existem, mesmo que não seja possível observarmos. O raciocínio é bastante razoável, pois a pressuposição de que existem prótons *explica* os fenômenos que podemos verificar experimentalmente (a existência dos rastros). E parece que nenhuma outra explicação é melhor do que a de que existem prótons. Seria insensato concluir que alguma entidade celeste causou os rastros de vapor, por exemplo. Portanto, é muito mais razoável concluir que os prótons melhor explicam a nossa observação e que, por isso, são reais ou existem. Neste sentido, a existência dessas entidades inobserváveis é inferida a partir do sucesso explicativo que desempenham quando pressupomos que

---

<sup>186</sup> WHER, M. R. & RICHARD, J. A. *Física do Átomo*. Trad. de Carlos Campos de Oliveira. Ao Livro Técnico S. A.: Rio de Janeiro, 1965, p. 269-272.

<sup>187</sup> Falemos em prótons a partir de agora, pois no exemplo discutido por Harman e nas várias discussões de seu exemplo, costuma-se falar em prótons ao invés de elétrons.

existem. Não podemos as observar diretamente, mas podemos observar certos fenômenos físicos cuja melhor explicação para a sua ocorrência é que essas entidades existem. Portanto, elas existem. Ou, pelo menos assim argumentam os realistas em ciência.

O ponto de Harman consiste em argumentar que esse tipo de raciocínio não pode ser adotado para a moralidade. Quer dizer, não podemos postular que existem fatos morais com base na inferência a favor da melhor explicação. Isso porque os fatos morais não são a nossa melhor explicação do porquê sustentamos as posições morais que sustentamos. A melhor explicação para a sua crença de que o ato dos jovens incendiando um gato por mera diversão é errado não é a existência de um fato moral (a *incorreção*) que pertence à prática de atear fogo em gatos, argumenta Harman. Mas sim, a sua sensibilidade moral diante de tal prática.

Harman, neste sentido, está argumentado que as propriedades morais não desempenham um papel explicativo para a formação de nossas crenças morais e, por isso, não podemos concluir que são reais ou que existem como podemos fazer com as propriedades que postulamos na ciência. Seu argumento pode ser organizado, como fez Miller, da seguinte forma:

- (1) P é uma propriedade real se, e somente se, P desempenha um papel na melhor explicação da experiência.
- (2) Propriedades morais não desempenham nenhum papel na melhor explicação da experiência.
- (3) Propriedades morais não são propriedades reais<sup>188</sup>.

Para mostrar que propriedades morais não são explicativas, Harman sugere que pensemos sobre o exemplo da câmara de gás. Quando um físico vê um rastro de vapor em uma câmara de gás e conclui que há prótons ali ele está fazendo uma observação. Mesmo que não esteja  *vendo* os prótons diretamente, ele está observando os rastros de vapor. E a melhor explicação para a existência desses rastros é a suposição de que há prótons. Assumindo que a existência dos prótons seja confirmação de sua teoria, temos uma confirmação baseada na inferência a favor da melhor explicação. Notemos que o que melhor explica a crença de que há prótons na câmara de gás é a observação dos

---

<sup>188</sup> Cf, MILLER, A. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press, 2011, p. 141.

rastros de vapor e não determinado estado psicológico do cientista. Não parece ser uma boa explicação da crença de que há átomos o apelar para os estados psicológicos, pois supostamente há uma evidência (os rastros de vapor) que independe de quem está observando. Então o estado psicológico é irrelevante neste caso. Por outro lado, quando alguém vê um grupo de jovens colocando gasolina em um gato e incendiando-o e conclui que o que eles estão fazendo é errado, não parece ser uma condição necessária a pressuposição de que há um fato moral que pertence a esta ação. A melhor explicação para a minha crença parece ser uma suposição sobre a minha sensibilidade moral ou estado psicológico de avaliar o ato como sendo errado<sup>189</sup>. Como afirma Harman,

[...] uma suposição sobre fatos morais parece ser totalmente irrelevante para a explicação de você fazer o juízo moral que você faz. Parece que tudo o que precisamos assumir é que você tem certos princípios morais mais ou menos articulados que são reflexo dos julgamentos que você faz, baseado na sua *sensibilidade moral*. Parece ser completamente irrelevante para a nossa explicação se o seu julgamento intuitivo imediato é verdadeiro ou falso<sup>190</sup>.

Além disso, devemos notar que na observação dos prótons há um fato que “pode *afetar* o que você observa”<sup>191</sup>. Quer dizer, você *vê* um rastro de vapor e, mesmo que isso não seja condição suficiente para a conclusão de que há prótons, concluir que há prótons é a melhor explicação para a existência dos rastros. Isso significa que houve uma

---

<sup>189</sup> Como notou Sayre MacCord, Cf. SAYRE-MCCORD, G. *Moral Theory and Explanatory Impotence. Essays on Moral Realism*, Ithaca, New York: Cornell University Press, 1988, p. 6, há um sentido fraco em que a observação do cientista também depende do seu estado psicológico. Para verificar os rastros de vapor e poder supor que há prótons na câmara de gás o cientista não poderia estar dormindo ou alucinando, por exemplo. Isso também é pressuposto para a observação moral de que os jovens estavam fazendo algo errado ao atear fogo no gato. Mas o sentido de ‘estado psicológico’ empregado por Harman não é este. Ele parece estar se referindo aos sentimentos morais, à capacidade de ver um ato de crueldade e poder reconhecer que é errado devido a algum sentimento de simpatia ou preocupação pela vítima mais os princípios morais endossados.

<sup>190</sup> HARMAN, *Op. Cit.*, p. 7. Itálico meu.

<sup>191</sup> *Idem*. Itálico meu.

observação de um fato. No entanto, no caso dos garotos colocando fogo no gato não há um fato moral que te afete perceptualmente. Sua visão pode ser afetada por um grupo de garotos, um gato, fogo, gasolina, etc., mas não pela *incorreção* deste ato<sup>192</sup>.

Se poderia pensar que é óbvio que a observação desempenha um papel determinante no nosso juízo de que é errado atear fogo em gatos. Se nós não víssemos os jovens fazendo o que estavam fazendo certamente não formaríamos o juízo moral. Se não víssemos que estavam infringindo sofrimento não diríamos que é errado. Esse é o sentido de ‘observação’ que atentei acima quando disse que em certo sentido a ética também pressupõe a observação assim como a ciência. No entanto, é importante chamarmos atenção para uma distinção entre dois sentidos de ‘observação’. No primeiro, você observa os garotos atear fogo no gato e vê que isso é um ato de crueldade, que estão infringindo dor deliberadamente, que o animal está sofrendo desnecessariamente, etc., e forma a crença de que o ato é errado. Harman concordaria que a observação é importante para a ética neste sentido. Mas não é a este sentido de observação que Harman se refere quando fala que não há observação na ética. O outro sentido é a observação de propriedades morais, como *incorreção*, *bondade*, *justiça*, etc. Estes fatos não estão presentes nas ações que você observa, segundo Harman. Por isso, não são parte da melhor explicação do porquê você formou a crença de que é errado atear fogo em gatos naquela situação. Este é o ponto de Harman.

Evidência observacional desempenha um papel na ciência que não parece desempenhar na ética, porque os princípios científicos podem ser justificados em última instância pelo seu papel em explicar observações [...] pelo seu papel explicativo. Aparentemente, princípios morais não podem ser justificados do mesmo modo. Parece verdade que não pode haver cadeia explicativa entre princípios morais e observações particulares no sentido que pode haver entre princípios científicos e observações particulares.<sup>193</sup>

Observe que, se as propriedades morais não desempenham um papel explicativo em nossas crenças morais, então não podemos defender a

---

<sup>192</sup> Cf. *Ibidem*, p. 7-8.

<sup>193</sup> *Ibidem*, p. 9.

existência de fatos morais de maneira análoga aos fatos científicos inobserváveis. Se não há fatos morais, então não podemos testar os nossos princípios ou confirmar se os nossos juízos morais são verdadeiros. Em outras palavras, não podemos observar o mundo para confirmar as nossas teorias morais. A isso Harman denomina “O problema com a ética”.

Esse é um poderoso argumento contra o naturalismo moral. O naturalista tem de negar a segunda premissa do argumento de Harman, isto é, tem de mostrar que as propriedades morais desempenham um papel explicativo se quiser sustentar a plausibilidade da existência de fatos morais de forma similar a como o realista científico sustenta a existência das entidades físicas. Vários naturalistas tentaram fornecer uma resposta ao “problema com a ética” argumentando que as propriedades morais fazem parte da melhor explicação das nossas crenças morais. Vejamos agora como eles defendem a naturalização da ética frente a este problema levantado por Harman.

### 3.2. As propriedades morais são explicativas

Nicholas Sturgeon e David Brink negam a segunda premissa do argumento de Harman argumentando que as propriedades morais desempenham um papel explicativo tanto para fatos não morais, em geral, quando para nossas crenças morais, em particular. Eles acreditam que é possível aplicar o mesmo raciocínio do realista científico para estabelecer que há fatos morais e que estes são parte do mundo natural. Podemos resumir o argumento do seguinte modo:

- (1) P é uma propriedade real se, e somente se, P desempenha um papel na melhor explicação da experiência.
- (2) Propriedades morais desempenham um papel na melhor explicação da experiência.
- (3) Propriedades morais são propriedades reais.

A primeira, e mais influente, resposta à tese de Harman foi sugerida por Sturgeon em *Moral Explanations*<sup>194</sup>. Neste texto, Sturgeon apresenta um argumento baseado em seu teste contrafactual para determinar se uma dada propriedade é explicativamente relevante ou não para a

---

<sup>194</sup> STURGEON, N. *Moral Explanations*. In David Copp and David Zimmerman (eds.), *Morality, Reason and Truth*. Totowa, NJ: Rowman & Littlefield, 1985a.

compreensão de algum fenômeno. Brink endossa o critério de Sturgeon e adiciona alguns comentários a fim de evitar algumas objeções. No que segue, irei apresentar a ideia de que as propriedades morais desempenham um papel explicativo, ao contrário do que argumentou Harman. Começamos com alguns exemplos em que, supostamente, os fatos morais contribuem para a explicação de fatos não morais.

Nós normalmente explicamos os feitos de alguém com base em seu caráter moral. Considere o caso de Hitler. É sabido que Hitler foi o responsável por uma das maiores barbáries morais da história ao ordenar e supervisionar a morte de milhares de pessoas. Todos nós facilmente concordaríamos que Hitler tinha de ser moralmente *depravado* para fazer o que fez. Obviamente, seria muito difícil sustentar, ao mesmo tempo, que Hitler tinha um *bom* caráter e que matou milhares de pessoas. Normalmente, os livros de história citam o seu caráter moral como sendo parte da explicação dos seus feitos. Portanto, parece que parte da explicação dos feitos de Hitler é que ele era uma pessoa moralmente deprimada. Nós formamos a crença de que ele era moralmente deprimado ao analisar o que ele fez. E se o que explica, em partes, o que ele fez era a sua personalidade moral (o fato de ser moralmente deprimado), então o que explica, em partes, a nossa crença de que Hitler era moralmente deprimado é o fato de que ele *realmente* era moralmente deprimado. O ele realmente ser deprimado é um fato moral acerca de sua personalidade. É claro que se pode citar outras coisas como sendo explicativamente relevantes para as ações de Hitler, tais como interesses e influências políticas, desejos pessoais, ideais sobre uma sociedade plena (por mais estranhos que fossem), etc. Como Sturgeon nota, explicações históricas normalmente não reivindicam *completamente* a versão dos fatos<sup>195</sup>. Mas é razoável supor que *parte* da explicação dos feitos de Hitler se deve a um fato moral a respeito de sua personalidade. Esse é o raciocínio de Sturgeon.

