

# FICCIONALIZAÇÃO DA MORAL: UM ROTEIRO

## FICTIONALIZING ETHICS: A GUIDELINE

WILSON MENDONÇA E IDIA LAURA FERREIRA  
(UFRJ – Brasil)

### RESUMO

O trabalho identifica algumas questões críticas em tentativas recentes de tratar o objeto da moralidade como uma ficção. Sobre o pano de fundo de uma distinção entre cognitivismo e não-cognitivismo moral exclusivamente em termos da natureza do estado mental expresso pelo enunciado moral, independentemente do conteúdo semântico do enunciado (§ 1), o trabalho mostra um *deficit* estrutural no *argumento da intransigência* proposto por Kalderon a favor do ficcionalismo moral hermenêutico (§ 2). Mediante a reconstrução dos passos que levam do cognitivismo irrealista ao ficcionalismo moral revolucionário, argumenta-se que não é claro como uma moral fictícia concebida por Joyce poderia ainda ter uma influência adequada na conduta cooperativa e na resolução de conflitos de interesses (§ 3). Ademais, a diferenciação entre a interpretação *de dicto* e a interpretação *de re* da conexão necessária entre o juízo moral e a motivação revela as lacunas no argumento internalista comum a favor do não-cognitivismo e no *argumento da bizarrice* em favor do cognitivismo irrealista (§ 4). O trabalho é encerrado com algumas observações sobre o recurso à teoria humiana das razões normativas na tentativa de justificação da teoria do erro invocada pelo ficcionalismo moral revolucionário (§ 5).

**Palavras-chave:** metaética, cognitivismo moral, ficcionalismo hermenêutico e revolucionário, internalismo motivacional, internalismo de razões.

### ABSTRACT

The paper identifies some critical issues in recent attempts to treat the subject matter of morality as a fiction. On the background of a distinction between moral cognitivism and noncognitivism exclusively in terms of the nature of the mental state expressed by a moral statement regardless of the semantic content of the statement (sec. 1), the paper shows a structural deficit in Kalderon's *argument from intransigence* for hermeneutic moral fictionalism (sec. 2). By reconstructing the steps that lead from irrealist cognitivism to revolutionary moral fictionalism, it is then argued that it is not clear how a fictitious morality as envisaged by Joyce could still have an adequate impact on cooperative conduct and the resolution of conflict of interests (sec. 3). Further, the differentiation of a *de dicto* and a *de re* interpretation of the necessary connection between moral judgment and motivation reveals the flaws in the familiar internalist argument for noncognitivism and in the *argument from queerness* for irrealist cognitivism (sec. 4). The paper closes with some remarks on the appeal to the Humean theory of normative reasons in the attempted justification of the error theory invoked by Joyce's revolutionary moral fictionalism (sec. 5).

**Keywords:** metaethics, moral cognitivism, hermeneutic and revolutionary fictionalism, motivational internalism, reasons internalism

## 1 COGNITIVISMO E NÃO-COGNITIVISMO MORAL

Embora seja apresentado (e discutido) por muitos filósofos como envolvendo essencialmente uma tese sobre o significado *sui generis* de termos e sentenças morais, o não-cognitivismo é uma teoria semântica somente em um sentido secundário. No sentido primário, o não-cognitivismo é uma teoria ‘psicológica’ sobre a natureza do estado mental convencionalmente associado ao enunciado (*statement*) moral. Os não-cognitivistas afirmam comumente que o enunciado moral *expressa* um estado mental radicalmente diferente das crenças (*beliefs*) comuns, um estado mental não-cognitivo. Deve-se notar, porém, que a existência de uma relação de expressão entre enunciados e estados mentais não é a característica distintiva do não-cognitivismo: a *mesma* relação de expressão liga o proferimento de sentenças não-morais (factuais) a estados mentais. Gibbard é especialmente claro nesse contexto: ‘Expressing we explain by analogy with prosaically factual statements ... Expressing a state of mind works the same in these two cases, but the states of mind expressed are different’ (Gibbard 2003, p. 75). É conveniente fixar uma terminologia para esta distinção entre estados mentais expressos. Alguns estados mentais—cujo exemplo paradigmático é o estado mental da crença factual comum—têm uma direção de ajuste (*direction of fit*) mente-mundo. Isto é, tais estados tendem a ser abandonados quando confrontados com a percepção de que o mundo não é como o conteúdo destes estados o representa. Outros estados têm direção de ajuste mundo-mente e tipicamente persistem mesmo após a constatação de que os fatos do mundo contrariam seu conteúdo. O exemplo paradigmático aqui é o estado de desejo. Para distinguir terminologicamente os estados com direções de ajuste opostas, muitos filósofos utilizam as expressões ‘crenças’ e ‘atitudes’—ou ‘crenças’ e ‘atitudes semelhantes ao desejo’ (*desire-like attitudes*). A tese que define o não-cognitivismo pode, então, ser formulada da seguinte maneira: o estado mental convencionalmente associado ao enunciado moral é uma atitude (ou uma atitude semelhante ao desejo). Em uma formulação alternativa, a tese não-cognitivista diz que a aceitação de uma sentença moral (o endosso no pensamento do conteúdo de uma sentença moral) não é uma crença, mas uma atitude de direção de ajuste mundo-mente.

A esta tese sobre a natureza dos estados mentais expressos o não-cognitivista normalmente associa uma teoria acerca do conteúdo semântico das expressões morais. Esta vai da negação peremptória de que termos morais possam ter qualquer conteúdo factual (Ayer) à

formulação de uma proposta construtiva que estabelece uma correspondência recursiva e composicional entre termos morais e estados mentais (Blackburn e Gibbard). Qualquer que seja a versão da teoria semântica tipicamente adotada pelo não-cognitivista, ela se baseia na rejeição do *factualismo semântico*, isto é, da visão semântica dominante segundo a qual sentenças declarativas expressam (em um sentido diferente da expressão de estados mentais por enunciados) proposições que (i) determinam as condições de verdade das sentenças em questão e (ii) atribuem propriedades a objetos. Na perspectiva factualista, os predicados significam propriedades e designam as extensões correspondentes. Eles contribuem assim para a determinação das condições de verdade das sentenças onde esses predicados ocorrem. (É por causa da ligação estreita entre proposições verdadeiras, por um lado, e fatos, por outro, que a visão semântica dominante é caracterizada como factualismo.) Para o não-factualista—para o não-cognitivista típico—predicados morais não funcionam semanticamente como designadores de extensões, mas como sinalizadores da presença de um estado mental. De acordo com isso, sentenças morais, na perspectiva não-factualista, tampouco têm condições de verdade em um sentido substancial ‘correspondentista.’ Para o não-factualista, sentenças morais podem ser verdadeiras somente em um sentido minimalista ou deflacionário.

Talvez haja uma ligação ‘natural’ entre a tese ‘psicológica’ do não-cognitivista e a semântica psicologista não-factualista. Da mesma forma, a teoria semântica mais ‘naturalmente’ adequada ao cognitivismo moral parece ser o factualismo. O ponto relevante, contudo, é que a questão ‘psicológica’ sobre a natureza dos estados mentais expressos é logicamente independente da questão sobre o significado das sentenças que os expressam. Isto quer dizer que há mais possibilidades de combinação do que é sugerido pela discussão típica da oposição cognitivismo *vs.* não-cognitivismo. Eis como Kalderon, por exemplo, descreve a situação padrão: ‘We appeared to be in the uncomfortable position of choosing between a plausible semantics wedded to an implausible cognitivism and an implausible semantics wedded to a plausible noncognitivism’ (Kalderon 2005a, p. 146).

