

ÉTICA EVOLUCIONISTA: A DIFERENÇA ENTRE ALTRUÍSMO E MORALIDADE

Roger Valério de Vargas Rex
UnB

Resumo: Autores que adotam a abordagem da ética evolucionista para explicar a conformação da psicologia humana costumam confundir altruísmo e moralidade, ignorando que a existência de comportamentos cooperativos e de punições contra indivíduos egoístas não depende da existência de juízos morais. No presente artigo, apresento argumentos extraídos de discussões entre cognitivistas e não-cognitivistas e resultados de estudos na área da Psicologia moral a fim de esclarecer a importância de distinguir os dois conceitos mencionados. Concluo que a compreensão das diferenças entre altruísmo e moralidade conduz à identificação das diversas capacidades intelectuais envolvidas nos julgamentos morais e é essencial para o projeto de averiguar se a capacidade para julgar moralmente é uma adaptação biológica.

Palavras-chave: Ética evolucionista; altruísmo; moralidade; não-cognitivismo; raciocínio moral.

Introdução

A ética evolucionista descritiva¹ apela para a teoria da evolução a fim de explicar a origem de certas capacidades humanas, tendências, comportamentos, padrões de pensamento e comportamento. Autores que adotam essa abordagem costumam, por exemplo, examinar a hipótese de que a capacidade humana de julgar moralmente é produto do processo de seleção natural (FITZPATRICK, 2016; JOYCE, 2006, p. 143).

Abordagens evolutivas da moralidade em geral se dedicam a explicar o altruísmo entre os seres humanos. Alguns pesquisadores acreditam que há alguma forma de moralidade inclusive em outros animais e fundamentam essa crença com base em observações de comportamentos altruístas entre outras espécies. Warneken e Tomasello (2006), *e.g.*, demonstraram que chimpanzés são prestativos, quando se trata de alcançar objetos caídos ‘acidentalmente’ aos pesquisadores. De Waal (1996) demonstrou que eles compartilham comida, consolam uns aos outros e retribuem os cuidados recebidos. No entanto, embora altruísmo e moralidade estejam relacionados, é importante não confundí-los. O fato de que certos animais praticam condutas que

¹Para a distinção entre ética evolucionista descritiva, ética evolucionista prescritiva e metaética evolucionista, ver Fitz Patrick (2016).

consideramos louváveis não nos permite inferir que eles tenham qualquer tipo de consciência moral. Eles podem cooperar por diversas razões, tais como empatia ou medo de represálias. Somente poderíamos dizer que eles agem moralmente, se eles fossem capazes de elaborar um juízo sobre qual seria a conduta correta. Caso contrário, também deveríamos dizer que as abelhas agem moralmente. Uma abelha é capaz de sacrificar a sua vida para salvar a colmeia, mas ela não faz isso movida por qualquer tipo de convicção moral de que está agindo pelo bem comum: ninguém diria que uma abelha ataca quem ameaça a sua colmeia em razão de julgar que tal atitude está moralmente errada e, portanto, merece punição. Equiparar altruísmo e moralidade implica ignorar a distinção entre praticar um ato bom e praticar um ato porque ele é bom.

Responder se a moralidade é uma adaptação não envolve apenas investigar a função adaptativa de comportamentos altruístas, envolve também investigar qual a função adaptativa de valorar condutas. Ayala, por exemplo, não considera que fazer julgamentos morais seja algo que aumente a aptidão (2010, p. 5). Teorias sobre a evolução da moralidade costumam enfatizar como as normas podem servir para estabilizar condutas cooperativas dentro do grupo. No entanto, não é algo óbvio que a existência de algum tipo de punição contra indivíduos egoístas dependa da existência de normas. Os indivíduos de um grupo poderiam ‘punir’ aqueles que apresentassem condutas egoístas simplesmente em razão de sentirem uma aversão a esse tipo de conduta, eles não precisam acreditar que condutas egoístas são erradas para atacar indivíduos egoístas.