O mesmo aplica-se à repressão da escravidão. Segundo Sturgeon, embora a escravidão fosse algo presente desde os povos mais antigos, a oposição moral à sua prática tornou-se mais evidente nos séculos XVIII e XIX. Além disso, a oposição teve início principalmente nos países britânicos, França e na América do Norte, mesmo sendo uma prática difundida em muitos países. Sturgeon sugere que normalmente se tem uma resposta padrão para isso. “A prática da escravidão na Grã-Bretanha e América Francesa, e depois nos Estados Unidos, era muito *pior* do que as formas anteriores de escravidão, e muito pior do que a

---

<sup>195</sup> STURGEON, 1985a, *Op. Cit.*, p. 64

escravidão na América Latina”<sup>196</sup>. Note que aqui, como no exemplo de Hitler, a propriedade moral *incorreção* está desempenhando um papel explicativo sobre a oposição à escravidão.

Se quisermos, podemos imaginar inúmeros exemplos que pressupõem fatos morais como parte da explicação para sua ocorrência<sup>197</sup>. Como afirma Brink,

[...] explicamos boa parte do comportamento de Calígula, tal como a execução de pessoas inocentes, através da sua *crudade*, e o fato de ele ter este vício de caráter explica, em grande parte, nossa opinião sobre ele e suas ações. De acordo com relatos populares, é o fato de que Lincoln era um homem *equitativo* e *justo* que explica a sua oposição à escravidão [...]. Além disso, frequentemente explicamos fatos e eventos sociais, assim como ações individuais, apelando aos fatos morais sobre instituições legais e sociais. Acreditamos que vícios políticos (e. g., *injustiça social*) às vezes causam, e assim ajudam a explicar, instabilidade, movimentos de protesto e revoluções; e achamos que virtudes políticas das instituições e leis de uma sociedade (e. g., *justiça social*) podem ajudar a explicar a sua estabilidade. Neste sentido, podemos citar a *injustiça* da escravidão como parte da explicação do seu desaparecimento ou a *injustiça* do sistema político e legal da África do Sul como causa da sua instabilidade e do protesto contra o apartheid<sup>198</sup>.

Nesses casos, parece que a propriedade moral é parte da explicação para os feitos de Hitler, de Calígula ou da oposição à escravidão. Se formos a crença de que Calígula era cruel ao analisar os seus feitos e o que explica, em partes, os seus feitos é a *crudade*, então parece que

---

<sup>196</sup> *Idem.*

<sup>197</sup> É importante notar que Sturgeon não estabeleceu previamente que há fatos morais para poder argumentar que fazem parte da explicação de fenômenos não morais. Ele apenas pressupõe, para o bem do argumento, que há fatos morais. O que irá determinar se realmente há ou não, em sua visão, é a propriedade de o fato moral desempenhar um papel explicativo, que será estabelecido a seguir.

<sup>198</sup> BRINK, 1989, *Op. Cit.*, p. 187. Itálicos meus.

esta propriedade moral desempenha um papel explicativo, mesmo que implícito, sobre nossas crenças. Parece que não formamos nossas crenças morais somente a partir do exercício da nossa sensibilidade moral frente aos feitos, como argumentou Harman. Há algo a mais que é parte da explicação dos fatos não morais, quer dizer, a propriedade moral.

Alguém inclinado a defender a posição de Harman, poderia levantar a seguinte objeção, como faz Quinn.

Mesmo que às vezes possamos usar hipóteses morais para explicar a ocorrência de uma observação, parece plausível pensar que nunca *precisamos* usá-las. Um explicador inteligente e reflexivo pode sempre substituí-las com hipóteses não morais sobre o mundo juntamente com hipóteses psicológicas sobre as crenças e atitudes morais das pessoas<sup>199</sup>.

Nesse sentido, se poderia explicar os feitos de Hitler sem a necessidade de se apelar para o seu caráter moral, mas apenas para os seus desejos, ambições políticas, ideais de sociedade etc. Se poderia explicar a oposição e fim da escravidão através das circunstâncias sociais e das crenças das pessoas sobre aquelas circunstâncias sociais, sobre os efeitos psicológicos que a prática da escravidão causava tanto nos escravos quanto nos opositores em geral. Poderíamos explicar os feitos de Calígula reivindicando os seus interesses políticos e assim por diante. Em resumo, nós poderíamos explicar a crença de que algo é correto ou incorreto, bom ou mal, sem apelar para nenhum fato moral a respeito da incorreção ou bondade morais.

Brink tenta responder a esse tipo de objeção do seguinte modo. Em primeiro lugar, a objeção não mostra que o apelo a fatos morais para a explicação de um evento é “completamente irrelevante”, como argumentou Harman. Quinn sugere que se podem utilizar outros meios explicativos, mas a possibilidade de a propriedade moral desempenhar um papel explicativo ainda continua sendo uma alternativa plausível, mesmo que não precise ser necessariamente reivindicada. Além disso, em segundo lugar, se poderia responder a Quinn dizendo que há casos possíveis em que certos fatos explicam a ocorrência dos eventos mesmo

---

<sup>199</sup> QUINN, W. Truth and Explanations in Ethics. *Ethics*, Vol. 96, No. 3 (Apr., 1986), pp. 524-544, p. 529.

que a crença ou o apelo das pessoas a tal fato seja irrelevante para a explicação dos eventos porque as pessoas não reconhecem tal fato. Brink sugere o seguinte exemplo. Imaginemos que  $x$  é casado com  $y$  e  $y$  está sendo infiel com  $x$  ao ter um caso com  $z$ .  $x$  não sabe da infidelidade de  $y$ . O ter um caso com  $z$  afeta levemente o comportamento de  $y$  para com  $x$ .  $y$  fica mais distante de  $x$ , menos atencioso, se irrita com mais frequência, etc. Isso faz com que  $x$  tenha um desempenho não tão bom no trabalho, que não durma adequadamente, que fique triste com mais frequência, etc. Se poderia explicar o comportamento de  $x$ , ou até mesmo  $x$  poderia explicar seu próprio comportamento, com base no comportamento não atencioso de  $y$ . E isso seria uma boa explicação. No entanto, a propriedade relevante que realmente explica o comportamento de  $x$  é o fato de  $y$  estar tendo um caso com  $z$ , mesmo que isso não seja reivindicado na explicação. Agora considere os casos morais. Mesmo que as pessoas não precisem recorrer a fatos morais para explicar certos eventos, porque não os reconhecem, não se segue que os fatos morais não possam explicar os eventos ou, mais enfaticamente, que os fatos morais não estejam desempenhando um papel determinante na *melhor explicação* dos eventos.

Note que até agora estivemos apenas considerando alguns exemplos em que as propriedades morais *parecem* desempenhar um papel explicativo para a compreensão de certos fenômenos. Não vimos ainda nenhum argumento definitivo a favor da tese de que propriedades morais são explicativas. Isso pode ser feito se considerarmos o teste contrafactual de Sturgeon para determinar a relevância explicativa de um dado fato. Sturgeon propõe o seguinte teste:

*Teste Contrafactual:* se uma suposição particular é completamente irrelevante para a explicação de certo fato, então o fato seria obtido, e o poderíamos ter explicado do mesmo modo, mesmo se a suposição tivesse sido falsa<sup>200</sup>.

Consideremos o exemplo de Hitler, novamente. O fato moral a seu respeito, quer dizer, o fato de ele ser moralmente depravado, seria “completamente irrelevante” para a explicação de seus feitos – e, por conseguinte, de nossa crença moral sobre sua personalidade – se ele tivesse ordenado e supervisionado a morte de milhões de pessoas mesmo se o ser moralmente depravado fosse falso, ou seja, mesmo se ele não fosse moralmente depravado. A questão é: poderia você ter tido

---

<sup>200</sup> STURGEON, 1985a, *Op. Cit.*, p. 65.

a crença de que Hitler era moralmente depravado mesmo se ele não fosse moralmente depravado? Se Hitler não fosse moralmente depravado ele teria ordenado e supervisionado a morte de milhões de pessoas? Parece insensato supor que sim. Não parece razoável supor que alguém pode ser, ao mesmo tempo, moralmente bom e ordenar a morte de milhões de pessoas, pois uma pessoa moralmente boa reconhece regras morais básicas como a de que o assassinato e a tortura são práticas incorretas. Então, parece que ser moralmente bom implica não matar milhões de pessoas. Se Hitler matou milhões de pessoas, então não poderia ter sido bom. Como já notamos acima, há outros fatores que também podem contribuir para a explicação dos feitos de Hitler, mas se a suposição moral a seu respeito *tivesse sido falsa*, então parece que ele não teria feito o que fez. Portanto, o fato moral não é “completamente irrelevante” para a explicação dos feitos de Hitler e, conseqüentemente, para a nossa crença moral de que ele era moralmente depravado. Mas isso é apenas uma visão mais intuitiva do teste contrafactual de Sturgeon. O seu argumento é mais sofisticado.

Retornemos ao exemplo proposto por Harman. Você vê um grupo de jovens colocando gasolina em um gato e incendiando-o. Harman sustenta que a incorreção deste ato é “completamente irrelevante” para a explicação de sua crença de que o que eles estão fazendo é errado. Mas se esta ação não fosse errada, você ainda assim teria a crença de que é errada? O teste contrafactual nos pede para imaginar um mundo possível em que os jovens estão colocando gasolina em um gato e incendiando-o, mas que essa ação não tem a propriedade da *incorreção*. Se isso é verdadeiro, se não é um fato moral que atear fogo em gatos é errado, então a *incorreção* é realmente irrelevante para a explicação da sua crença de que eles estão fazendo algo errado. Mas, o problema consiste em imaginar um mundo possível próximo em que os jovens estão fazendo exatamente o que estão fazendo no mundo atual e que sua ação não seja errada. Vejamos o porquê.

Como vimos na primeira parte deste trabalho, as propriedades morais são supervenientes às propriedades naturais. Aparentemente, tanto um naturalista quanto um não naturalista concordam com isso. Neste sentido, dado que a incorreção é superveniente à propriedades como infringir dor deliberadamente, ser cruel com a vítima, fazer a vítima sofrer sem causa relevante, etc., então imaginar um mundo possível que os jovens não estão fazendo algo errado é imaginar um mundo possível em que eles não estão causando dor deliberadamente ao animal, não estão fazendo-o sofrer e assim por diante. Mas imaginar um mundo possível em que os jovens não estão causando dor

deliberadamente ao animal não é imaginar um mundo possível em que estão colocando gasolina e incendiando-o. Portanto, num mundo possível em que os jovens estão colocando gasolina num gato e incendiando-o, eles estão fazendo algo errado, pois a *incorreção* é superveniente às propriedades naturais que constituem o que eles estão fazendo.

Considere agora a sua crença de que eles estão fazendo algo errado. O que explica a sua crença? Sturgeon acredita que a *incorreção* desempenha um papel explicativo na sua crença de que a ação dos jovens é errada. Se a *incorreção* fosse completamente irrelevante para a explicação de sua crença, então eles teriam que poder estar fazendo exatamente o que estavam fazendo e a ação não ser errada. Se este fosse o caso então parece que a *incorreção* não contribuiu para a formação de sua crença. Mas como vimos, dado que a propriedade da *incorreção* é superveniente àquilo que eles estão fazendo, então não temos como supor uma situação possível em que eles estão ateando fogo no gato e não estão fazendo algo errado. Se sua crença é formada a partir do que eles estão fazendo, e o que eles estão fazendo é errado, então parece que a *incorreção* explica, ao menos em parte, a sua crença de que o que eles estão fazendo é errado. Portanto, se eles não estivessem fazendo algo errado você não teria a crença de que eles estão fazendo algo errado<sup>201</sup>.