Se o ponto de vista ‘psicológico’ for separado analiticamente do ponto de vista semântico e se distinção entre cognitivismo e não-cognitivismo moral for feita exclusivamente em termos da natureza do estado mental convencionalmente associado ao enunciado moral, há pelo menos duas outras possibilidades além do cognitivismo clássico e do não-cognitivismo expressivista. De fato, algumas dessas possibilidades teóricas inovadoras foram invocadas

recentemente no debate metaético. Por razões que serão expostas mais tarde, a questão ‘psicológica’ sobre a natureza dos estados mentais expressos por enunciados morais pode ser reconstruída com vantagem como a questão sobre se o proferimento de sentenças morais é uma asserção ou outro ato de fala. Sob estas condições, uma exposição das posições teóricas possíveis com indicação de algumas teorias já propostas pode tomar a seguinte forma:

	enunciado moral = asserção	enunciado moral ≠ asserção
factualismo	realismo moral (Moore) e teoria do erro (Mackie)	ficcionalismo moral hermenêutico (Kalderon)
não-factualismo	cognitivism não-descritivista (Horgan e Timmons)	expressivismo (Blackburn, Gibbard)

Evidentemente, a mera exposição das possibilidades teóricas não decide sobre a plausibilidade das propostas correspondentes. De um ponto de vista intuitivo, o cognitivism não-descritivista é relativamente estranho. Por outro lado, o ficcionalismo moral tem-se mostrado crescentemente influente no debate metaético. A investigação das possibilidades e dos limites da adoção da atitude ficcional frente à moralidade define o foco principal das próximas seções.

## 2 FICCIONALISMO MORAL HERMENÊUTICO

Ficcionalismo a propósito de uma região do discurso é, em termos gerais, a teoria segundo a qual os enunciados nesta região *não* são asserções genuínas—não almejam a verdade, não pretendem representar os fatos, não atribuem propriedades e relações a objetos etc.—mas são ficções aceitas e mantidas por nós por causa do seu valor prático ou emotivo. O *ficcionalismo moral hermenêutico* é o resultado da aplicação do ponto de vista ficcionalista geral ao discurso e à prática da moralidade convencional atual. Como teoria não-cognitivista no sentido da definição restritiva comentada na seção anterior, o ficcionalismo moral hermenêutico desenvolve as implicações da tese segundo a qual a aceitação de uma sentença moral não é a crença na verdade da sentença em questão. De acordo com o ficcionalista hermenêutico, nós usaríamos, como participantes do jogo da linguagem da moralidade, sentenças declarativas de caráter factual ou

representacional para *expressar* nossas atitudes conativas, *não* para afirmar os fatos que determinam o conteúdo das nossas crenças. Ao declarar o aborto como moralmente errado, por exemplo, nós expressaríamos nossa desaprovação do aborto e faríamos isso invocando um mundo ficcional no qual o aborto instância a propriedade de ser moralmente errado. A sentença moral ‘o aborto é errado’ expressaria semanticamente uma proposição que determina suas condições de verdade, mas o proferimento desta sentença não afirmaria a proposição correspondente. Ao contrário, o proferimento transmitiria somente a nossa atitude negativa frente ao aborto mediante o uso não-assertórico (expressivo) de uma sentença declarativa. O *ficcionalismo moral revolucionário*, por outro lado, se apóia no cognitivismo irrealista de Mackie e propõe a substituição eficaz, do ponto de vista prático, da crença moral (*belief*) pelo assentimento ficcional (*make-belief*)—não como interpretação da moralidade atual, mas como programa de reconstrução das nossas atitudes mais ou menos abaladas pelo reconhecimento de que não há justificção epistêmica para nossas crenças morais. A questão relevante para quem está convencido que asserções morais são irremediavelmente não-verdadeiras é saber se é possível continuar de alguma maneira o discurso moral de tal forma que possamos (i) desfrutar ainda dos benefícios práticos instrumentais do discurso moral, mas (ii) sem incorrer nos custos de acreditar em falsidades (Joyce 2001, p. 185). A teoria que dá uma resposta afirmativa a essa questão, o ficcionalismo moral revolucionário, é o objeto da próxima seção. A presente seção discute criticamente uma formulação do ficcionalismo hermenêutico.

A plausibilidade do ficcionalismo hermenêutico depende da tese não-cognitivista segundo a qual o proferimento de uma sentença moral não é uma asserção genuína. O que subjaz a essa maneira de colocar o problema é o ponto de vista já mencionado na seção anterior, a saber: a questão ‘psicológica’ sobre a natureza do estado mental expresso por uma sentença moral deve ser transformada na questão sobre se o proferimento da expressão sentencial moral é ou não uma asserção. A fórmula ‘o estado mental *convencionalmente* associado ao proferimento da sentença moral’ já indicava que o problema não é propriamente psicológico, mas diz respeito à natureza das convenções e normas linguísticas que permitem a classificação de alguns atos de fala como asserções e outros não.

O ficcionalista moral Kalderon tem um argumento sobre o caráter não-assertórico do proferimento moral. O argumento da intransigência (*argument from intransigence*) é concebido por Kalderon como uma alternativa aos argumentos familiares internalistas a favor do não-

cognitivismo. (Os argumentos internalistas serão examinados na seção 4.) Centrado na consideração do desacordo moral e da investigação como empreendimento coletivo, o argumento da intransigência pode ser esquematizado da seguinte maneira:

1. No desacordo resultante de asserções contraditórias, cada uma das partes em desacordo está submetida à norma que obriga, em princípio, a continuar a investigação, reconsiderando criticamente a própria posição.
2. No caso do desacordo moral é racionalmente permitido ser intransigente. Nesse caso não existe, nem mesmo em princípio, a obrigação de continuar a investigação, reconsiderando criticamente a própria posição.
3. Portanto, os proferimentos morais não são asserções genuínas. (não-cognitivismo moral)

A premissa problemática é a segunda. Não é óbvio que a norma mencionada por Kalderon na primeira premissa não seja também aplicável ao caso do desacordo moral. Por que a intransigência seria racionalmente permitida na investigação moral, mas proibida na investigação não-moral? Em ambos os casos, o desacordo que Kalderon tem em mente é o ‘desacordo sobre razões,’ onde as partes em disputa discordam não somente sobre se devem aceitar ou rejeitar alguma sentença S, mas *principalmente* sobre o que conta, nas circunstâncias dadas, como uma razão para aceitar ou rejeitar S. Isto é possível se as partes em disputa operarem sob princípios distintos e mais ou menos incomensuráveis que levam à consideração de características distintas das circunstâncias dadas como razões para a aceitação ou rejeição de S. Ademais, como participante de um empreendimento coletivo, cada uma das partes em disputa estará interessada na aceitação ou rejeição de S não meramente para si mesma (*on behalf of oneself*, na terminologia de Kalderon), mas também e principalmente para os outros (*on behalf of others*), por exemplo, porque a verdade ou falsidade de S é relevante para a prática comum na qual todas as partes estão engajadas. Nesse caso, ao aceitar S, qualquer parte deve supor que as outras partes não aceitam razões que levem à rejeição de S, desde que elas sejam igualmente racionais e razoáveis, informadas e interessadas no valor cognitivo de S. A constatação de que uma outra parte aparentemente aceita razões para rejeitar S, sem que isso signifique irracionalidade ou qualquer outro vício cognitivo, gera *dívida reflexiva*: a parte que aceita S está racionalmente motivada a investigar mais a fundo as bases da aceitação de S, na esperança de descobrir razões mais fortes que outras pessoas igualmente racionais, razoáveis e informadas não possam rejeitar. ‘Inquiring

further into the grounds of acceptance is an obligatory end of public inquiry for those engaged in a disagreement about reasons' (Kalderon 2005a, p. 26). Ao contrário do argumento internalista motivacional 'clássico' pró não-cognitivismo, que procura estabelecer uma *diferença motivacional* entre a aceitação moral necessariamente motivadora e a crença intrinsecamente inerte, o argumento da intransigência revelaria uma *diferença epistêmica* entre a crença e a aceitação moral. Isto é, na pesquisa cognitiva os sujeitos da pesquisa têm um motivo que lhes falta na pesquisa moral: a disposição para reconsiderar as bases racionais das respectivas aceitações e rejeições de proposições em face do desacordo de razões.