Igualar esses conceitos implica uma ampliação injustificada do conceito de moralidade que ignora o seu caráter normativo. Sensibilidade a situações que causam danos a terceiros, empatia pelos membros do grupo, capacidade para generalização, capacidade para memorizar condutas não-cooperativas, todas essas características podem possibilitar o bom convívio no mundo social sem implicar a existência de moralidade. Podemos imaginar um ser que colabore com os membros de sua espécie, esteja disposto a sacrificar-se conscientemente em benefício de seu grupo, mas seja incapaz de julgar que alguma ação merece punição, ou que ela é desejável e não simplesmente desejada, ou que a autoridade de uma norma independe de convenção (JOYCE, 2013, p. 547).

Nesse sentido, a distinção feita por Darwin entre ‘social instincts’ e ‘intellectual powers’ é ainda hoje relevante, na medida em que explicita o fato de que é concebível a existência de grupos de cooperadores/altruístas que não possuam faculdade moral (JOYCE, 2006). Para que os animais se comportem de uma forma ou de outra não é necessário que eles julguem determinada conduta como boa ou ruim: para que um comportamento seja selecionado basta que os indivíduos o adotem, eles não precisam fazê-lo conscientemente. Isto é, poderia haver condutas que hoje consideramos moralmente louváveis sem que houvesse qualquer ser capaz de fazer julgamentos morais. Poderíamos

agir altruisticamente por alguma inclinação sem que acreditássemos dever agir dessa forma. Fazer algo porque gostamos de fazer é diferente de fazer algo por acreditar que temos um dever (JOYCE, 2006, p. 50).

Obviamente o fato de moralidade e altruísmo não se confundirem não implica que o surgimento do altruísmo psicológico não tenha sido um precursor essencial do julgamento moral (JOYCE, 2013, p. 537). Emoções que favorecem o comportamento social evoluíram antes da cognição moral e tornaram possível a cooperação e o aprendizado cultural, preparando o caminho para que o pensamento normativo fosse possível (STERELNY, 2012, p. 140).

Para conseguirmos distinguir claramente esses dois conceitos é indispensável reconhecermos que os julgamentos morais envolvem crenças, em algum sentido relevante. Ao reconhecermos esse ponto, estaremos em uma posição melhor para compreender as diversas capacidades intelectuais envolvidas nos julgamentos morais, e essa compreensão é essencial para o projeto de averiguar se a moralidade é uma adaptação biológica. Uma vez constatado que a moralidade envolve crenças, qualquer explicação adaptacionista precisará fornecer uma explicação a respeito de como as capacidades cognitivas necessárias para a formação dessas crenças evoluíram, isto é, uma simples explicação acerca da origem de nosso comportamento cooperativo não será suficiente para defender a tese de que a moralidade é uma adaptação.

Contudo, essa ideia de que julgamentos morais envolvem crenças enfrenta objeções. A discussão sobre essa questão é um dos pontos centrais do debate metaético entre cognitivistas e não-cognitivistas. Essa discussão envolve tanto aspectos conceituais quanto questões empíricas acerca da psicologia humana. No entanto, os argumentos apresentados pelos filósofos da tradição analítica, como regra, enfatizam os aspectos conceituais, e, portanto, não são suficientes para esclarecer esse debate.

Ainda assim, os argumentos empregados pelos cognitivistas são relevantes na medida em que mostram que, uma vez aceitos os pressupostos não-cognitivistas, o raciocínio moral não faria qualquer sentido. Se dizer que algo é errado equivale apenas a expressar uma atitude de desaprovação, teríamos dificuldade para explicar a atitude de quem se questiona se as suas próprias atitudes estão moralmente corretas.

Os estudos psicológicos mais recentes acerca dos papéis desempenhados pela nossa capacidade de raciocínio nos julgamentos morais ajudam a preencher os espaços deixados pelos argumentos conceituais. Esses estudos indicam que, embora os nossos julgamentos sejam em larga medida influenciados por nossas emoções, eles envolvem também processos de raciocínio consciente. A análise desses estudos nos ajudará a compreender as capacidades intelectuais envolvidas nos julgamentos morais.