No nosso exemplo científico, o físico vê os rastros de vapor na câmara de nuvem e conclui que há prótons, pois o que melhor explica os rastros é a existência dessas propriedades. No caso dos jovens ateando fogo ao gato a *incorreção* explica em partes a sua crença de que eles estão fazendo algo errado. Isso porque, se a *incorreção* é explicativamente irrelevante, então você teria que poder ter a crença de que é errado mesmo que não fosse. Mas para imaginar uma situação em que não há *incorreção* no ato de atear fogo em gatos, há que se imaginar uma situação em que eles não estão causando dor deliberada e cruelmente ao animal, pois estas são as propriedades a partir das quase a *incorreção* sobrevêm. Mas fazer isso é alterar completamente o caso, dado que o que estamos considerando é uma situação em que você vê jovens colocando fogo no gato. Então, imaginar uma situação em que um grupo de jovens coloca fogo em um gato deliberada e cruelmente é imaginar uma situação em que eles estão fazendo algo errado. E se você forma a crença de que é errado a partir daquilo que eles estão fazendo, então o que explica em partes a sua crença de que é errado é o fato de que é realmente errado. Portanto, tem de haver propriedades morais

---

<sup>201</sup> Cf. *Ibidem*, p. 65ss.

como *incorreção*, pois se não houvesse você não formaria a crença de que atear fogo em gatos deliberadamente é errado.

Se quisermos colocar o argumento de Sturgeon na forma canônica, podemos organizá-lo do seguinte modo:

- (1) Se a *incorreção* é explicativamente irrelevante para a sua crença de que  $x$  é moralmente errado, então você deve poder formar a crença de que  $x$  é moralmente errado mesmo que a *incorreção* não seja uma propriedade de  $x$ .
- (2) Você não pode formar a crença de que  $x$  é moralmente errado se a *incorreção* não for uma propriedade de  $x$ .
- (3) Portanto, a *incorreção* não é explicativamente irrelevante para a sua crença de que  $x$  é moralmente errado.

Este é um *mudus tollens* válido. A segunda premissa é justificada com base na ideia da superveniência das propriedades morais sobre as naturais. Se aceitamos a superveniência, então temos que aceitar que um mundo possível em que a ação de atear fogo em gatos não tem a propriedade da *incorreção* é um mundo possível em que o atear fogo em gatos não possui a propriedade de ser uma ação de causar dor desnecessariamente, cruelmente e deliberadamente. Mas uma ação que não tenha estas propriedades não é uma ação similar a que estamos considerando, quer dizer, a ação tem de ter estas propriedades. Mas o ter essas propriedades implica que a ação tem, também, a propriedade da *incorreção* e, portanto, é errada. Conclui Sturgeon, você não poderia formar a crença de que a ação é errada se ela não fosse realmente errada.

Podemos aplicar o teste contrafactual aos inúmeros exemplos em que a propriedade moral desempenha um papel explicativo. Se Calígula não fosse uma pessoa *cruel* não teria matado pessoas inocentes. Se Lincoln não fosse *correto* e *justo* não teria sido um opositor da escravidão. As pessoas não lutariam contra o *apartheid* se este movimento não fosse *injusto*. E assim por diante. Quinn, por exemplo, poderia dizer que se pode explicar os protestos dos negros contra o *apartheid* apelando não para a injustiça, mas para o próprio interesse dos negros em impedir que tal política continuasse suprimindo seus direitos. Brink sugere uma resposta interessante. Em primeiro lugar, devemos insistir no papel explicativo da injustiça. É certo que os negros tinham o interesse em banir o *apartheid*. Mas o que motivou o seu

interesse foi o fato de que o *apartheid* era injusto, incorreto e ilegítimo, quer dizer, o que motivou a oposição fora a incorreção moral. Em segundo lugar, se o que explica a oposição ao *apartheid* for apenas o interesse dos negros que estavam em uma situação desfavorável, então parece que brancos não teriam nenhuma razão para apoiar os protestos contra o *apartheid*. No entanto, é sabido que muitos brancos também foram grandes opositores desta política. E o que parece explicar isso é o fato de que era algo moralmente *errado*<sup>202</sup>.

Além disso, devemos lembrar que Brink endossa a posição de Boyd sobre a constituição múltipla das propriedades morais. Assim, as propriedades morais não são *idênticas* às propriedades naturais, mas são *realizadas* ou *constituídas* pelas propriedades naturais. Isso significa que, a primeira vista, pode parecer que a propriedade moral não desempenha um papel explicativo, pois se poderia citar vários fatos não morais relevantes como melhores candidatos à explicação. Por exemplo, se poderia dizer que são os interesses políticos, desejos pessoais, ideais sociais, etc., ao invés de sua personalidade depravada, que melhor explicam o assassinato de milhões de pessoas por Hitler. Esta é uma resposta bastante razoável. Como temos visto, o fato moral a respeito de Hitler não reivindica a explicação completa de seus atos, mas apenas parte. Mas dado que as propriedades morais são constituídas pelas propriedades naturais, os interesses políticos, ideais sociais e desejos pessoais de Hitler *constituem* a propriedade moral, seja ela a incorreção ou a depravação. Portanto, mesmo que se reivindique propriedades não morais para a explicação de determinado evento, se essas propriedades forem constitutivas do fato moral então o fato moral desempenha um papel explicativo.

Essa é outra estratégia utilizada pelos naturalistas não reducionistas para defender que os fatos e propriedades morais são naturais. Vimos que o raciocínio é similar ao realismo científico que se baseia na inferência a favor da melhor explicação. Harman defende que as propriedades morais não desempenham um papel explicativo a respeito de nossas crenças morais. Na visão de Harman, formamos juízos morais não porque observamos que certos atos têm certas propriedades morais como *incorreção* ou *bondade*, mas porque a nossa sensibilidade moral nos compele a fazermos tais e tais juízos morais. A ética, a seu ver, não é passível de observação e este é “o problema com a ética”. Como vimos, os naturalistas tentam responder a este problema argumentado que se as propriedades morais desempenham um papel

---

<sup>202</sup> Cf. BRINK, 1989, *Op. Cit.*, p. 190-191.

explicativo, então deve haver propriedades morais e que as propriedades morais desempenham um papel explicativo. A conclusão, obviamente, é de que há fatos morais. O argumento principal se baseia no teste contrafactual de Sturgeon.

## **OBSERVAÇÕES FINAIS**

Neste capítulo, tentei apresentar as principais ideias do naturalismo não reducionista mostrando em que sentido se diferencia do naturalismo analítico e como evita algumas objeções tradicionais. Começamos por uma série de avanços na filosofia da linguagem, metafísica, epistemologia e filosofia da mente que propiciaram o desenvolvimento de uma forma renovada de naturalismo moral. Vimos que este naturalismo não é vulnerável a objeções como o argumento da questão aberta e a versão hareana deste argumento, dada a recusa da tese de que se dois termos são sinônimos, então devem referir o mesmo objeto. Com a ideia de que alguns termos singulares e tipos naturais são designadores rígidos o naturalista pode agora defender que sentenças que expressam identidade entre duas propriedades e que são compostas por designadores rígidos são verdades necessárias sem que sejam analíticas. Isso significa que ele pode reivindicar que propriedades morais são naturais sem a necessidade de apresentar um método reducionista. Além disso, ao se apropriar da ideia dos filósofos da mente de que existem certas entidades que são multiplamente realizadas ou constituídas por um conjunto de outras propriedades, o naturalista pode dizer sustentar que a definição das propriedades morais não é analítica, mas constituída por um número indefinido de propriedades naturais. Com a adoção de uma teoria semântica externalista, o naturalismo não precisa se comprometer com a ideia de que o significado dos termos é conferido por um conjunto de descrições que são atribuídas pelos falantes. Agora, ele pode sustentar que o significado de termos como ‘bom’, ‘justo’ e ‘incorreto’ é fixado por estados de coisas que são externos às opiniões e crenças dos indivíduos. Tendo essas ideias em mãos, pudemos perceber em que sentido o novo naturalismo se distingue do naturalismo reducionista.

Na segunda parte, nos ocupamos da teoria de Boyd. Vimos que há certos termos cujas definições dependem de um conjunto de propriedades, que são a sua extensão. No entanto, essas definições não são exaustivas, quer dizer, é sempre possível encontrarmos outros indivíduos que, embora sejam em alguns aspectos distintos dos

indivíduos que compõem a extensão em questão, compartilham propriedades relevantes e, portanto, também fazem parte da definição. Nesse sentido, essas definições vão sendo constituídas à medida que o nosso conhecimento progride. Boyd denomina essas definições de agregados de propriedades homeostáticas. Além disso, vimos que ele defende que as propriedades morais são agregados de propriedades homeostáticas. Se isso é verdadeiro, então o naturalista pode dizer que as propriedades morais são multiplamente realizadas por um conjunto indefinido de propriedades naturais; que não é necessário termos uma definição precisa sobre a extensão das propriedades morais; que as teorias normativas nos fornecem uma explicação aproximada sobre quais propriedades naturais fazem parte do agregado das propriedades morais; e que a definição da propriedade moral vai sendo revelada à medida que o nosso conhecimento moral progride. Novamente, o naturalista não precisa se comprometer com a ideia de que aquilo que é bom depende, em algum sentido, das crenças e opiniões dos agentes.

E, na última parte, consideramos a defesa de que as propriedades morais são reais e, por conseguinte, naturais, com base na inferência a favor da melhor explicação. Harmam defendeu que as propriedades morais não desempenham nenhum papel explicativo no que diz respeito às nossas crenças morais. Por isso, segundo ele, o naturalista não pode adotar o raciocínio do realista científico que sustenta que propriedades como prótons e elétrons existem tendo como base o papel explicativo desempenhado por estas entidades. No entanto, Brink e Sturgeon tentam fornecer uma defesa de que as propriedades morais desempenham um papel explicativo. O ponto forte do argumento destes autores é o teste contrafactual de Sturgeon que diz que se uma suposição particular é completamente irrelevante para a explicação de certo fato, então o fato seria obtido, e o poderíamos ter explicado do mesmo modo, mesmo se a suposição tivesse sido falsa. Brink e Sturgeon sustentam que as propriedades morais não são completamente irrelevantes para a explicação de nossas crenças morais. Nesse sentido, as propriedades morais desempenham um papel explicativo. Portanto, são reais.



### III – NATURALISMO E SEMÂNTICA MORAL

“[...] the reliability of our intuitions diminishes in proportion to the complexity of the case at hand”.  
(Geirsson, 2003).

## INTRODUÇÃO

Há pouco mais de meio século, R. M. Hare apresentou um argumento contra o naturalismo moral sustentando que as propriedades morais são *algo a mais* do que um conjunto de propriedades descritivas. A este ‘algo a mais’ ele chamou de significado prescritivo e, indo além, argumentou que o significado prescritivo é primário e, portanto, característica mais importante das propriedades morais. Mas o argumento de Hare dependia do pressuposto de que o significado dos juízos morais é determinado pelos falantes de uma comunidade linguística. Assim, com o desenvolvimento das teorias externalistas da referência, o naturalista pôde finalmente ver uma luz no fim do túnel e sustentar que as propriedades morais são independentes dos sujeitos.

No entanto, graças a dois filósofos de Tennessee, o naturalista têm tido que enfrentar um argumento direcionado a ofuscar sua visão novamente. Em uma série de artigos<sup>203</sup>, Terence Horgan e Mark Timmons (H&T) têm argumentado que o naturalismo é vulnerável a uma objeção muito similar ao argumento de Hare. Mas, diferentemente de Hare, o argumento de H&T se aplica ao naturalismo não reducionista também. Eles pensam que a objeção é fatal e que, por conseguinte, o naturalista está destinado à escuridão de uma vez por todas.