Na visão de Kalderon, isto não é, porém, uma afirmação empírica sobre a psicologia dos envolvidos em disputas morais, sobre os motivos ou disposições que eles de fato têm ou deixam de ter quando em situações de desacordo de razões. Ao contrário, tratar-se-ia de uma observação *conceitual* sobre as normas epistêmicas que governam a aceitação e a rejeição de sentenças na pesquisa (Kalderon 2005a, p. 36). Como única evidência de que as normas epistêmicas atuais no caso do juízo moral têm esse caráter que as distingue radicalmente das normas aplicáveis ao caso cognitivo, Kalderon cita o registro, por parte de Putnam, do seu desacordo recorrente com Nozick sobre as bases morais do liberalismo político:

But what of the fundamentals on which one cannot *agree*? It would be dishonest to pretend that one thinks that there are no better and worse views *here*. I don't think that it is just a matter of *taste* whether one thinks that the obligation of the community to treat its members with compassion takes precedence over property rights; nor does my co-disputant. Each of us regards the other as lacking, at this level, a certain kind of sensitivity and perception. To be perfectly honest, there is in each of us something akin to *contempt*, not for the other's *mind*—for we each have the highest regard for each other's minds—nor for the other as a *person*—, for I have more respect for my colleague's honesty, integrity, kindness, etc., than I do for that of many people who agree with my 'liberal' political views—but for a certain complex of emotions and judgments in the other. (Putnam 1981, p. 165 *apud* Kalderon 2005a, p. 36)

Há espaço para pelo menos três comentários sobre o exemplo ilustrativo de Kalderon que são relevantes para a avaliação de sua tese.

Em primeiro lugar, embora o testemunho citado de Putnam possa ser visto como o relato sincero da ausência de qualquer motivação para investigar mais profundamente as bases de sua visão político-moral em desacordo com a de Nozick, a passagem citada é perfeitamente compatível com a interpretação segundo a qual a motivação relevante *estava* presente nos seus



primeiros embates com Nozick. Depois de repetidos fracassos, a disposição em Putnam para a reconsideração da sua própria posição e para a continuação da investigação conjunta simplesmente pode ter-se extinguido. Mas isso quer dizer, de acordo com o critério proposto por Kalderon, que os proferimentos morais de Putnam—e certamente também os de Nozick—*eram* asserções e o ambiente da investigação *era* cognitivo. Ocorre que a discussão chegou a um ponto em que os proferimentos morais *nesse contexto específico e especial do embate entre Putnam e Nozick* deixaram de expressar crenças e passaram a ser utilizados também—e talvez sobretudo—como veículos para a manifestação de atitudes de desprezo por uma sensibilidade e percepção morais reciprocamente consideradas como defeituosas. Não há a menor razão para generalizar as propriedades especiais dos proferimentos morais nesse contexto para os mesmos proferimentos por parte desses mesmos agentes ou de *outros* agentes em *outros* contextos. Por que os proferimentos morais de Putnam frente a outros interlocutores deveriam ser dissociados da disposição para a revisão da própria posição e para a continuação da investigação no caso de desacordos? Ademais, a ausência da motivação relevante deveria ser uma questão conceitual sobre as normas epistêmicas que governam a aceitação de sentenças morais na pesquisa. O relato sobre o que se passa na mente de Putnam quando ele discute com Nozick não pode corroborar teses sobre ligações conceituais ou sobre a forma de normas da pesquisa moral.

Em segundo lugar, tampouco há razões para pensar que todo e qualquer desacordo moral resulte nesse ponto degenerado onde não há mais espaço para a continuação da investigação e a busca de razões que se impõem a todos os envolvidos. Talvez alguns desacordos morais possam encontrar uma solução construtiva. De qualquer forma, o exemplo de Putnam e Nozick *não mostra* que isso não é possível.

Em terceiro lugar, Kalderon não pode excluir que a discussão teórico-cognitiva possa chegar a impasses semelhantes, sem que isso a descaracterize como cognitiva. Não é um ‘privilegio’ da discussão moral que o desacordo não conduza à ‘dúvida reflexiva’ e disposição para a revisão das aceitações ou rejeições correspondentes. O próprio Kalderon menciona a possibilidade de um desacordo que não gera dúvidas reflexivas na discussão teórica entre um defensor e um detrator da tese de que a terra é plana. E esse é um caso claro de investigação cognitiva. É verdade que Kalderon afirma que, nesse caso, a parte que rejeita a tese em questão pode plausivelmente questionar que haja simetria epistêmica entre as partes em disputa. E o *reconhecimento* da simetria epistêmica é, segundo Kalderon, uma das condições necessárias para



que o desacordo gere dúvida reflexiva (Kalderon 2005a, p. 21). Mas o que distingue a afirmação enfática, por parte de Putnam, de que ao seu opositor falta ‘a certain kind of sensitivity and perception’ do não-reconhecimento da simetria epistêmica entre ele e seu oponente? Se a ausência da dúvida reflexiva no caso da disputa entre o defensor e o detrator da opinião segundo a qual a terra é plana não é suficiente para caracterizar a disputa como não-cognitiva, por que isto não se aplica ao embate entre Putnam e Nozick?

A conclusão que se impõe é que Kalderon mostra, na melhor das hipóteses, que a discussão entre interlocutores tão radicalmente díspares como Putnam e Nozick pode *tornar-se* não-cognitiva, por não estar mais submetida à norma epistêmica que governa a aceitação e rejeição de proposições. A tese distintiva do ficcionalismo moral hermenêutico diz, porém, que a aceitação moral *também por parte de agentes morais normais nas circunstâncias atuais de confronto em condições de simetria epistêmica* é não-cognitiva.

Que o não-cognitivismo possa ser a descrição adequada, não da moralidade atual, mas da moralidade transformada por nossas atitudes ficcionalizadoras motivadas pelo reconhecimento do erro fundamental revelado por Mackie em nossas crenças morais atuais—esta é a tese do ficcionalismo revolucionário de que trata a próxima seção.