1. Cognitivismo e não-cognitivismo

O não-cognitivismo apoia-se em duas teses negativas centrais: *não-fatualismo semântico* e *não-cognitivismo psicológico*. A primeira tese nega que sentenças nas quais se atribui predicados morais expressem proposições que possam ser verdadeiras ou falsas em algum sentido substancial, ou seja, elas não possuem condições de verdade. A segunda tese nega que os estados mentais manifestados por meio de discursos morais sejam crenças. Esses estados mentais não seriam ‘cognitivos’, pelos menos não no mesmo sentido que as crenças são. Eles serviriam para expressar atitudes ‘não-cognitivas’ similares a desejos e atitudes de aprovação ou desaprovação (VAN ROOJEN, 2015; JOYCE, 2015). No entanto, as diferentes teorias não-cognitivistas – emotivismo, prescritivismo, expressivismo, etc. – divergem em relação ao tipo exato de atitude manifestada pelos julgamentos morais. Conforme, A.J. Ayer, por exemplo, quando dizemos ‘roubar é errado’, é como se estivéssemos apenas dizendo ‘roubando dinheiro!!!’ com um tom de voz indicando um sentimento de desaprovação (AYER, 2002, p. 107). Já para Carnap, dizer que matar é errado seria o mesmo que dizer ‘não mate!’ (CARNAP, 1935, p.23-24 e 29)².

Cognitivistas, por sua vez, negam o não-cognitivismo. Eles afirmam que as sentenças contendo julgamentos morais estão aptas a serem verdadeiras ou falsas e que o estado mental de quem emite ou aceita um julgamento moral equivale, como regra, a uma crença. Portanto, sentenças com predicados morais seriam semelhantes a outras sentenças com predicados descritivos (VAN ROOJEN, 2015; MAMELI, 2013, p. 903-4).

Alguns filósofos contemporâneos que estudam a moralidade a partir de uma perspectiva evolutiva adotam alguma forma de não-cognitivismo. Mameli, *e.g.*, defende uma teoria expressivista sobre a moralidade, segundo ele a habilidade de manifestar estados mentais a respeito do que é moralmente exigível é uma habilidade de adquirir certos tipos de disposições emocionais direcionadas aos violadores de certos tipos de demandas práticas. Ou seja, julgamentos morais, segundo Mameli, são constituídos apenas por disposições para sentir certos tipos de emoção. De acordo com essa visão, se você não está disposto a sentir raiva em relação a um assassino, então você não considera genuinamente que assassinatos são moralmente proibidos (MAMELI, 2013, p. 903-7)³. Ele resume assim a sua posição expressivista sobre a moralidade:

² Essas duas teses enfatizam aspectos distintos dos julgamentos morais. A primeira delas envolve concebê-los como ‘entidades linguísticas’; enquanto a segunda tese envolve concebê-los como estados mentais (JOYCE, 2015). No presente artigo, abordo os julgamentos morais, enquanto estados mentais manifestados por indivíduos. Desta forma, a minha preocupação principal gira em torno da segunda tese (não-cognitivismo psicológico).

³ Mameli reconhece que os estados mentais acerca do que é exigido pela moralidade possuem uma estrutura proposicional, pois eles incluem a representação mental de uma ação e a atribuição de um predicado (a ideia de que essa ação é moralmente proibida ou obrigatória, *e.g.*). No entanto, segundo

Let *a* be a specific action and *p-for-a* the corresponding practical demand in its favor. Consider the following dispositions: 1. A disposition D1 to feel anger when you (the person with the disposition) believe or imagine that someone else acts (or did act, or will act) in *p-for-a*-violating ways, and to feel anger about such behavior. 2. A disposition D2 to feel guilt when you believe or imagine that you act (or did act, or will act) in *p-for-a*-violating ways, and to feel guilt about such behavior. 3. A disposition D3 to feel anger when you believe or imagine that someone else does not (or did not, or will not) have or manifest dispositions D1 and D2, and to feel anger about this. 4. A disposition D4 to feel guilt when you believe or imagine that you do not (or did not, or will not) have or manifest dispositions D1 and D2, and to feel guilt about this. Our hypothesis is that, for any action *a*, for an individual to token the moral judgement that *a* is morally required just is for that individual to acquire the dispositions D1–D4 with respect to *a* (MAMELI, 2013, p. 905-907).