Neste breve capítulo final, gostaria de apresentar o argumento proposto por H&T no intuito de mostrar que mesmo o naturalismo desenvolvido por Brink, Boyd e Sturgeon como uma tentativa de reabilitar o realismo moral tem de enfrentar um problema difícil. A

---

<sup>203</sup> Ver HORGAN, T. & TIMMONS, M. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, Volume 16, 1991, p. 447-465, HORGAN, T. & TIMMONS, M. Troubles on Moral Twin Earth: Moral Queerness Revived. *Synthese*, 92, 1992a, p. 221-260 e HORGAN, T. & TIMMONS, M. Troubles for New Wave Moral Semantics: The Open Question Argument Revived. *Philosophical Papers*, Volume 21, No. 3, 1992b, p. 153-175.

posição desenvolvida por esses três autores pareceu ser a alternativa correta para dar plausibilidade ao realismo moral naturalista, como vimos no capítulo anterior. No entanto, o *Argumento da Terra Gêmea Moral*, como ficara conhecido o problema proposto por H&T, parece ser o entrave principal que o naturalista tem de enfrentar. Além de apresentar esse argumento, irei reconstruir algumas estratégias de defesa do naturalismo. Essas estratégias são negativas, no sentido de que tentam formular razões para recusarmos o argumento de H&T ao invés de modificarmos positivamente a própria posição naturalista. Irei me limitar a apresentar a discussão com o intuito de mostrar que há um problema que o naturalista tem de enfrentar se quiser reivindicar plausibilidade à sua posição e, além disso, irei apresentar alguns caminhos iniciais possíveis.

Vou dividir essa tarefa em duas partes. Na primeira, apresentarei o argumento de H&T mostrando em que medida é supostamente uma refutação do naturalismo não reducionista. Este argumento é construído sobre os pilares de um experimento mental para os termos morais baseado no cenário de Putnam. Se o experimento é legítimo, H&T sustentam, então se pode construir uma nova versão do argumento da questão aberta; dessa vez para o naturalismo não reducionista. Buscarei apresentar esta reformulação do argumento de Moore também. Na segunda parte, quero mostrar algumas objeções a que o argumento de H&T é vulnerável.

## **1. O argumento da Terra Gêmea Moral**

Começemos relembando algumas coisas. O naturalismo moral pode ser dividido em duas subáreas. À primeira, denominamos naturalismo analítico, ou naturalismo reducionista, que sustenta que os termos morais têm definições naturalistas analiticamente verdadeiras. À segunda, naturalismo sintético, ou naturalismo não reducionista, que sustenta que os termos morais têm definições naturalistas sinteticamente verdadeiras. Antes do desenvolvimento do naturalismo sintético acreditava-se o naturalismo analítico era a única via para se defender o naturalismo moral. O argumento da questão aberta objetivou refutar o naturalismo analítico. Com isso, Moore acreditou estar refutando o naturalismo moral em geral; algo que, como temos visto, é falso. O naturalismo sintético não se compromete com a analiticidade entre termos morais e não morais. A teoria semântica endossada pelos naturalistas sintéticos é externalista, o que significa que a referência dos termos morais é determinada pelo mundo externo e não pelo sujeito

unicamente. Seguindo Boyd, podemos dizer que  $k$  regula causalmente o uso de  $t$ , em que  $k$  é a propriedade natural e  $t$  é o predicado usado para se referir a  $k$ . H&T denominaram isso de *tese da regulação causal*.

Dado que aqui o assunto é os termos morais, então podemos reformular a tese da regulação causal para os termos morais dizendo o seguinte.

*Tese da Regulação Causal:* para cada termo moral  $t$  (e.g., ‘bom’), há uma propriedade natural  $N$  tal que  $N$  somente, e nenhuma outra propriedade, regula causalmente o uso de  $t$  pelos humanos<sup>204</sup>.

Esta é a tese semântica do naturalismo não reducionista. Os naturalistas acreditam que, dado que o naturalismo semântico sintético é verdadeiro, então o naturalismo moral é verdadeiro.

O experimento mental de Putnam sobre a terra gêmea visou mostrar que a extensão dos termos não é determinada somente pelas intensões dos falantes. Como vimos, temos a intuição de que ‘água’ tal como usada pelos habitantes da Terra (‘T-água’) e ‘água’ tal como usada pelos habitantes da Terra Gêmea (‘TG-água’) tem significados diferentes, pois o que *regula causalmente* um é  $H_2O$  e outro é  $XYZ$ . E dado que falantes competentes têm intuições fortes sobre as regras que regem o comportamento semântico e sintático da sua linguagem, temos uma boa evidência de que os termos realmente têm significados diferentes. O experimento de Putnam mostrou, portanto, que ‘T-água’ e ‘TG-água’ são designadores rígidos.

Agora, consideremos o seguinte. O naturalista está sustentando que as propriedades morais são naturais e que, então, os termos morais se comportam exatamente como os termos naturais. Quer dizer, assim como para termos como ‘água’ há um estado de coisas que regula causalmente o seu uso, para termos morais também há. Se assim é, então se substituirmos o termo ‘água’ do experimento mental de Putnam por termos morais, deveremos ter as mesmas intuições que temos para com o termo ‘água’. Isso significa que os termos morais também são designadores rígidos. Em outras palavras, dizer que o naturalismo sintético é verdadeiro implica em dizer que uma reflexão sobre os termos morais no cenário de Putnam deve gerar exatamente as mesmas intuições que as intuições geradas para com termos não morais e que isso deve fornecer um juízo intuitivo de que a tese sobre a regulação

<sup>204</sup> Cf. H&T, 1991, *Op. Cit.*, p. 455 e 1992b, *Op. Cit.*, p. 159.

causal para termos morais é verdadeira, e que, portanto, os termos morais se comportam como o naturalismo nos diz que se comportam<sup>205</sup>. Todavia, se a reflexão sobre os termos morais no cenário de Putnam não gerar as mesmas intuições que os termos não morais, então deve haver algo de errado com o naturalismo sintético.

H&T sustentam que se colocarmos termos morais no cenário de Putnam não teremos as mesmas intuições que temos com os termos não morais. Portanto, deve haver alguma diferença semântica entre os termos morais e os termos não morais. Isso significa, em última instância, que o naturalismo sintético é falso. E, da mesma forma que Moore argumentou que o naturalismo moral é falso porque o naturalismo analítico é falso, H&T argumentam que o naturalismo moral é falso porque o naturalismo sintético é falso. Vejamos o seu argumento.

Como sugeriram Brink e Boyd, as propriedades morais são propriedades funcionais. Ao invés de serem constituídas por uma única propriedade natural (como *prazer*, por exemplo), são multiplamente realizadas por um conjunto indefinido e propriedades naturais que constituem a sua definição. Essas propriedades naturais regulam o nosso uso dos termos morais, pois só podemos dizer que algo é *bom*, por exemplo, se este algo for uma instância que realiza *bom*. Quais são as propriedades naturais que constituem as propriedades morais é determinado pela teoria normativa que se adota. Neste sentido, H&T nos pedem para supormos que a teoria normativa que regula o uso das nossas propriedades morais é uma teoria consequencialista, de modo que as propriedades funcionais que regulam o nosso uso das propriedades morais são propriedades consequencialistas. Denominemos esta teoria de T<sup>c</sup>.

Agora entra em jogo o velho experimento mental de Putnam, com algumas modificações, obviamente. Imagine um planeta chamado Terra Gêmea Moral (TGM) que é similar a Terra (T).

[...] Mesma geografia e ambiente natural, as pessoas que vivem nos Estados Unidos falam Inglês-gêmeo; há um estado que eles chamam “Tennessee” que é situado diretamente ao sul de um estado que eles chamam “Kentucky”; e todo ano um grande número de Terráqueos Gêmeos faz

---

<sup>205</sup> Cf. H&T, 1991, *Op. Cit.*, p. 458 e 1992b, *Op. Cit.*, p. 162-163.

uma peregrinação para o Memphis Gêmeo para visitar a sepultura do Elvis Gêmeo.

Já conhecemos a ideia. *Quase* tudo o que há em TGM é exatamente similar ao que há em T. É importante chamar atenção para o fato de que os habitantes de TGM tem um vocabulário *moral* muito similar aos de T.

[...] eles usam os termos “bom” e “mau”, “certo” e “errado” para avaliar ações, pessoas, instituições e assim por diante [...]. De fato, se um grupo de exploradores da Terra visitassem a Terra Gêmea Moral eles seriam fortemente inclinados a traduzir os termos da Terra Gêmea Moral “bom”, “correto” e o restante como idênticos em significado para os nossos termos do Inglês ortograficamente idênticos. Não obstante, os usos desses termos na Terra Gêmea Moral suporta todas as marcas “formais” que tomamos como características do vocabulário e da prática moral. Em particular, os termos são usados para raciocinar sobre considerações relacionadas sobre o bem estar da Terra Gêmea Moral; os habitantes da Terra Gêmea Moral estão normalmente dispostos a agir de certos modos em correspondência com os julgamentos sobre o que é “bom” e “correto”; eles normalmente tomam considerações sobre o que é “bom” e “correto” como sendo especialmente importante [...] para decidir o que fazer, e assim por diante<sup>206</sup>.

Assim como em T, em TGM as propriedades morais são funcionais e, portanto, multiplamente realizadas por um conjunto indefinido de propriedades naturais. De maneira similar, o que determina quais propriedades naturais constituem as propriedades morais é a teoria normativa que se assume. No entanto, em TGM as propriedades morais não são reguladas por um conjunto de propriedades funcionais consequencialistas. Em TGM o uso dos termos morais é regulado por propriedades não consequencialistas que são captadas por uma teoria deontológica. Denominemos tal teoria de T<sup>d</sup>.

---

<sup>206</sup> H&T, 1992, *Op. Cit.*, p. 164.

Notemos que, embora os habitantes de T e TGM façam uso do discurso moral de maneira muito similar, as propriedades naturais que regulam o uso de ‘bom’ em T (‘T-bom’) e o uso de ‘bom’ em TGM (‘TGM-bom’) são distintas. O que regula o uso de ‘T-bom’ é o conjunto de propriedades consequencialistas. E o que regula o uso de ‘TGM-bom’ é o conjunto de propriedades deontológicas. Portanto, a referência de ‘T-bom’ é diferente da referência de ‘TGM-bom’. Ou, ao menos assim argumentam H&T.

Agora, suponhamos que alguns habitantes de T viagem para TGM ou que alguns habitantes de TGM viagem para T. Assumamos que ambos estão cientes a respeito da teoria normativa que capta o conjunto de propriedades naturais. Os habitantes de T sabem que o que regula o seu uso de ‘bom’ é T<sup>c</sup> e os habitantes de TGM sabem que o que regula o seu uso de ‘bom’ é T<sup>d</sup>. Suponhamos que eles venham a discordar sobre algum assunto moral. Digamos que os habitantes de T estão sustentando que mentir não é sempre errado, pois há circunstâncias em que seria muito melhor dizer uma falsidade para o bem geral. Já os habitantes de TGM argumentam que mentir é errado em qualquer circunstância, pois dizer falsidades viola uma regra fundamental do seu sistema moral. O ponto é: eles estão tendo um desacordo real ou estão apenas a “jogar palavras ao vento”? Quer dizer, quando os habitantes de TGM dizem que ‘Mentir é errado’ e os habitantes de T dizem que ‘Mentir não é errado’ eles estão negando a afirmação um do outro ou estão apenas, cada um à sua maneira, reivindicando um conjunto de propriedades naturais distintas (os habitantes de TGM um conjunto de propriedades funcionais deontológicas e os habitantes de T um conjunto de propriedades funcionais consequencialistas)?