### 3 FICCIONALISMO MORAL REVOLUCIONÁRIO

Mesmo entre ficcionalistas morais é possível encontrar argumentos contra a interpretação não-assertórica dos proferimentos morais proposta por Kalderon. Joyce, por exemplo, pensa que os proferimentos morais estão sujeitos ao debate ‘que tem todas as marcas características do desacordo factual’ (Joyce 2001, p. 13). E ainda que houvesse uma distinção entre o desacordo factual e o moral no que tange à validade da norma mencionada por Kalderon, muitas outras evidências falam a favor da interpretação assertórica do discurso moral: sentenças morais são expressas no modo indicativo; podem ser transformadas em sentenças interrogativas; podem figurar em contextos de atitudes proposicionais; são consideradas verdadeiras ou falsas, corretas ou incorretas; são vistas como tendo um caráter impessoal e objetivo; aparecem em contextos logicamente complexos (por exemplo como antecedentes de condicionais); aparecem como premissas de argumentos válidos (Joyce 2001, p. 13). Não haveria assim razões fortes para

a revisão da opinião intuitivamente plausível de que o caráter das nossas convenções linguísticas atuais é tal que o estado mental convencionalmente associado ao enunciado moral é uma crença.

A diferença entre os ficcionalismos de Kalderon e Joyce pode ser comentada com base no exemplo já mencionado do proferimento moral ‘o aborto é errado’. Em oposição ao expressivismo de Gibbard, por exemplo, ambos concordam que este proferimento expressa semanticamente uma proposição. Para Kalderon, contudo, o proferimento não registra a crença na proposição. A aparente descrição do mundo ficcional onde o aborto tem a propriedade de ser errado é invocada somente como um meio para transmitir nossas atitudes conativas. O discurso ficcional imita o discurso factual, mas, ao contrário deste, não transmite crenças. No caso da ficção propriamente dita, isto se manifesta na tendência a responder negativamente à pergunta sobre se nós acreditamos realmente, por exemplo, que Bentinho tinha ciúmes de Capitu, embora estejamos dispostos a aceitar que Bentinho tinha ciúmes de Capitu. Esta tendência não existe no caso da moralidade convencional atual. A não ser que tenha permitido que uma teoria filosófica contraintuitiva determine suas respostas, quem endossa o pensamento de que o aborto é errado não nega, via de regra, que o aborto instancie realmente a propriedade de ser errado. É claro que alguém poderia também não ser sincero e proferir ainda ‘o aborto é errado,’ por exemplo para influenciar de forma mais ou menos manipulativa os destinatários do seu proferimento. Mas também o proferimento insincero é uma asserção. Isso mostra, incidentalmente, que a relação de expressão entre o enunciado e o estado mental correspondente não é uma relação concreta entre entidades no tempo e no espaço (uma delas na mente do proferidor), mas deve ser compreendida em termos de convenções linguísticas. Como o caso da mentira (do proferimento insincero) mostra, o enunciado ‘o aborto é errado’ expressa ainda a crença que o aborto é errado, mesmo que o agente manipulador não acredite que o aborto seja errado. O agente manipulador explora as convenções linguísticas em vigor, que não podem ser suspensas por um *fiat* psicológico. São elas—e não o que se passa na mente dos proferidores individuais—que determinam se o discurso moral é assertórico ou não. Considerando os fatos não controversos da análise das convenções que regem os enunciados morais, Joyce conclui, contra Kalderon, que o discurso moral é assertórico. ‘In the absence of an explanation of *why* we would have a non-assertoric discourse bearing all the hallmarks of an assertoric one ... we must assume that if something walks and talks like a bunch of assertions, it’s highly likely that it *is* a bunch of assertions’ (Joyce 2001, p. 14).

Esta interpretação de Joyce se aplica à moralidade como ela se apresenta atualmente. Joyce, como Mackie ou Moore, é um cognitivista moral. Sua proposta definitiva, contudo, é que nós estabeleçamos uma *nova* prática de compreensão da moralidade. Isso envolve a adoção coletiva e mais ou menos institucionalizada de uma atitude ficcional frente à moralidade. O não-cognitivismo moral concebido por Joyce não descreve a moralidade como ela é, mas poderia vir a ser uma descrição adequada da moralidade transformada por nossas atitudes ficcionais. ‘... the proposal is not that noncognitivism is true of our actual moral discourse. Rather, fictionalism amounts to the claim that noncognitivism might *become* true, if we were to alter our attitude towards moral discourse’ (Joyce 2001, p. 201). A transformação é motivada pelo reconhecimento de que a moralidade atual cognitiva pressupõe um ponto de vista essencialmente errado, que a moralidade atual está comprometida com um universo de entidades inexistentes. O aborto, por exemplo, não instância a propriedade categórica de ser errado, pois nada pode instanciar tal propriedade. Isto quer dizer que ‘o aborto é errado’ é essencialmente falso. Mas também ‘o aborto é certo’ é falso nesse sentido. O conjunto de todas as nossas asserções morais (o discurso moral) é falso. Esse é o resultado da teoria do erro proposta por Mackie e endossada por Joyce. Mas isso implica que, uma vez convencidos por Mackie, nós não podemos mais manter o discurso moral assertórico, sob pena de irracionalidade. O que fazer com o discurso da moralidade? ‘For all I know, “Jettison the discourse” is the correct answer, and it is not my intention to convince the reader otherwise’ (Joyce 2001, p. 180).

Contudo, Joyce acredita que um aspecto importante e valioso da moralidade, associado nas condições atuais ao juízo moral, pode ser preservado se nós mudarmos nossas atitudes frente ao discurso moral, não simplesmente descartando, mas ‘ficcionalizando’ a moralidade. Isto é, se nós continuarmos a expressar-nos *como se* o aborto instanciasse a propriedade de ser errado, *sabendo* que o aborto não é errado—mais ou menos como aceitamos que Bentinho tinha ciúmes de Capitu, mesmo sem acreditar realmente na existência de Bentinho, Capitu e qualquer relação entre eles. É importante notar que, sob as condições da ficcionalização da moralidade, nós não estaríamos mentindo ou proferindo insinceramente sentenças morais. A mentira, como já vimos, é ainda uma forma de asserção que pressupõe as convenções linguísticas que conferem caráter assertórico aos proferimentos morais. O que Joyce imagina é a implementação de convenções linguísticas alternativas, convenções similares àquelas que se aplicam hoje ao discurso da ficção. Contra o pano de fundo destas convenções, o assentimento que damos à sentença ‘Bentinho tinha

ciúmes de Capitu’ não sinaliza/expressa nossa crença na verdade da proposição correspondente. Mas tampouco estamos mentindo ou sendo insinceros quando proferimos ou aceitamos ‘Bentinho tinha ciúmes de Capitu’ contra o pano de fundo das convenções próprias à ficção. O ponto importante aqui é que a ficcionalização da moralidade proposta por Joyce equivale à instituição de uma nova prática social, ao estabelecimento de novas convenções linguísticas governando o uso de sentenças morais. Da mesma forma que a ausência da crença na mente do proferidor individual insincero não impede que seu proferimento expresse uma crença, o ‘ficcionalizador solitário,’ por mais convicto que ele esteja da verdade da teoria do erro de Mackie, não pode transformar o discurso moral em um discurso não-assertórico (cf. Joyce 2001, p. 203). Somente uma nova prática coletiva poderia ficcionalizar a moral. Também por isso o ponto de vista de Joyce é melhor caracterizado como um *ficcionalismo revolucionário*, em oposição ao ficcionalismo hermenêutico.