Prinz adota posição similar à adotada por Mameli. Ele entende que emoções além de *necessárias* são também *suficientes* para avaliarmos algo como bom ou mau; ou seja, valores morais seriam sentimentos e nada mais (PRINZ, 2013). Pensar que roubar é errado equivaleria a ter sentimentos negativos frente a esta conduta (PRINZ, 2013, p. 103)⁴. Prinz resume a sua compreensão sobre a moral da seguinte forma:

Basic moral values are moral sentiments directed towards various acts. To believe that stealing is bad is to have a negative moral sentiment towards stealing. To have a negative moral sentiment is to be disposed to feel one of the other-directed moral emotions when someone else commits the act, and a self-directed moral emotion when you commit the act (PRINZ, 2012, p. 267).⁵

A adoção de posições desse tipo abre caminho para a confusão entre moralidade e altruísmo. Autores não-cognitivistas que adotam posições

Mameli isso não implica que esses estados mentais sejam crenças. Ele menciona exemplos de desejos que, embora não sejam crenças, também possuem essa estrutura proposicional, tais como o pai que deseja que seu filho tire boas notas. Proposições desse tipo não seriam crenças, pois elas seriam direcionadas da mente para o mundo, enquanto crenças autênticas são formadas de modo inverso, isto é, do mundo para a mente (MAMELI, 2013, p. 903-4).

⁴Prinz tenta explicar o surgimento do julgamento moral a partir da emoção de culpa, mas essa é uma emoção complexa, rica em elementos cognitivos. Ver Joyce (2006, p. 101–4).

⁵ A principal diferença entre esses dois filósofos nesse ponto é que Prinz não menciona nada equivalente às disposições D-3 ('meta-raiva') e D-4 ('metaculpa'), elencadas por Mameli. Para Prinz o julgamento moral funcionaria da seguinte forma: 'An agent A's doing/having/being X is bad iff by X, A (potentially or actually) has an effect on victim V, where V is construed as a person, a community, or nature, and, depending on that construal, an evaluator E who so construes A's X-ing will have the corresponding emotion of blame (anger, contempt, or disgust, if E is a third party, and guilt or shame if E = A)' (PRINZ, 2013, p. 104-5).

semelhantes às de Mameli e Prinz enfatizam aspectos emocionais e deixam de lado muitas das capacidades cognitivas envolvidas nos julgamentos morais que estão presentes apenas na espécie humana. Em consequência disso, eles têm dificuldade em explicar por que parece tão plausível compreender os julgamentos morais como um modo de expressar crenças e de atribuir propriedades a certas condutas. Eles também têm dificuldade em explicar por que o pensamento sobre questões morais parece se submeter às regras da lógica, ao contrário do que ocorre com as atitudes às quais eles buscam equipará-lo – manifestações de emoções, prescrições, etc. (SAYRE-MCCORD, 2014)⁶.

O não-cognitivista procura explicar os significados de expressões morais quando utilizadas como predicados em sentenças simples (*freestanding predicative uses*) e os estados mentais que manifestamos por meio dessas expressões. No entanto, ele tem dificuldade para explicar a utilização de expressões morais em outras sentenças, nas quais os predicados aparecem embutidos em estruturas mais complexas (*embedded sentences*). Dito de outro modo, ele se propõe a explicar os usos de predicados morais quando esses predicados são afirmados diretamente, mas não oferece explicações satisfatórias para os casos em que esses predicados aparecem de outras formas.

Nesse sentido, ainda que a explicação não-cognitivista de que o significado de ‘mentir é errado’ é simplesmente uma expressão de aversão à mentira possa parecer plausível; ela não é uma boa explicação para a atitude manifestada por essa mesma expressão quando ela está embutida em outras sentenças, tais como ‘é verdade que mentir é errado’; ‘eu me pergunto se mentir é errado’; ‘Paulo acredita que mentir é errado’, etc. É completamente implausível afirmar que quem emprega a expressão ‘mentir é errado’ em sentenças desse tipo está tentando expressar uma atitude de aversão. Quem se pergunta se mentir é errado não está concordando que mentir é errado, tampouco está manifestando uma emoção negativa em relação ao ato de mentir ou exprimindo um comando como ‘não minta!’ (VAN ROOJEN, 2015).