Temos duas opções agora. A primeira é dizer que o significado de ‘bom’ tal como usado pelos habitantes de T designa rigidamente o conjunto de propriedades naturais que regula causalmente o *seu* uso de ‘bom’. E o significado de ‘bom’ tal como usado pelos habitantes de TGM designa rigidamente o conjunto de propriedades naturais que regula causalmente o *seu* uso de ‘bom’. Isso significa que ‘T-bom’ designa o conjunto de propriedades consequencialistas. E ‘TGM-bom’ designa o conjunto de propriedades deontológicas. Portanto, os habitantes de T e os de TGM não estão tendo um desacordo moral. Como dizem H&T, “o termo moral e o termo gêmeo moral *diferem em significado*, e não são intertraduzíveis”<sup>207</sup>. A outra opção é dizer que eles estão realmente tendo um desacordo moral. ‘T-bom’ e ‘TGM-bom’

---

<sup>207</sup> *Ibidem*, p. 165.

“não diferem em significado ou referência, e então qualquer desacordo moral aparente que pode ser levantado entre os Terráqueos e os Terráqueos Gêmeos seria um desacordo *genuíno*”<sup>208</sup>. Ao invés de serem desacordos meramente semânticos, seriam “desacordos em crenças morais e em teoria moral normativa”<sup>209</sup>.

Se o naturalismo não reducionista é verdadeiro e os fatos e propriedades morais são naturais, então os termos morais devem se comportar exatamente como os termos não morais se comportam. Isso quer dizer que, neste experimento mental moral devemos ter o mesmo tipo de intuição que tivemos quando consideramos o experimento de Putnam sobre o termo ‘água’. Se Brink e Boyd estão certos, então não deve haver dúvida de que é a primeira opção que está em questão. Quando os habitantes de T e os habitantes de TGM discordam sobre algum assunto moral, na verdade não há desacordo genuíno algum, pois a referência dos termos morais por eles utilizados é diferente. Uns reivindicam propriedades funcionais consequencialistas e outros propriedades funcionais deontológicas, assim como os habitantes de T se referiam a H<sub>2</sub>O e os habitantes de TG a XYZ no caso do experimento sobre ‘água’.

No entanto, segundo H&T, não é este tipo de intuição que temos quando consideramos o cenário moral. Parece-nos muito mais que eles estão tendo um desacordo moral genuíno, que as suas reivindicações não são meramente verbais, mas sobre crença e teoria moral<sup>210</sup>. Quando pensamos seriamente sobre o cenário moral a nossa intuição é de que ambas as partes estão usando os termos morais com o mesmo significado, com a mesma referência, e não como o naturalista teria que sustentar.

[...] se o NSC (Naturalismo Semântico Causal) fosse verdadeiro, e os termos em questão designassem rigidamente aquelas propriedades naturais que regulam causalmente o seu uso, então a reflexão neste cenário deve gerar intuições análogas àquelas geradas no cenário original da Terra Gêmea de Putnam. I. e., deveria parecer intuitivamente natural dizer que aqui nós temos uma diferença de significado, e que termos

---

<sup>208</sup> *Idem.*

<sup>209</sup> *Idem.*

<sup>210</sup> H&T, 1991, *Op. Cit.* p. 460-461, 1992a, *Op. Cit.*, p. 245ss e 1992b, *Op. Cit.*, p. 165, 166.

“morais” do Inglês gêmeo não são traduzíveis para os termos morais do Inglês. No entanto, quando paramos para caracterizar as diferenças entre os Terráqueos e os Terráqueos gêmeos sobre essa questão, a coisa natural a dizer é que as diferenças envolvem crença e teoria, não significado. [...] se o NSC fosse verdadeiro, então o reconhecimento dessas diferenças deveria resultar numa aparente tolice, para os membros de cada grupo, se engajar num debate intergrupar sobre a bondade [...] Mas tal debate intergrupar certamente tomaria ambos os grupos não como uma tolice, mas como bastante apropriado, pois eles considerariam um ao outro como diferindo em crenças morais e teoria moral, não em significado<sup>211</sup>.

É certo que o ter a intuição de que os habitantes de T e de TGM estão realmente engajados em um desacordo moral genuíno é um pressuposto empírico somente. Mas, parece que é realmente esta intuição que temos ao considerar o experimento da TGM, afirmam H&T. Como disse Van Roojen,

[...] falantes competentes têm a intuição que H&T procuram extrair com seu exemplo. A maioria das pessoas que tem lido o seu artigo parece concordar que as duas populações estão engajadas num debate sobre a *bondade moral*<sup>212</sup>.

E, se a intuição dos falantes competentes conta como uma forte evidência para a tese de que o significado não é inteiramente determinado pelo conjunto de descrições que os falantes associam a um objeto, no caso do cenário não moral de Putnam, então neste cenário moral a nossa intuição também deve contar como uma forte evidência a favor de que os termos morais não se comportam da maneira como os naturalistas dizem que se comportam. Neste sentido, o naturalismo não

---

<sup>211</sup> H&T, 1991, *Op. Cit.*, p. 60-61. Cf. também H&T, 1992b, *Op. Cit.*, p. 165ss. Parêntesis adicionado.

<sup>212</sup> VAN ROOJEN, M. Knowing Enough to Disagree: A New Response to the Moral Twin Earth Argument. In: SHAFER-LANDAU, R. (Org.) *Oxford Studies in Metaethics*. Volume 1. Oxford: Clarendon Press, 2006, p. 162-193, p. 168. Itálico adicionado.

reducionista está *provavelmente* falso, sustentam H&T<sup>213</sup>. E, ao mesmo estilo de Moore que acreditou estar refutando o naturalismo moral ao refutar o naturalismo semântico analítico, H&T acreditam estar refutando o naturalismo moral ao refutar o naturalismo semântico sintético<sup>214</sup>.

Note que a pressuposição central de H&T para sustentar que os termos morais ‘T-bom’ e ‘TGM-bom’ tem a mesma referência (ou que, pelo menos, temos a intuição de que tem a mesma referência) é que *nos parece* que ambos não têm significados diferentes. Neste sentido, podemos estruturar o seu argumento da seguinte forma<sup>215</sup>. Assumamos primeiro o seguinte: que  $M^T$  significa ‘termo moral tal como é usado Terra’, que  $M^{TGM}$  significa ‘termo moral tal como é usado na Terra Gêmea Moral’, que  $N^T$  significa ‘o conjunto das propriedades naturais na Terra’ e que  $N^{TGM}$  significa ‘o conjunto das propriedades naturais na Terra Gêmea Moral’. Teríamos algo como

P1. Suponhamos que o naturalismo não reducionista é verdadeiro.

P2.  $M^T$  refere  $N^T$  e  $M^{TGM}$  refere  $N^{TGM}$ .

P3.  $N^T$  e  $N^{TGM}$  são diferentes.

C1. Então,  $M^T$  e  $M^{TGM}$  possuem significado diferente.

P4. Mas,  $M^T$  e  $M^{TGM}$  não possuem significado diferente.

C2. Portanto, o naturalismo não reducionista é falso.

---

<sup>213</sup> H&T tomam o cuidado de dizer ‘provavelmente’ porque o argumento não se funda em um pressuposto indubitável, quer dizer, funda-se na intuição. Mas, como argumentam, a intuição é uma ferramenta muito importante em reflexões que envolvem a linguagem, por exemplo. De modo que é muito mais preferível uma teoria que se adéqua às intuições que os falantes competentes têm sobre a linguagem do que uma teoria que contraria as intuições.

<sup>214</sup> Cf. H&T, 1992b, *Op. Cit.*, p. 159.

<sup>215</sup> A seguinte organização do argumento é inspirada em Geirsson. Algumas modificações, mínimas, foram feitas. Cf. GEIRSSON, H. Moral Twin-Earth and Semantic Realism. *Erkenntnis*, 62, pp. 353-378, 2005.

Obviamente, a premissa central é a 5. A evidência de H&T para a sua verdade é a intuição de que há *desacordo moral genuíno* entre habitantes da T e da TGM. Neste sentido, podemos organizar o seu argumento para justificar 5 da seguinte forma.

P1. Se  $M^T$  e  $M^{TGM}$  diferem em significado, então os habitantes de T e os habitantes de TGM não podem ter desacordos morais legítimos.

P2. Os habitantes de T e os habitantes de TGM podem ter desacordos morais legítimos.

C1. Portanto,  $M^T$  e  $M^{TGM}$  não diferem em significado.

A conclusão de que o naturalismo é falso se segue do fato de que ‘T-bom’ e ‘TGM-bom’ não parecem ter significados diferentes, pois a nossa intuição é de que tem o mesmo referente. E isso se segue do fato de que nos parece que há *desacordo moral genuíno* entre os habitantes de T e de TGM.

### 1.1. O “novo” argumento da questão aberta

Vimos que o argumento da questão aberta de Moore não refuta o naturalismo moral porque, se atinge, atinge no máximo o naturalismo analítico. O naturalista não reducionista não está comprometido com a analiticidade entre propriedades morais e não morais. H&T procuram mostrar que o naturalismo sintético é falso através do seu experimento mental da Terra Gêmea Moral. Mas, para além disso, se o experimento mental de H&T é realmente plausível, então podemos construir um novo argumento da questão aberta, dessa vez direcionado ao naturalismo sintético e fornecer uma refutação definitiva do naturalismo moral. Agora, vou apresentar o argumento de H&T.

Mas, primeiro, relembremos mais algumas coisas, novamente. O argumento da questão aberta dependia de um teste para se determinar se uma questão era aberta ou fechada. Segundo Moore, uma questão é aberta se e somente se é possível para um falante competente entender completamente a questão e ainda assim não saber a resposta. Se compreender a questão e souber a resposta com base apenas nos termos da questão, então é fechada. O seu raciocínio era o seguinte. Se o naturalismo analítico é verdadeiro, então enunciados como ‘É  $x$ , que é  $N$ , bom?’ são analiticamente verdadeiros. Qualquer falante competente

saberia, atentando apenas para os termos da questão, que o enunciado é verdadeiro. Neste sentido, questões da seguinte forma teriam que ser necessariamente fechadas: ‘É  $x$ , que tem a propriedade  $N$ , bom?’ Se o naturalismo analítico é verdadeiro, então é certo que se um  $x$  tem a propriedade  $N$ , e  $N$  é analiticamente equivalente a ‘bom’, então  $x$  tem a propriedade de ser bom. Um falante competente saberia isso *a priori*. No entanto, a questão não é fechada, sustentou Moore. Portanto, o naturalismo analítico é falso. H&T sugerem que há um modo reverso de testar a validade do naturalismo semântico analítico ao modo de Moore<sup>216</sup>. Basta formular a mesma questão substituindo ‘bom’ por ‘ $N$ ’. A questão ficaria assim: ‘É  $x$ , que é bom,  $N$ ?’ A questão é aberta, pelos mesmos motivos que a anterior. Então, o naturalismo analítico é supostamente falso.

Agora, vamos considerar as seguintes questões.

- (a) É  $x$ , que é  $H_2O$ , água?
- (b) É  $x$ , que é água,  $H_2O$ ?

A definição de água é  $H_2O$ . Esta é uma definição sintética. Do mesmo modo, o naturalista está sustentando que a definição das propriedades morais em termos de um conjunto de propriedades naturais é também sintética. Como afirmam H&T, (a) e (b) em um sentido são questões abertas, pois os falantes competentes não precisam necessariamente saber que a definição de ‘água’ é ‘ $H_2O$ ’. Mas se considerarmos o experimento mental de Putnam, vemos que ‘água’ é um designador rígido. Isso quer dizer que fixa o referente  $H_2O$  em todos os mundos possíveis em que existe. Se assim é, e pressupondo a tese sobre a regulação causal, então podemos reformular (a) e (b) no seguinte sentido.