Sob condições da ficcionalização efetiva da moralidade, nós daríamos ainda assentimento à sentença ‘o aborto é errado,’ por exemplo. Porém, logo que nos retirássemos do ‘contexto prático (ou deliberativo),’ no qual estamos imersos quando discutimos ‘questões de primeira ordem’ sobre o que devemos fazer ou sobre o que é certo ou errado, e déssemos entrada no ‘contexto crítico’ (o contexto mais reflexivo da conversação metaética sobre ‘questões de segunda ordem’), nós tenderíamos a responder negativamente à pergunta sobre se nós acreditamos *realmente* que o aborto é errado. Joyce pensa que a substituição da crença moral (*belief*) atual pelo assentimento ficcional (*make-belief*) (i) deixaria intocados os aspectos emotivos e conativos do juízo moral e (ii) preservaria as vantagens práticas da crença moral atual. A moralidade atual é especialmente eficaz na resolução de conflitos de interesses que emergem em situações de escassez moderada de recursos. E isso é assim porque o juízo moral sincero geralmente motiva à ação benéfica, silenciando o cálculo auto-interessado das conseqüências. Tudo isso (o valor instrumental social da moralidade) deveria permanecer também sob condições da enunciação não-assertórica de sentenças morais e do correspondente abandono da crença moral em favor do assentimento meramente ficcional.

Nesse ponto, há uma tensão mais ou menos óbvia no ficcionalismo revolucionário de Joyce. É bastante plausível, pelo menos de um ponto de vista intuitivo, que os acompanhamentos conativos e emotivos do juízo moral—e, por extensão, as vantagens sociais da moralidade—dependam do fato de que nosso assentimento moral tenha o caráter de uma crença na natureza

impessoal, objetiva e categórica do conteúdo ao qual damos nosso assentimento. Se a superestrutura conativa/emotiva da moralidade não sobreviver à transformação infraestrutural de sua natureza cognitiva em um universo ficcional, o programa ficcionalista de Joyce terá fracassado (cf. Sainsbury 2009 sobre ‘o valor da moralidade de um ponto de vista ficcionalista.’). Por que a moralidade ficcionalizada teria ainda o valor instrumental, o potencial para a resolução de conflitos de interesses e todas os outros benefícios próprios da moralidade à qual damos o nosso assentimento categórico?

A solução engenhosa proposta por Joyce para esse problema depende da caracterização mais detalhada de nossas atitudes no contexto prático e no contexto crítico. Para Joyce, o assentimento dado por nós a um conteúdo moral qualquer no contexto prático partilharia com as crenças morais atuais a propriedade de não ser questionado. Assim como a função social associada à moralidade (por exemplo, possibilitar a cooperação entre seres humanos) é bem servida quando a moralidade é vista internamente pelos agentes morais como a representação verdadeira do universo das razões externas objetivas e impessoais (mesmo que a visão interna seja um erro), o não-questionamento de nossos assentimentos (nossas *pseudo*-crenças) no contexto prático ou deliberativo garantiria a realização satisfatória da função prática da moralidade. É somente no contexto crítico que temos consciência do caráter ficcional de nossas *pseudo*-crenças morais, que não são (não podem ser) realmente verdadeiras. Na medida em que os contextos podem ser separados, não há maiores problemas. Mas se o contexto de deliberação incorporar também uma dimensão crítica reflexiva—e é razoável supor que as coisas sejam assim—nós nos encontraríamos na mesma situação desconfortável em que se encontrava o teórico do erro que ainda não deu o passo rumo ao ficcionalismo: a situação de agentes morais que precisam agir de acordo com crenças que a nossa reflexão revela como essencialmente falsas (cf. Wallace 2003).

A natureza da atitude ficcional e o ficcionalismo tal como ele foi exemplarmente aplicado a diferentes regiões do discurso têm sido amplamente discutidos nos últimos anos. Existe hoje uma extensa literatura sobre essa abordagem inovadora de problemas metafísicos na matemática, na ciência, no discurso sobre mundos possíveis, no discurso sobre atitudes proposicionais e, *last but not least*, na moralidade. As considerações na presente seção e na seção anterior sinalizam alguns dos problemas estruturais que a ficcionalização da moral deve ainda resolver. Em primeiro lugar, há um *deficit* argumentativo no ficcionalismo hermenêutico no que

diz respeito à sua orientação não-cognitivista. O argumento da intransigência desenvolvido por Kalderon como uma alternativa ao argumento internalista comum não estabelece uma distinção conceitual de princípio entre as normas que regem o discurso factual e as que regem o discurso moral. Talvez a natureza não-cognitiva da moralidade possa ser fundamentada de outra maneira. A questão ainda está aberta. Quanto ao ficcionalismo revolucionário, é preciso reconstruir o caminho que leva do cognitivismo moral irrealista de Mackie à moral ficcionalizada defendida por Joyce. Em particular, é especialmente relevante a apreciação crítica da tese segundo a qual uma moralidade fictícia poderia ainda ter uma influência adequada na conduta cooperativa e na resolução de conflitos de interesses.

#### 4 INTERNALISMO E MOTIVAÇÃO MORAL

Para muitos filósofos existe uma conexão necessária (interna) entre o juízo moral (a aceitação de uma sentença moral) e a motivação para a ação. De acordo com uma versão desse princípio, quem pensa que  $\phi$  é certo está necessariamente motivado a realizar  $\phi$ . A motivação é uma disposição para agir que pode não se manifestar em uma circunstância ou outra, mas ainda assim ela está presente se o agente mantém o juízo moral correspondente. A motivação pode ser superada por uma motivação contrária. Uma versão mais fraca deste princípio condiciona a conexão necessária à racionalidade do agente: o agente que endossa a sentença ‘ $\phi$  é certo’ está necessariamente disposto a realizar  $\phi$ , sob pena de irracionalidade prática. Essa versão abre espaço para a ação acrática, que é normalmente considerada como manifestação irracional. A tese da conexão necessária (condicionada ou não à racionalidade do agente) entre o juízo moral e a motivação adequada é normalmente conhecida como ‘internalismo de juízo,’ de acordo com uma sugestão terminológica de Darwall (1983). Dela se distingue a afirmação de que há uma conexão interna ou necessária entre as *razões normativas* de um agente e os elementos do seu ‘conjunto motivacional’ atual (ou simplesmente dos seus ‘desejos’ atuais), onde as razões normativas são as considerações que, independentemente da percepção por parte do agente, isto é, independentemente das *razões motivacionais* subjetivas do agente, falam objetivamente a favor da ação realizadora de  $\phi$ . Este é o ‘internalismo de existência’ na terminologia de Darwall. Aqui serão utilizadas as expressões ‘internalismo motivacional’ e ‘internalismo de razões.’ As

implicações metaéticas do primeiro tipo de internalismo serão tratadas na presente seção. Considerações sobre o internalismo de razões formam o objeto da próxima seção.

Um argumento muito comum pró não-cognitvismo mobiliza o internalismo motivacional como sua premissa principal:

1. Juízos morais motivam necessariamente. (internalismo motivacional)
2. Estados mentais têm direção de ajuste mente-mundo ou mundo-mente.
3. Somente estados mentais com direção de ajuste mundo-mente são essencialmente motivadores. (teoria humiana da motivação)
4. Logo, juízos morais têm direção de ajuste mundo-mente. (não-cognitvismo moral)

A segunda premissa desse argumento não é controversa. Ela só foi negada pelos defensores de estados mentais híbridos (*'besires'*) como estados com dupla direção de ajuste. Mas isso é visto hoje como uma hipótese extravagante. A defesa mais plausível da terceira premissa foi formulada por Smith, que (i) situa a motivação para agir em uma estrutura teleológica—‘estar motivado é basicamente ter uma meta’—e (ii) identifica o estabelecimento de uma meta com um estado de direção de ajuste mundo-mente—‘ter uma meta é estar em um estado com o qual o mundo deve ajustar-se’ (Smith 1994, p. 116). Embora tenha sido alvo de considerações críticas por alguns filósofos, a terceira premissa será considerada aqui como estabelecida. Isto permite isolar a primeira premissa como o ponto crucial do argumento reproduzido acima. Quão forte é o argumento que leva do internalismo motivacional ao não-cognitvismo?