O não-cognitivista ainda poderia sustentar que os predicados morais possuem um significado quando são afirmados diretamente e outro quando estão embutidos em outra sentença, afinal as mesmas palavras podem ter usos distintos em ocasiões distintas. No entanto, para ser consistente, o não-

⁶ Há teóricos, como os defensores do quasi-realismo, que encaram esse desafio. O quasi-realismo tenta explicar, a partir de bases não-cognitivistas, como e por que os nossos discursos morais aparentam ser cognitivistas e realistas. Quasi-realistas começam afirmando que o mundo não possui propriedades morais nem pessoas que acreditam nessas propriedades e então tentam explicar como surge a prática de pensar e falar sobre propriedades morais, crenças morais e fatos morais (SAYRE-MCCORD, 2014). Eles tentam explicar por que podemos agir como se os julgamentos morais fossem aptos a serem verdadeiros, ainda que estritamente falando eles não possam ser nem verdadeiros nem falsos em qualquer sentido robusto. De acordo com eles, há uma justificativa para a nossa prática de tratá-los como se eles fossem verdadeiros (GIBBARD, 2003; VAN ROOJEN, 2015).

cognitivista precisará negar a possibilidade de argumentos logicamente válidos serem elaborados a partir de predicados morais. Para que um argumento seja logicamente válido é necessário que as expressões empregadas preservem o mesmo significado ao longo do argumento, caso contrário o argumento será falacioso. Se atribuirmos um significado a um predicado moral quando ele é afirmado e outro quando ele aparece embutido, não conseguiremos sequer elaborar um argumento na forma *modus ponens*. Considere, por exemplo, o seguinte argumento mencionado por Geach:

If doing a thing is bad, getting your little brother to do it is bad.
Tormenting the cat is bad.
Ergo, getting your little brother to torment the cat is bad
(1965, p.463).

Para que esse argumento seja válido é indispensável que a palavra 'bad' tenha o mesmo significado nas quatro vezes em que é mencionada. Ela não pode ser descritiva em um dos usos, manifestar uma emoção ou expressar um imperativo em outro e assim por diante. Ora, está claro que a palavra 'bad' presente na premissa maior não está ali para expressar uma atitude de condenação ou algo semelhante: ninguém diria que na primeira parte dessa premissa se está simplesmente condenando o ato de 'doing a thing'. Desta forma, uma vez que o significado de 'bad' deve permanecer constante, também não se poderia dizer que a premissa menor expressa simplesmente uma atitude de condenação ao ato de torturar um gato (GEACH, 1965, p.464)⁷.

De modo mais geral, para que qualquer argumento da forma 'se *p*, então *q*; *p*; logo *q*' seja válido, é indispensável que *p* tenha o mesmo significado quando aparece sozinho e quando aparece embutido na premissa 'se *p*, então *q*'. No entanto, o não-cognitivista que diz que os predicados morais significam uma coisa quando afirmados isoladamente e outra quando estão embutidos, diz justamente que *p* (*e.g.*, mentir é errado) tem um significado quando está isolada e outro quando está embutida (*e.g.*, se mentir é errado, então...).

Não há dúvidas de que os julgamentos morais são influenciados pelas emoções e servem muitas vezes para expressar atitudes, mas isso não nos permite concluir que esses julgamentos não refletem crenças e não podem ser avaliados como falsos ou verdadeiros (*truth-evaluable judgements*). É difícil explicar até mesmo certas emoções sem recorrermos a esse tipo de julgamento no qual estão presentes condições de verdade. A emoção de culpa, *e.g.*, não envolve apenas um sentimento (*feeling*), mas também certos pensamentos como

⁷ Para Geach, as falhas de teorias não-descritivistas, nas quais as teorias não-cognitivistas se enquadram, resultam da falta de percepção do que ele chama de 'the Frege point': 'A thought may have just the same content whether you assent to its truth or not; a proposition may occur in discourse now asserted, now unasserted, and yet be recognizably the same proposition' (1965, p. 449).

‘eu transgredi uma norma’ ou ‘eu não cumpri com as minhas obrigações’ (JOYCE, 2016, p. 1-2).