- (c) Dado que o uso de ‘água’ pelos humanos é causalmente regulado pelo tipo natural  $H_2O$ , é  $x$ , que é  $H_2O$ , água?
- (d) Dado que o uso de ‘água’ pelos humanos é causalmente regulado pelo tipo natural  $H_2O$ , é  $x$ , que é água,  $H_2O$ ?

O experimento mental de Putnam nos mostra as seguintes coisas: (i) que ‘água’, tal como usada pelos habitantes da Terra (T-água), designa

---

<sup>216</sup> Cf. H&T, 1992b, *Op. Cit.*, p. 155.

rigidamente o tipo natural  $H_2O$ ; (ii) que ‘água’, tal como usada pelos habitantes da Terra Gêmea (TG-água), designa rigidamente o tipo natural XYZ; (iii) que, portanto, ‘T-água’ e ‘TG-água’ diferem em significado; (iv) que, dado i, ii e iii, a resposta para as questões (c) e (d) é, obviamente, ‘sim’. Isso porque a propriedade natural que regula causalmente o uso de ‘T-água’ é  $H_2O$ , e, como o experimento de Putnam mostra, isso é uma verdade necessária. Então, a questão sobre se um  $x$ , que é  $H_2O$ , é água, é necessariamente fechada, pois ‘água’ designa rigidamente  $H_2O$ . O mesmo vale para a questão sobre se um  $x$ , que é água, é  $H_2O$ <sup>217</sup>. Portanto, (c) e (d) são questões fechadas.

Lembremos agora que o naturalista está sustentando que a semântica dos termos morais é similar à semântica dos termos naturais, como em ‘água =  $H_2O$ ’. Isso significa que os termos morais têm de ser designadores rígidos. Se o naturalismo é verdadeiro, então devemos poder substituir os termos ‘água’ e ‘ $H_2O$ ’ das questões (c) e (d) por quaisquer termos morais e obter o mesmo resultado, quer dizer, a questão deve continuar sendo fechada. Teríamos algo como o seguinte.

- (e) Dado que o uso de ‘bom’ pelos humanos é causalmente regulado pela propriedade natural  $N$ , é  $x$ , que é  $N$ , bom?
- (f) Dado que o uso de ‘bom’ pelos humanos é causalmente regulado pela propriedade natural  $N$ , é  $x$ , que é bom,  $N$ ?

Se o naturalismo é verdadeiro, então ‘bom’ é um designador rígido. Além disso, devemos poder tirar as quatro mesmas conclusões que tiramos de (c) e (d) acima. Mas, é isso possível?

H&T sustentam que não. Se ‘bom’, tal como usado pelos habitantes da Terra (‘T-bom’), fosse um designador rígido, então fixaria a mesma referência em todos os mundos possíveis. Isso também vale para o ‘bom’ tal como usado pelos habitantes da Terra Gêmea Moral (‘TGM-bom’). Para lembrarmos o experimento moral, ‘T-bom’ designaria rigidamente o conjunto de propriedades funcionais consequencialistas e ‘TGM-bom’ designaria rigidamente o conjunto de propriedades funcionais deontológicas. Se isso fosse o caso, então as questões (c) e (d) seriam fechadas. No entanto, como o experimento da TGM tem mostrado, se ‘T-bom’ e ‘TGM-bom’ fossem designadores

---

<sup>217</sup> Obviamente, ‘água’ nas questões (c) e (d) se referem somente aos habitantes da Terra.

rígidos, então o suposto desacordo sobre o exemplo da mentira, na verdade, não seria um desacordo genuíno, pois cada um estaria a usar o termo moral à sua maneira. Mas, como sustentam H&T, a nossa intuição é de que realmente há um desacordo moral genuíno. Portanto, ‘T-bom’ e ‘TGM-bom’ não podem ser designadores rígidos. Se não, então não podemos tirar as mesmas conclusões de (e) e (f) que tiramos de (c) e (d). Assim, (e) e (f) são questões abertas. A conclusão de H&T é de que, se o naturalismo se compromete com a tese de que as questões (e) e (f) são fechadas, então o naturalismo deve ser falso.

Esta é uma versão renovada do argumento da questão aberta, especialmente adaptada para o naturalismo não reducionista. A estratégia de H&T é clara. Se era o naturalismo não reducionista que nos impedia de dizer que o naturalismo moral é falso já que o argumento da questão aberta se aplicava somente ao naturalismo reducionista, agora podemos dizer que o naturalismo moral é definitivamente falso, pois temos uma versão do argumento de Moore que se aplica também ao naturalismo não reducionista. A validade deste argumento, obviamente, depende da veracidade da conclusão extraída por H&T do experimento sobre a TGM, pois a premissa fundamental do argumento é de que termos morais não são designadores rígidos e o experimento mental busca mostrar justamente isso.

## 2. Algumas objeções

H&T propuseram um persuasivo argumento contra o naturalismo não reducionista, de modo que não há como ser um naturalista e desconsiderar esse argumento. Diante desse problema, o naturalista pode fazer duas coisas. Ou fornece razões negativas, sustentando que o argumento de H&T falha por tais e tais motivos, ou aceita as consequências do argumento e tenta defender o naturalismo positivamente, quer dizer, revendo sua teoria semântica, por exemplo. O que se tem feito em maior proporção ultimamente, como estratégia para defender o naturalismo moral, se refere à adoção da primeira dessas alternativas, isto é, fornecer razões para recusarmos o argumento de H&T. Nesse sentido, a partir de agora irei apresentar algumas objeções com o intuito de mostrar que há algumas saídas possíveis para o adepto do naturalismo. Essas objeções certamente são passíveis de réplicas, não sendo, portanto, conclusivas. Como já afirmei acima, o objetivo não é apresentar uma posição definitiva, pois isso está além dos objetivos desse trabalho. O que gostaria de apontar aqui é apenas que o naturalista

moral possui alguns caminhos iniciais de escape para a objeção supostamente fatal de H&T.

A epígrafe deste capítulo expressa a ideia central por trás da recusa do argumento de H&T. O ponto principal do argumento é que se o naturalismo é verdadeiro, então deveríamos ter, no cenário moral, exatamente as mesmas intuições que temos no cenário de Putnam. O ponto deles – de H&T – é que não temos as mesmas intuições e, por isso, o naturalismo é falso. Mas isso pressupõe que haja alguma simetria entre o experimento de Putnam e o experimento moral e que, além disso, que as nossas intuições sobre o experimento contam como fortes razões a favor do significado dos termos morais. Esse é, supostamente, o ponto fraco do argumento. E uma das principais estratégias tem sido a de rejeitar a intuição de que há desacordo moral entre terráqueos e terráqueos gêmeos. Se analisarmos a literatura filosófica que tenta recusar o argumento, veremos que há pelo menos seis objeções mais usuais.

E. Gampel e S. Laurence, E. Margolis e A. Dawson<sup>218</sup> sugerem que a intuição não é uma boa ferramenta a favor de H&T. O argumento é o seguinte. Eles dizem que no experimento de Putnam, há a pressuposição de que a referência de ‘água’ para os terráqueos é H<sub>2</sub>O e que, a não ser que pressuponhamos um cenário cético, não é uma questão discutível se a referência de ‘água’ é *realmente* H<sub>2</sub>O. Não há outra teoria diferente, ou outras teorias, que reivindicuem uma referência distinta para o termo em questão. O mesmo vale para a referência de ‘água’ na Terra Gêmea. Somos apresentados ao fato de que se trata de uma substância denominada XYZ, e não há disputa sobre se é XYZ ou não. Em termos simples, podemos dizer que não há teorias *em competição* no experimento de Putnam. As teorias que dizem que ‘água’ é H<sub>2</sub>O na Terra e XYZ na Terra Gêmea não estão em disputa, são teorias isoladas. Mas no experimento de H&T, é uma questão discutível se a referência de ‘bom’ é melhor caracterizada por uma teoria deontológica ou consequencialista. Na verdade, a maior parte das discussões em ética se dá devido à discussão dessas teorias competidoras. E do fato de não haver um consenso acerca de qual é a melhor teoria para estipular a referência do termo moral ‘bom’, se segue que não tendemos a ver com certeza, como no caso de Putnam, qual é a

---

<sup>218</sup> Cf. GAMPEL, E. H. Ethics, Reference and Natural Kinds. *Philosophical Papers* 26: 2, p- 147-163, 1997 e LAURENCE, S., MARGOLIS, E. & DAWSON, A. Moral Realism and Twin Earth. *Facta Philosophica* 1, p. 135-165, 1999.

real referência do termo moral. Quer dizer, não somos apresentados de antemão a duas referências inteiramente diferentes como ‘H<sub>2</sub>O’ e ‘XYZ’. Isso porque é uma questão de disputa se a referência é melhor caracterizada por uma teoria consequencialista ou não consequencialista. Se não vemos com certeza que a referência do termo moral na Terra e na Terra Gêmea Moral é de fato diferente, então ficamos em dúvida se as referências são realmente diferentes. Segundo esses autores, isso *favorece* a nossa intuição a favor de que eles estão tendo um desacordo moral genuíno.

Gampel ainda insiste e apresenta uma segunda objeção: o naturalista poderia aceitar o relativismo e aceitar que as propriedades morais são reguladas por propriedades naturais diferentes, ele diz. Assim, não haveria desacordo<sup>219</sup>.

Uma terceira objeção ao argumento de H&T têm sido a de apontar para a diferença entre as propriedades em questão no experimento de Putnam e no experimento de H&T. No experimento de Putnam, ‘água’ não é uma propriedade *funcional*. Quer dizer, não se trata de um termo cuja referência é dada por um número indeterminado de propriedades de outro nível, que são diferentes entre si, mas que compartilham alguma característica comum entre todas que as faz serem fixadas pelo termo em questão. A referência de ‘água’ não depende de um conjunto de propriedades que realiza multiplamente ‘água’, mas apresenta condições necessárias e suficientes. Para um objeto ser água na Terra, é condição necessária e suficiente que seja H<sub>2</sub>O. Para ser água na Terra Gêmea, é condição necessária e suficiente que seja XYZ. Mas no experimento de H&T estamos lidando com uma propriedade funcional. Os objetos que compõem a referência de ‘bom’ não são condições suficientes para determinar a sua extensão. ‘Bom’ é multiplamente realizado por um número indefinido de instâncias naturais. Assim, tendemos a ver ‘bom’ como pertencendo a uma ordem superior tendo a referência composta por objetos de ordem inferior. Isso, como a objeção sugere, pode explicar o seguinte. Pode explicar porque temos a intuição de que há desacordo genuíno entre os habitantes dos dois mundos. Temos a *tendência* a ver a referência de ‘bom’, consequencialista e deontológica, como extensão funcional de uma única e mesma propriedade. Por isso temos a intuição de que eles estão falando de uma mesma coisa quando falam de moralidade<sup>220</sup>.

<sup>219</sup> Cf. GAMPEL, 1997, *Op. Cit.*, p. 152-153.

<sup>220</sup> Cf. LAURENCE, MARGOLIS & DAWSON, 1999, *Op. Cit.*, GEIRSSON, H. Moral Twin Earth: The Intuitive Argument. *Southwest Philosophy Review*,

Uma quarta objeção a H&T foi sugerida por D. Copp<sup>221</sup>. Copp tenta fornecer uma explicação do porquê temos a intuição de que os habitantes da Terra e da Terra Gêmea estão em desacordo. Ele diz que da forma como H&T construíram o experimento mental, nos parece que ‘T-bom’ e ‘TG-bom’, embora tenham *referentes* distintos, têm o mesmo *significado*. Isso porque tanto na Terra quanto na Terra Gêmea se usa ‘bom’ para as mesmas coisas; para aprovar, aplicar a certas situações e não em outras, direcionar cursos de ação, e assim por diante. Segundo Copp, isso explicaria porque temos a intuição de que há desacordo entre terráqueos e terráqueos gêmeos, mas que, na verdade, a referência dos termos não é a mesma como o argumento de H&T supostamente mostra.