Antes de abordar esta questão, é conveniente considerar outro argumento que combina o internalismo motivacional com o cognitvismo moral para concluir com a afirmação de que os fatos morais são inexistentes ou pelo menos ‘bizarros’ (*queer*):

1. Juízos morais motivam necessariamente. (internalismo motivacional)
2. Juízos morais são estados mentais de direção de ajuste mente-mundo, isto é, crenças. (cognitvismo moral).
3. Logo, os fatos morais (os objetos de nossas crenças morais) são intrinsecamente motivacionais, no sentido de que a representação de tais fatos (as crenças correspondentes) são necessariamente motivadoras.
4. Os fatos não-morais com que estamos familiarizados não são intrinsecamente motivacionais.



5. Logo, fatos morais são bizarros, muito provavelmente inexistentes.

A quarta premissa desse argumento é uma extensão da teoria humiana da motivação. Uma crença comum (não-moral)—a representação subjetiva dos fatos mundanos—faz parte dos antecedentes causais motivadores das nossas ações. Intrinsecamente, porém, a crença comum é inerte: somente em combinação com um estado ativo independente e de direção de ajuste mente-mundo (desejo) ela contribui para a motivação. Como na exposição do primeiro argumento esquematizado no início desta seção, a quarta premissa pode ser considerada aqui como estabelecida. A segunda premissa (o cognitivismo moral), por outro lado, pode ser vista como uma assunção. Nesse caso, o argumento leva do internalismo motivacional ao irrealismo moral. Na opinião de muitos filósofos, essa é a estrutura do *argument from queerness*, no qual Mackie (1977, pp. 38-42) baseia a teoria do erro. Alternativamente, a primeira parte do argumento pode ser vista como a base de uma *reductio*: se, depois da conclusão da primeira parte, acrescentarmos a premissa plausível de que não existem crenças essencialmente motivadoras, a falsidade do cognitivismo moral poderá ser inferida com alguma segurança.

O que foi dito até aqui justifica a relevância atribuída ao internalismo motivacional no debate metaético. Os não-cognitivistas vêem comumente no internalismo motivacional um aliado poderoso, ao passo que os cognitivistas o consideram um adversário. Não por acaso, os cognitivistas são, via de regra, também externalistas motivacionais que (i) não negam a existência de uma conexão regular e relativamente estreita entre o juízo moral e a motivação correspondente, mas (ii) criticam os internalistas por terem aparentemente confundido uma contingência recorrente com uma relação necessária. Para o cognitivista externalista típico é perfeitamente concebível e coerente a descrição de um ‘amoralista,’ isto é, de um agente racional no qual a perda da motivação correspondente deixa ainda intacta sua crença moral. Todavia, alguns cognitivistas não gostariam de adotar o externalismo motivacional. Estes vêem na compatibilização do internalismo motivacional com o cognitivismo—e a consequente refutação da hipótese amoralista—a solução do problema mais central da metaética, o que Smith (1994, p. 4-13) chama de ‘problema moral’ (cf. tb. McNaughton 1988, p. 23).

Como Jon Tresan observou em alguns ensaios recentes (Tresan 2006, 2009a e 2009b), o debate sobre o internalismo motivacional repousa em uma ambiguidade. A idéia de uma conexão necessária entre juízos morais e motivações pode ser interpretada em dois sentidos excludentes. De acordo com a interpretação *de dicto*, a necessidade é atribuída a uma proposição:

necessariamente, juízos morais são motivadores—ou talvez: necessariamente, juízos morais são acompanhados de motivações. (Em termos mais formais, o internalismo motivacional *de dicto* pode ser simbolizado como um condicional material universalmente quantificado e precedido do operador proposicional de necessidade.) De acordo com a interpretação *de re*, a necessidade se aplica aos juízos morais e não a proposições sobre os juízos morais: juízos morais são necessariamente motivadores—ou talvez: juízos morais são necessariamente acompanhados de motivações. (Em termos mais formais, o internalismo motivacional *de re* pode ser simbolizado como um condicional material universalmente quantificado onde o conseqüente do condicional é precedido do operador de necessidade.) A distinção pode ser esclarecida com base em um exemplo trivial. *De dicto*, é necessariamente verdadeiro que pais têm filhos. Mas não é verdadeiro que pais, *de re*, necessariamente têm filhos. Os pais no mundo atual podem existir em outros mundos possíveis sem filhos. Eles não seriam pais nesses mundos, mas ainda assim seriam os mesmos seres que no mundo atual são pais e existem, em alguns mundos possíveis, sem filhos. Isso é o mesmo que dizer que os pais atuais não são essencialmente pais. Em todos os mundos possíveis, porém, os pais nesses mundos têm filhos. No que diz respeito à relação entre pais e filhos, nós devemos ser internalistas *de dicto*, mas externalistas *de re*. Qual deve ser a nossa posição quando o interesse é dirigido não à relação pais e filhos, mas à relação entre juízo moral e a motivação correspondente?

Aqui é importante observar que a validade dos argumentos esquematizados acima depende da interpretação *de re* da conexão necessária entre juízo moral e motivação. As premissas relevantes afirmam que a crença moral (a representação dos fatos morais) não pode deixar de ser motivadora. Contudo, se a necessidade em questão fosse *de dicto*, a crença moral *poderia* não motivar, como qualquer pai atual poderia não ter filhos. Em um mundo possível onde um pai atual não tem filhos, ele não seria pai, mas ainda a mesma pessoa com todas as suas características essenciais. Da mesma forma a representação dos fatos morais sem a propriedade de motivar ou desacompanhada de motivação não seria uma crença *moral*, mas ainda assim a mesma crença com todas as suas propriedades essenciais. O que subjaz a este raciocínio é a idéia independentemente plausível de que a propriedade de motivar para a ação é uma propriedade não-essencial das crenças morais, assim como ser pai não é uma propriedade essencial dos que são pais. O resultado é que somente a interpretação *de re* da conexão necessária entre juízos

morais, por um lado, e estados de direção de ajuste mundo-mente, por outro, tornaria válidos os argumentos acima. Mas talvez devamos ser internalistas motivacionais *de re*.

A ambigüidade notada por Tresan se manifesta na discussão sobre o internalismo motivacional—não na formulação das consequências do internalismo motivacional. Nesse contexto a interpretação relevante é a *de dicto*. É a verdade do internalismo motivacional *de dicto* que está em jogo na discussão metaética. Argumentos pró e contra o internalismo motivacional giram em torno da figura do amoralista e apelam às nossas intuições. Defensores do externalismo afirmam ser natural e intuitiva a interpretação dos proferimentos do amoralista como genuinamente morais, apesar da ausência de toda e qualquer motivação. Os internalistas, ao contrário, consideram que, na pior das hipóteses, os proferimentos do amoralista não são atos de fala significativos, não expressam mais juízos—o amoralista deixou de *falar*. Na melhor das hipóteses, os internalistas consideram que os proferimentos do amoralista devem ser interpretados como expressando juízos não-morais, por exemplo, sobre o que seus ouvintes (mas não ele próprio) julgam ser certo ou errado. O sentido do proferimento do amoralista ‘ $\phi$  é errado, mas eu não estou motivado a fazer  $\phi$ ’ é o da sentença ‘ $\phi$  é o que vocês chamam de errado, mas eu não estou motivador a fazer  $\phi$ .’ De acordo com esta interpretação do proferimento do amoralista, este não é verdadeiramente um amoralista, mas um não-conformista moral cuja motivação está perfeitamente alinhada com o juízo moral negativo sobre as convenções locais. Evidentemente, os internalistas motivacionais pensam que essa é a interpretação que está mais de acordo com nossas intuições. Qualquer que seja a decisão sobre esse conflito de interpretações, o que está sendo discutido aqui são as condições de aplicabilidade do conceito de juízo moral. O internalista motivacional e o externalista não estão discutindo, com referência ao amoralista, se um juízo moral no mundo atual é acompanhado de motivação em qualquer outro mundo possível. O que o internalista motivacional e o externalista discutem é se o juízo proferido pelo amoralista, *no mundo possível do amoralista*, é ou não um juízo *moral*. O internalista insiste—e o externalista nega—que a presença da motivação adequada seja uma condição para a aplicabilidade do conceito de juízo moral ao proferimento do amoralista. Nesse sentido, a tese internalista estabelece uma conexão necessária e *a priori* entre o conceito de juízo moral, por um lado, e o de motivação, por outro, não uma relação *de re* entre juízos morais, *independentemente da descrição como morais*, e motivações. Como Tresan define o internalismo motivacional *de dicto*: ‘moral beliefs are accompanied by conations, and are conceived by us precisely as so accompanied

(perhaps inter alia)’ (Tresan 2006, p. 147, grifo nosso). De forma semelhante, Jackson afirma a compatibilidade entre o internalismo motivacional e o cognitivismo moral: ‘... a moral judgement is accompanied by the relevant pro-attitude ... And we can accommodate this view by refusing to *call* something a *moral* belief unless it is accompanied by the relevant pro-attitude (Jackson 1998, p. 161).

Considerando, agora, os dois sentidos possíveis para a tese internalista, o resultado é o seguinte: o não-cognitivismo e o irrealismo moral só podem ser justificados com base na premissa internalista se esta for compreendida *de re*. Os dois argumentos esquematizados nesta seção são simplesmente inválidos quando a premissa internalista é interpretada *de dicto*. Isto remete o problema para a justificação da premissa internalista. A verdade do não-cognitivismo e do irrealismo moral depende da validade dos argumentos reproduzidos acima e da verdade da premissa internalista. Porém, se alguma forma de internalismo puder ser justificada no debate entre internalistas e externalistas, esta será a do internalismo *de dicto*. A verdade do internalismo motivacional *de re* não é o objeto da discussão centrada no caso do amoralista. Pode ser que haja uma justificação para o internalismo *de re*, mas ela parece não ter sido ainda proposta. O que se segue é que não temos um argumento internalista pró não-cognitivismo, que possa ser utilizado pelo ficcionalista hermenêutico. Como o argumento da permissibilidade da intransigência (comentado na seção 2) também parece ser insatisfatório, o não-cognitivismo moral—e com ele o ficcionalismo hermenêutico—ainda são problemáticos. Mas isso não quer dizer que o ficcionalismo revolucionário esteja em melhor situação: não existe ainda um argumento válido e baseado em uma premissa internalista justificada que leve à teoria do erro—e desta depende o ficcionalismo revolucionário.

## 5 RAZÕES NORMATIVAS E MOTIVAÇÃO MORAL

O segundo argumento esquematizado na seção anterior reconstrói, como pensam muitos filósofos, em termos atualizados o núcleo do *argument from queerness* proposto por Mackie, a aproximação do valor moral objetivo com a Forma Platônica do Bem que ‘é tal que o conhecimento dela provê o sujeito do conhecimento com uma orientação e um motivo preponderante; o fato de que algo é bom diz à pessoa que sabe isso que o persiga e, ao mesmo, faz com que ela o persiga’ (Mackie 1977, p. 40). Se as considerações na seção anterior estiverem

corretas, o ficcionalista moral revolucionário precisa encontrar uma alternativa ao *argument from queerness*, que não recorra mais ao internalismo motivacional. Joyce propõe esta alternativa em *The Myth of Morality* (Joyce 2001):

1. Necessariamente, se S deve<sub>M</sub> fazer  $\phi$ , então S deve fazer  $\phi$  independentemente de seus desejos ou interesses. (moralidade categórica)
2. Se S deve fazer  $\phi$ , S tem uma razão para fazer  $\phi$ .
3. Logo, se S deve<sub>M</sub> fazer  $\phi$ , então S tem uma razão para fazer  $\phi$  independentemente de seus desejos ou interesses.
4. Não existem tais razões. (internalismo de razões)
5. Logo, não há coisa alguma que S deva<sub>M</sub> fazer.

O verbo ‘dever<sub>M</sub>’ representa o dever especificamente moral categórico, que, de acordo com a primeira premissa, independe dos desejos ou interesses do agente. (É isso que o torna categórico.) A segunda premissa estabelece uma relação analítica entre obrigações (morais ou não) e razões para agir. À obrigação moral correspondem razões categóricas. A quarta premissa expressa o internalismo de razões, segundo a qual há uma conexão necessária entre as *razões normativas* de um agente e os elementos do seu ‘conjunto motivacional’ atual, seus desejos e interesses atuais. A conclusão do argumento nega a existência de qualquer obrigação moral.

Esse argumento reproduz o espírito da teoria do erro de Mackie sem envolver o comprometimento original com a tese problemática do internalismo motivacional. De fato, Mackie formula inicialmente sua tese em termos da ‘inexistência de valores objetivos,’ mas logo acrescenta que ela inclui ‘não apenas a bondade moral, que poderia ser mais naturalmente identificada com o valor moral, mas também outras coisas que poderiam ser chamadas, num sentido mais fraco, de valores ou desvalores morais—retidão e erro, dever, obrigação, o fato de uma ação ser corrupta e desprezível etc.’ (Mackie 1977, p. 15). O argumento de Joyce afeta diretamente só a noção de obrigação moral. Contudo, devido à conexão holística entre as noções morais, o ataque ao núcleo deontológico da moralidade tem efeitos destrutivos indiretos no resto da moralidade (cf. Joyce 2001, pp. 175s.).

A premissa chave do argumento de Joyce é, evidentemente, o internalismo de razões. Resumidamente, este foi justificado por Bernard Williams (1981) com base no princípio de que as razões normativas para um agente devem ser capazes de motivá-lo, ou elas não são razões genuínas para esse agente. Disso segue que as razões normativas são extensões racionais (no

sentido da racionalidade prática deliberativa) das motivações atuais do agente, isto é, extensões que têm sua origem no interior do conjunto motivacional atual (desejos e interesses atuais) desse agente. Embora as razões normativas não sejam idênticas, elas são relativas às motivações atuais do agente. Na terminologia de Williams, todas as razões normativas são ‘internas,’ isto é, centradas no agente para o qual elas são razões (*agent-centered*) e dependentes, no que diz respeito à sua própria existência, das motivações atuais contingentes do agente.