Uma quinta objeção ao argumento de H&T foi proposta por vários filósofos<sup>222</sup>. Esses autores sugerem o seguinte. A forma como H&T constroem o seu experimento de algum modo contribui para o alcance da intuição que eles querem extrair. Como já consideramos anteriormente, no experimento de Putnam, não há controvérsias sobre as teorias que reivindicam que ‘água’ é H<sub>2</sub>O na Terra e XYZ na Terra Gêmea. Somos apresentados a duas teorias que fixam referências distintas, H<sub>2</sub>O e XYZ. Mas no experimento moral, o consequencialismo e o deontologismo concordam sobre muita coisa. Se poderia dizer que há um nível comum de moralidade que não é objeto de disputa entre essas duas teorias. Ambas concordam que há algo de errado no assassinado, que a igualdade entre as pessoas é uma coisa boa, que a honestidade é uma virtude e assim por diante. Isso significa, que ambas as teorias fixam a referência de ‘bom’ de maneira similar. Grande parte dos objetos que constituem a referência de ‘T-bom’ constitui, também, a referência de ‘TG-bom’. Se é assim, a objeção conclui, então é natural que tenhamos a intuição de que há desacordo moral genuíno entre terráqueos e terráqueos gêmeos, pois, mesmo que os termos morais sejam causalmente regulados por propriedades funcionais putativamente distintas, há uma grande extensão de casos que são similares. E no experimento de Putnam, não há similaridade alguma entre H<sub>2</sub>O e XYZ.

19 (1): p. 115-124, 2003 e GEIRSSON, H. Moral Twin-Earth and Semantic Realism. *Erkenntnis*, 62, p. 353-378, 2005.

<sup>221</sup> Cf. COPP, D. Milk, Honey, and the Good Life on Moral Twin Earth. *Synthese*, Volume 124, No. 1, 2000, p. 113-137.

<sup>222</sup> Cf. GEIRSSON, 2003, 2005 *Op. Cit.*, COPP, 2000, *Op. Cit.*, GAMPEL 1997, *Op. Cit.*, MERLI, D. Return to Moral Twin Earth. *Canadian Journal of Philosophy* 32: 2, p. 207-240, 2002; BRINK, D. Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy*, 18 (2), 2001, p. 154-176.

Há ainda uma sexta objeção a H&T que chama atenção para a dificuldade de se isolar as propriedades morais da forma como o mundo é organizado. E essa proposta pode ser vista em trabalhos de autores como Merli, Brink, Gampel e Laurence, Margolis e Dawson<sup>223</sup>. A objeção propõe o seguinte. Como isolar as propriedades morais do sistema conceitual, crenças, a forma de organizar o mundo, etc? Parece que as propriedades morais têm direta influência na forma como as pessoas organizam uma sociedade, um país, um modo de vida. As pessoas tendem a prender quem comete um assassinato porque isso é *errado*; tendem a arrecadar dinheiro através de impostos para investir em educação e saúde porque isso é o *correto*; tendem a aceitar formas de governo democrático porque essa é a *melhor* forma de governo; e assim por diante. De alguma forma, as propriedades morais têm relação direta com o modo como organizamos o mundo. Agora, se os dois mundos são similares em tudo, exceto com respeito às propriedades naturais que regulam o uso dos termos morais, então como garantir que as propriedades morais são diferentes na terra e na terra gêmea? Se forem diferentes, então isso parece implicar que a forma como as pessoas organizam o mundo deva ser diferente também. Mas esse não é o caso no modo como H&T constroem o argumento. Eles dizem que os mundos são exatamente similares, com exceção somente das propriedades naturais que regulam o uso de ‘bom’. Nesse sentido, a objeção conclui que ao sermos apresentados a dois mundos exatamente iguais, e dado que há uma relação estrita entre as propriedades morais e a forma como o mundo é organizado, *tendemos* a achar intuitivo que quando os habitantes da Terra e da Terra Gêmea falam em ‘bom’, ‘correto’ e ‘errado’, eles estão falando sobre as mesmas coisas. Isso supostamente explica porque temos a intuição de que há desacordo. Um argumento similar a isso seria apelar para a ideia de superveniência e colocar H&T numa espécie de dilema, como Merli faz<sup>224</sup>: se propriedades morais supervem às propriedades naturais e os dois mundos são similares, então as propriedades morais devem ser as mesmas; se não são as mesmas, então deve haver diferença no mundo natural, o que implica que os mundos não são similares.

Como vemos, há uma série de respostas possíveis que o naturalista pode reivindicar contra o argumento da terra gêmea moral. Se essas respostas apresentam condições suficientes para recusarmos o

---

<sup>223</sup> MERLI, 2002, *Op. Cit.*, BRINK, 2001, *Op. Cit.*, LAURENCE, MARGOLIS & DAWSON, 1999, *Op. Cit.*, e GAMPEL, 1997, *Op. Cit.*,

<sup>224</sup> MERLI, 2002, *Op. Cit.*, p. 210ss.

argumento proposto por H&T é uma questão em aberto<sup>225</sup>. Assim como é uma questão em aberto se o argumento da terra gêmea moral nos dá razões suficientes para recusarmos o realismo moral naturalista. No entanto, mesmo que não nos dê, é notório que H&T formularam um problema para o naturalismo moral digno de muita atenção. Como temos percebido, a estratégia dos naturalistas tem sido negativa, quer dizer, eles têm apresentado objeções ao argumento de H&T. Possivelmente, uma resposta mais consistente da parte do naturalista seja uma resposta positiva. Isso iria requerer que se revisitasse a teoria semântica proposta por Boyd, algo que não tem sido feito de forma muito notável hoje, mas que merece um estudo mais aprofundado. Talvez o próximo grande passo a favor do naturalismo moral seja apresentar uma defesa positiva diante do argumento da terra gêmea moral propondo uma teoria semântica alternativa.

## OBSERVAÇÕES FINAIS

Neste capítulo final, tentei fazer duas coisas: apresentar o argumento da terra gêmea moral e fornecer algumas estratégias de resposta. Vimos que H&T formularam um experimento mental com o intuito de mostrar que o naturalismo não reducionista é falso. O naturalista está alegando que a semântica dos termos morais é similar à semântica dos termos não morais. E, dado que a teoria semântica proposta por Boyd é um desenvolvimento da teoria externalista de Putnam, então, se o naturalismo for verdadeiro, os termos morais devem se comportar do mesmo modo que os termos não morais no cenário de Putnam, afirmam H&T. No entanto, segundo H&T, não é isso que acontece. Quando imaginamos dois mundos em que a extensão de ‘bom’ é diferente e supomos que os habitantes de cada mundo discordam sobre uma questão moral qualquer, temos a intuição de que realmente há desacordo moral genuíno, e não meramente semântico – o que teria que ser o caso se o naturalismo moral fosse verdadeiro. E isso significa que o significado dos termos morais não pode ser determinado inteiramente pela regulação causal natural, como quer Boyd, por exemplo. Assim, H&T sustentam que o naturalismo é falso. Como vimos, isso lhes possibilita construir uma reformulação do argumento da questão aberta com base na suposta não rigidez dos termos morais. Além disso, vimos que o naturalista pode lançar mãos de várias respostas ao argumento

---

<sup>225</sup> Algo que será objeto de minhas pesquisas futuras.

proposto por H&T apelando principalmente para a dissimilaridade entre o experimento mental de Putnam para com termos naturais e o experimento proposto por H&T e para a intuição como critério confiável para extrairmos conclusões sobre semântica moral.



## CONCLUSÃO

A proposta deste trabalho foi analisar a resposta naturalista para o problema do que são as propriedades morais. O objetivo fora mostrar que, mesmo que o naturalismo não reducionista não seja vulnerável a maioria das objeções tradicionais, ele precisa fornecer uma resposta ao argumento da terra gêmea moral se quiser reivindicar plausibilidade, pois este é o seu principal problema. Além disso, ao final, algumas estratégias iniciais de resposta ao problema apresentado por H&T foram traçadas.

Para finalizar, devo lembrar sobre alguns pontos principais deste trabalho. O primeiro capítulo fora dedicado a uma série de esclarecimentos introdutórios, bem como, principalmente, a alguns críticos do naturalismo moral. Vimos que o argumento da questão aberta buscou refutar o naturalismo moral apelando para a suposta disparidade semântica entre os termos morais e não morais e que exerceu profunda influência no desenvolvimento de teorias metaéticas posteriores. O emotivismo moral fora uma estratégia notável como teoria que aceitou o argumento da questão aberta, mas se recusou a aceitar o intuicionismo de Moore, desenvolvendo-se, assim, o não cognitivismo. Como pudemos perceber, para os não cognitivistas as propriedades morais não são parte do mundo tal qual o naturalista assume serem, mas são expressão de atitudes dos agentes morais. Hare, mesmo concordando com os emotivistas sobre a não factualidade semântica e o não cognitivismo psicológico, argumentou que o emotivismo implica na não racionalidade da moral e desenvolveu o prescritivismo universal. Como vimos, essa teoria representa uma série de críticas ao naturalismo moral, tais como: uma reformulação do argumento da questão aberta; a ideia de que não se pode derivar conclusões sobre o *dever ser* a partir de premissas sobre o *ser* somente; a tese da prescritividade, que implica que o naturalismo não consegue dar conta do suposto elemento motivacional das propriedades morais; e a tese de que o naturalismo moral implica em relativismo. Ainda, ao final do primeiro capítulo, analisamos a teoria do erro de Mackie. Como vimos, a ideia de Mackie constitui uma crítica ao naturalismo na medida que estipula que as propriedades morais são uma espécie de erro e que não existem objetivamente, sendo as sentenças morais uniformemente falsas. Pode-se dizer que depois do argumento da questão aberta e das teorias metaéticas que se seguiram, houve um período não naturalista, por assim dizer.

No entanto, como vimos no decorrer deste trabalho, principalmente a partir da década de oitenta os tempos começaram a mudar com o surgimento de um novo naturalismo. Nesse sentido, dediquei o segundo capítulo para apresentar as principais ideias do naturalismo não reducionista. Pudemos perceber que esta vertente do naturalismo moral pretendeu se beneficiar de vários avanços na metafísica, filosofia da linguagem, filosofia da mente e epistemologia para argumentar a favor de um novo naturalismo. Destaquei algumas ideias principais, como a recusa da ideia de que se dois termos são sinônimos então devem referir o mesmo objeto, a ideia de que alguns termos singulares e tipos naturais são designadores rígidos, a tese de que os estados mentais são propriedades funcionais, a recusa da teoria semântica descritivista e o desenvolvimento das teorias semânticas externalistas, e a ideia de que propriedades morais desempenham um papel explicativo. Como tentei mostrar, os não reducionistas, representados por Brink, Boyd e Sturgeon, argumentam que algumas objeções tradicionalmente endereçadas à ideia de que propriedades morais são naturais não se aplicam a essa nova vertente do naturalismo moral. Além disso, busquei apresentar a teoria semântica do naturalismo não reducionista tal como formulada e defendida por Boyd. Vimos que essa teoria semântica se caracteriza por sustentar que a extensão das propriedades morais são agregados de propriedades homeostáticas reunidos por certos mecanismos. Nesse sentido, a referência dos termos morais não é constituída por uma propriedade natural, como um naturalista reducionista poderia sustentar, mas por um conjunto indeterminado de propriedades naturais que são descobertos pelas melhores teorias normativas. E, por fim, procurei apresentar outra estratégia que os naturalistas não reducionistas normalmente fazem uso para sustentar que propriedades morais são naturais. Essa estratégia consiste em dizer que propriedades morais desempenham um papel explicativo. O raciocínio é similar ao dos filósofos que defendem o realismo científico, quer dizer, a melhor explicação para certos fatos é a postulação de que certas entidades existem, mesmo que não consigamos as observar. Vimos que o naturalista sustenta que a melhor explicação para a atribuição de certas propriedades morais a certos atos ou pessoas é a postulação de que existem fatos morais. O que tentei deixar claro nesse capítulo foi que, com o surgimento desses novos argumentos dos não reducionistas, o realismo moral naturalista voltou a atrair adeptos.