Assim, como consequência principalmente da quarta premissa, o núcleo deontológico da moralidade convencional cai por terra. Devido à conexão holística das noções morais, a moralidade convencional atual se revela um empreendimento cognitivo, porém essencialmente falho. Isso estabelece uma das bases para o ficcionalismo revolucionário.

Portanto, a questão imediatamente relevante é se o internalismo de razões *à la* Williams pode ser efetivamente mantido. Isto requer a investigação pormenorizada do que veio a ser chamado de *teoria humiana das razões*, para a qual têm sido propostas recentemente diversas formulações desde aquela original de Williams (cf. Schroeder 2007). Contudo, mesmo que o cognitivismo não-realista possa ser visto como a descrição filosoficamente adequada da moralidade atual, o ficcionalista moral revolucionário tem ainda de resolver o problema mencionado no final da terceira seção.

## 6 FICCIONALIZANDO A MORAL: UM ROTEIRO

As considerações feitas até aqui podem ser resumidas no seguinte roteiro para a ficcionalização da moral:

1. O ponto de partida pode ser a articulação satisfatória da distinção entre cognitivismo e não-cognitivismo. Se as observações na primeira seção desse trabalho estiverem corretas, existe uma relativa confusão na caracterização comum sobretudo do ponto de vista não-cognitivista. É necessário separar analiticamente a dimensão semântica da dimensão ‘psicológica’ do problema metaético. Isto exige o tratamento pormenorizado da semântica moral, por um lado, e da teoria dos atos de fala aplicada ao discurso moral, por outro.
2. A extensão natural das considerações anteriores é a avaliação crítica do programa não-cognitivista do ficcionalismo hermenêutico com o objetivo de superar o *deficit*

argumentativo das teorias baseadas no argumento da intransigência e no argumento internalista motivacional.

3. De forma similar, é preciso explorar o caminho que leva do cognitivismo moral irrealista de Mackie ao ficcionalismo moral revolucionário defendido por Joyce, com ênfase na apreciação crítica da tese segunda a qual uma moralidade fictícia poderia ainda ter influência adequada na conduta pessoal e na resolução de conflitos sociais de interesse.
4. Ato contínuo, é necessário diferenciar analiticamente a interpretação *de dicto* e a interpretação *de re* aplicadas à tese da conexão interna entre juízo moral e motivação. Isto serve à problematização do argumento internalista pró não-cognitivismo e do argumento motivacional do cognitivismo irrealista.
5. Finalmente, trata-se da investigação das propostas mais atuais de formulação da ‘teoria humiana das razões’ como formas de justificação racional do internalismo de razões. Isto deve ser aplicado ao exame da tentativa de substituição do argumento motivacional original de Mackie (*argument from queerness*) por um argumento centrado na relação entre razões normativas e razões motivacionais.



## REFERÊNCIAS

- BRINK, David (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- CURRIE, Greg (1990). *The Nature of Fiction*. Cambridge: Cambridge University Press.
- DARWALL, Stephen (1983). *Impartial Reason*. Ithaca: Cornell University Press.
- FIELD, Hartry (1980). *Science Without Numbers*. Oxford: Blackwell.
- FINLAY, Stephen (2008). ‘The Error in the Error Theory.’ *Australasian Journal of Philosophy* 86, pp. 347 – 369.
- GARNER, Richard (1993). ‘Are Convenient Fictions Harmful To Your Health?’ *Philosophy East & West* 43, pp. 87–106.
- GIBBARD, Allan (2003). *Thinking How to Live*. Cambridge Massachusetts: Harvard University Press.
- HINCKFUSS, Ian (1993). ‘Suppositions, Presuppositions, and Ontology.’ *Canadian Journal of Philosophy* 23, pp. 595–617.
- HORGAN, Terry e Timmons, Mark (2000). ‘Nondescriptivist Cognitivism: Framework for a New Metaethic.’ *Philosophical Papers* 29, pp. 121-153.
- JACKSON, Frank (1998). *From Metaphysics to Ethics. A Defence of Conceptual Analysis*. Oxford: Oxford University Press.
- JAY WALLACE, R. (2003). ‘Review of Richard Joyce’s *The Myth of Morality*.’ *Notre Dame Philosophical Reviews*. <http://ndpr.nd.edu/review.cfm?id=1301>.
- JOYCE, Richard (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.
- KALDERON, Mark Eli (2005a). *Moral Fictionalism*. Oxford: Clarendon.
- \_\_\_\_\_ (2005b). *Fictionalism in Metaphysics*. Oxford: Oxford University Press.
- LEWIS, D. K. (1978). ‘Truth in Fiction.’ *American Philosophical Quarterly* 15, pp. 37-46
- LILLEHAMMER, Hallvard (2004). ‘Review: The Myth of Morality.’ *Mind* 113, pp.760-763.
- MACKIE, J. L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin.
- MCNAUGHTON, David (1988). *Moral Vision: An Introduction to Ethics*. Oxford: Blackwell.
- NOLAN, Daniel, Restall, Greg e West, Caroline (2005). ‘Moral Fictionalism vs. The Rest.’ *Australasian Journal of Philosophy* 83, pp. 307-330.
- PUTNAM, Hilary (1981). *Reasons, Truth, and History*. Cambridge: Cambridge University Press.
- RAILTON, P. (1986). ‘Moral Realism.’ *Philosophical Review* 94, pp. 163-207.

- ROSEN, Gideon (1990). 'Modal Fictionalism.' *Mind* 99, pp. 327–354
- \_\_\_\_\_ (1993). 'A Problem for Fictionalism about Possible Worlds.' *Analysis* 53, pp. 71–81
- \_\_\_\_\_ (1995). 'Modal Fictionalism Fixed.' *Analysis* 55, pp. 67–73.
- SAINSBURY, R. M. (2009). *Fiction and Fictionalism*. London: Routledge.
- SCHROEDER, Mark (2007). *Slaves of the Passions*. Oxford: Oxford University Press.
- SMITH, Michael (1994). *The Moral Problem*. Oxford: Blackwell.
- TRESAN, Jon (2006). 'De Dicto Internalist Cognitivism.' *Noûs* 40, pp. 143-165.
- \_\_\_\_\_ (2009a). 'Metaethical Internalism: Another Neglected Distinction.' *The Journal of Ethics* 13, pp. 51-72.
- \_\_\_\_\_ (2009b). 'The Challenge of Communal Internalism.' *The Journal of Value Inquiry* 43, pp. 179-199.
- VAN FRAASSEN, Bas C. (1980). *The Scientific Image*. Oxford: Clarendon.
- WALTON, Kendall (1990). *Mimesis as Make-Believe*. Cambridge: Harvard University Press.
- WEDGWOOD, Ralph (2007). *The Nature of Normativity*. Oxford: Clarendon.
- WILLIAMS, Bernard (1981). 'Internal and External Reasons.' In Bernard Williams. *Moral Luck*. Cambridge: Cambridge University Press, pp. 101-113.
- YABLO, Stephen (2001). 'Go Figure: A Path Through Fictionalism.' *Midwest Studies in Philosophy* 25, pp. 72–102.
- ZALTA, Edward N. (1992). 'Fictional Truth, Objects, and Characters.' In J. Kim e E. Sosa (orgs.). *Blackwell Companion to Metaphysics*. Oxford: Blackwell.
- \_\_\_\_\_ (2000). 'The Road Between Pretense Theory and Object Theory.' In A. Everett e T. Hofweber (orgs.). *Empty Names, Fiction, and the Puzzles of Non-Existence*. Stanford: CSLI Publications, pp. 117–147.