Finalmente, no último capítulo, busquei mostrar que, mesmo que admitamos as supostas vantagens do naturalismo não reducionista, há um problema fundamental que o naturalista tem de enfrentar, o

argumento da terra gêmea moral. Desde a publicação dos artigos de H&T, esse problema tem direcionado boa parte da discussão sobre a plausibilidade do naturalismo moral, de modo que defender o naturalismo moral implica em se ter uma resposta a esse problema. Como vimos, o argumento desafia a teoria semântica proposta por Boyd de que a extensão dos termos morais é causalmente regulada por mecanismos externos aos agentes. Busquei apresentar também a nova versão do argumento da questão aberta que deriva o experimento mental de H&T. E, para concluir, falei brevemente sobre algumas alternativas de resposta que o naturalista poderia reivindicar contra o problema formulado por H&T. Como o leitor pôde perceber, não me comprometi em apresentar uma posição definida sobre se o argumento da terra gêmea moral é condição suficiente para recusarmos o naturalismo moral. Limitei-me a mostrar que há alguns caminhos possíveis que o adepto do naturalismo moral pode tentar adotar para escapar ao problema proposto por H&T. Se esses caminhos constituem razão última para recusarmos o argumento de H&T é um problema que deixo em aberto.



## REFERÊNCIAS BIBLIOGRÁFICAS

- AUSTIN, J. L. *How to do things with words*. Oxford: Clarendon Press, 1962.
- AYER, A. J. *Language, Truth and Logic*. Harmondsworth: Penguin Books, 1971.
- BOYD, R. Metaphor and Theory Change. In: ORTONY, A. (ed), *Metaphor and Thought*, Cambridge: Cambridge University Press, 1979.
- \_\_\_\_\_. How to be a Moral Realist. In: SAYRE-MCCORD, G. (ed) *Essays on Moral Realism*, Ithaca: Cornell University Press, 1988.
- BRINK, D. O. Moral Realism and Skeptical Arguments for Disagreement and Queerness. *Australian Journal of Philosophy*, 62:2, p. 111-125, 1984
- \_\_\_\_\_. *Moral Realism and the Foundations of Ethics*, Cambridge: Cambridge University Press, 1989.
- \_\_\_\_\_. Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy*, 18 (2), p. 154-176, 2001
- CARNAP, R. *Meaning and Necessity: A Study in Semantics and Modal Logic*. The University of Chicago Press: Chicago, 1948.
- COPP, D. Milk, Honey, and the Good Life on Moral Twin Earth. *Synthese*, Volume 124, No. 1, p. 113-137, 2000.
- \_\_\_\_\_. Why Naturalism? *Ethical Theory and Moral Practice*, Vol. 6, p. 179-200, 2003.
- DARWALL, S., GIBBARD, A. & RAILTON, P. *Moral discourse & Practice: Some Philosophical Approaches*. New York, Oxford: Oxford University Press, 1997.
- DE CARO, M & MACARTHUR, D. Science, Naturalism, and the Problem of Normativity. In: DE CARO, M & MACARTHUR, D. *Naturalism and Normativity* (eds). Columbia University Press: New Yourk, p. 1-19, 2010.

DEVITT, M. & STERELNY, K. *Language and Reality: An Introduction to the Philosophy of Language*. Second Edition, Blackwell Publishers: United States, MIT Press, 1999.

FISHER, A. *Metaethics: An Introduction*. Acumen, 2011.

FRANKENA, W. The Naturalistic Fallacy. *Mind*, No. 48, p. 464-477, 1938.

FREIMAN, C. Goodness and Moral Twin Earth. *Erkenntnis* 79: p. 445-460, 2014.

GAMPEL, E. H. Ethics, Reference and Natural Kinds. *Philosophical Papers* 26: 2, p. 147-163, 1997.

GEACH, P. Ascriptivism. *Philosophical Review*, 69, p. 221-225, 1960.

GEIRSSON, H. Moral Twin Earth: The Intuitive Argument. *Southwest Philosophy Review*, 19 (1): p. 115-124, 2003.

\_\_\_\_\_. Moral Twin-Earth and Semantic Realism. *Erkenntnis*, 62, p. 353-378, 2005.

HARE, E. J. *God and Morality: a Philosophical History*. Oxford, Blackwell Publishing, 2007.

HARE, R. M. Imperative Sentences. In: *Mind*, New Series, Vol. 58, No. 229, p. 21-39, 1949.

\_\_\_\_\_. *The Language of Morals*, New York: Oxford University Press, 1952.

\_\_\_\_\_. *Freedom and Reason*. Oxford: Oxford Clarendon Press, 1963.

\_\_\_\_\_. *Moral Thinking, its levels, method and point*. Oxford: Oxford Clarendon Press, 1981.

\_\_\_\_\_. *Sorting Out Ethics*. Oxford: Oxford Clarendon Press, 1997.

\_\_\_\_\_. Internalism and Externalism in Ethics. In: HARE, R. M. *Objective Prescriptions and Other Essays*. Oxford: Clarendon Press, 1999.

HARMAN, G. *Moral Relativism Explained*, p. 1. (Artigo não publicado), 2012.

\_\_\_\_\_. *The Nature of Morality*. New York: Oxford University Press, 1977.

HORGAN, T. & TIMMONS, M. New Wave Moral Realism Meets Moral Twin Earth. *Journal of Philosophical Research*, Volume 16, p. 447-465, 1991.

\_\_\_\_\_. Troubles on Moral Twin Earth: Moral Queerness Revived. *Synthese*, 92, p. 221-260, 1992a.

\_\_\_\_\_. Troubles for New Wave Moral Semantics: The Open Question Argument Revived. *Philosophical Papers*, Volume 21, No. 3, p. 153-175, 1992b.

\_\_\_\_\_. Copping Out On Moral Twin Earth. *Synthese* 124, p. 113-137, 2000.

\_\_\_\_\_. Analytical Moral Functionalism Meets Moral Twin Earth. In: RAVENSCROFT, I. (Ed.) *Essays on the Philosophy of Frank Jackson*. Oxford: Blackwell Publishers, p. 221-230, 2009.

KANT, I. *Fundamentação da Metafísica dos Costumes*. Tradução de Guido Antônio de Almeida. São Paulo: Discurso Editorial: Barcarolla, 2009.

KIRCHIN, S. *Metaethics*. Palgrave Macmillan, 2012.

KRIPKE, S. Identity and Necessity. In: SCHWARTZ, S. P. *Naming, Necessity and Natural Kinds*. Ithaca: Cornell University Press, p. 67-101, 1977.

\_\_\_\_\_. *Naming and Necessity*. Cambridge, MA: Harvard University Press, 1980.

LAURENCE, S., MARGOLIS, E. & DAWSON, A. Moral Realism and Twin Earth. *Facta Philosophica* 1, p. 135-165, 1999.

LEVIN, J. "Functionalism", *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2013/entries/functionalism/>.

LEVINSON, I. J. *Introduction to Mechanics*. Second Edition: Prentice-Hall, Englewood Cliffs, 1968.

LOCKE, J. Ensaio Acerca do Entendimento Humano. In.: LOCKE, J. *Carta Acerca da Tolerância; Segundo Tratado Sobre o Governo; Ensaio Acerca do Entendimento Humano*. Tradução de Anoar Aiex e E. Jacy Monteiro, 2. Ed., São Paulo: Abril Cultural, 1978.

MACKIE, J. *Ethics: Inventing Right and Wrong*. New York: Penguin, 1977.

MERLI, D. Return to Moral Twin Earth. *Canadian Journal of Philosophy* 32: 2, p. 207-240, 2002.

MILL, J. S. *Utilitarianism*. New York, Bantan Books, 1993.

MILLER, A. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press, 2011.

MOORE. G. E. *Principia Ethica*. (Revised Edition). Cambridge: Cambridge University Press, 1993.

OGDEN, C. K. & RICHARDS, I. A. *Meaning of Meaning: A Study of the Influence of Language upon Thought and of the Science of Symbolism*. Harcourt, Brace & World: New York, 1946.

OKASHA, S. *Philosophy of Science: a very short introduction*. Oxford: Oxford University Press, 2002.

PUTNAM, H. The Meaning of "Meaning". In.: *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University Press, p. 215-271, 1975.

\_\_\_\_\_. Brains and Behaviour. In.: *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University press, p. 325-341, 1975.

\_\_\_\_\_. The Nature of Mental States. In.: *Mind, Language and Reality*. Philosophical Papers, Volume 2, Cambridge University press, p. 429-440, 1975.

QUINN, W. Truth and Explanations in Ethics. *Ethics*, Vol. 96, No. 3, p. 524-544, 1986.

RAWLS, J. *A Theory of Justice* (Revised Edition). Cambridge, Massachusetts: Harvard University Press, 1971/1999.

RICHARDS, I. A. *Principles of Literary Criticism*. Routledge Classics: London and New York, 2004.

RUBIN, M. Synthetic Ethical Naturalism. University of Massachusetts, *Dissertations*, Paper 24, 2009.

SAYRE-MCCORD, G. Moral Theory and Explanatory Impotence. *Essays on Moral Realism*, Ithaca, New York: Cornell University Press, 1988.

\_\_\_\_\_. 'Good' On Twin Earth. *Philosophical Issues* 8, p. 267-292, 1997.

SHAFER-LANDAU, R. *Moral Realism: A Defense*. Clarendon Press: Oxford, 2003.

\_\_\_\_\_. Ethics as Philosophy: A Defense of Ethical Non naturalism. In.: HORGAN, T. & TIMMONS, M. (eds). *Metaethics after Moore*. New York: Oxford University Press, p. 209-232, 2006.

SMART, J. J. C. Sensations and Brain Processes. *The Philosophical Review*, Vol. 68, No. 2. 1959.

SOAMES, S. *Philosophical Analysis in the Twentieth Century*. Volume 2, Princeton University Press, 2003.

STEVENSON, C. L. *Ethics and Language*. New Heaven, Yale University Press, 1958.

STURGEON, N. Moral Explanations. In: David Copp and David Zimmerman (eds.), *Morality, Reason and Truth*. Totowa, NJ: Rowman & Littlefield, 1985a.

\_\_\_\_\_. Moore on Ethical Naturalism. *Ethics*, 113 (3), p. 528-556, 2003.

\_\_\_\_\_. Ethical Naturalism. In: COPP, D. (ed) *The Oxford Handbook of Ethical Theory*. Oxford: Oxford University Press, 2006.

TIMMONS, M. *Morality without Foundations*. Oxford: Oxford University Press, 1999.

VAN ROOJEN, M. Knowing Enough to Disagree: A New Response to the Moral Twin Earth Argument. In: SHAFER-LANDAU, R. (Org.) *Oxford Studies in Metaethics*. Volume 1. Oxford: Clarendon Press, p. 162-193, 2006.

WHER, M. R. & RICHARD, J. A. *Física do Átomo*. Trad. de Carlos Campos de Oliveira. Ao Livro Técnico S. A.: Rio de Janeiro, 1965.