Inclusive valorações feitas a partir de uma perspectiva puramente egocêntrica costumam ser parcialmente descritivas. Obviamente elas não são puramente descritivas, pois um elemento essencial desse tipo de valoração é a aprovação, por parte de quem julga, de certos requisitos – estejam eles explícitos ou não – que o objeto avaliado supostamente possui. Mas ela é parcialmente descritiva, pois afirma que o objeto avaliado possui as características necessárias para satisfazer esses requisitos, sejam eles quais forem. Quando alguém que acredita na existência de valores morais objetivos julga moralmente, o caráter descritivo do julgamento se torna ainda mais claro. Quando essa pessoa diz que alguma conduta é boa, ela pretende afirmar algo sobre o mundo, isto é, o suposto fato de que essa conduta satisfaz certos requisitos intrínsecos (MACKIE, 1990, p. 61-3). Se, *e.g.*, um utilitarista diz que é mentir é errado, posso inferir que ele acredita que não mentir ajuda a maximizar a utilidade.

Críticas como essa a respeito do significado de expressões morais quando embutidas levaram a uma reformulação das posições não-cognitivistas. O não-cognitivismo surgiu como uma alternativa bastante distinta em relação às, então predominantes, teorias cognitivistas e realistas. À proporção que seus defensores tentaram explicar determinadas características do discurso moral que serviam de base para os seus críticos, as teorias não-cognitivistas tornaram-se mais sutis e menos contrastantes com as posições contrárias. As suas duas teses centrais (*não-fatualismo semântico* e *não-cognitivismo psicológico*) tornaram-se mais moderadas. A tese de que os julgamentos morais não têm significado descritivo passou a ser a tese de que esse significado existe, mas é secundário⁸. A tese de que esses julgamentos não podem ser verdadeiros ou falsos tornou-se a tese de que eles podem ser verdadeiros ou falsos apenas em um sentido mínimo⁹. Naturalmente essa transição não foi aceita por todos os não-cognitivistas, mas ela é uma tendência bastante nítida nas teorias mais recentes (VAN ROOJEN, 2015).

As teorias expressivistas híbridas, por exemplo, combinam ideias expressivistas e cognitivistas. Desta forma, as sentenças morais seriam instrumentos convencionais utilizados para expressar atitudes de aprovação ou desaprovação, mas ao mesmo tempo elas serviriam para atribuir propriedades; e as atitudes morais seriam parcialmente cognitivas. Qual a propriedade exata atribuída pelo julgamento moral dependeria das atitudes morais de quem o

⁸ Há, por exemplo, versões sofisticadas de emotivismo e de prescritivismo que propõem que os julgamentos morais possuem simultaneamente um significado descritivo e um significado emotivo ou prescritivo. Os julgamentos morais serviriam para predicar propriedades naturais e a aparência de que eles fazem algo mais do que isso decorreria da existência de certo componente expressivo em seu significado (VAN ROOJEN, 2015).

⁹ A respeito da concepção minimalista de verdade, ver Ritchie (2008, p. 175).

profere. No entanto, quem atribui predicados por meio de julgamentos morais não pretende apenas descrever o mundo, mas expressar o seu posicionamento perante aquilo que é descrito (VAN ROOJEN, 2015).

Como resultado da crescente semelhança entre teorias cognitivistas e não-cognitivistas, os argumentos para sustentar uma ou outra posição foram se tornando cada vez mais intrincados. Chegamos a um ponto em que é difícil dizer se as teorias elaboradas para lidar com as críticas feitas pelos cognitivistas ainda merecem ser rotuladas como não-cognitivistas. Filósofos importantes sugerem que a distinção entre cognitivismo e não-cognitivismo não tem mais utilidade (GIBBARD, 2003, p. 183–84; KORSGAARD, 2008, p. 309). Talvez essa distinção tenha colapsado à medida que as teorias não-cognitivistas se modificaram a fim de responder as críticas feitas pelos cognitivistas (VAN ROOJEN, 2015).

De qualquer modo, as teorias não-cognitivistas desempenharam o importante papel de demonstrar que os julgamentos morais desempenham outros papéis além daqueles que os cognitivistas lhes costumavam atribuir. Hare, por exemplo, estava certo ao afirmar que expressões valorativas são utilizadas com a função de recomendar algo (HARE, 1991, p. 91). No entanto, como os defensores de certas teorias híbridas e de algumas teorias não-cognitivistas mais sofisticadas perceberam, isso não implica que elas necessariamente não possuam um significado descritivo ou que não expressem crenças.

Não pretendo aqui ser o árbitro dessa discussão. A forma como estou procurando caracterizar a moralidade poderia ser aceita por cognitivistas de diversos matizes e também por aqueles que adotam posições não-cognitivistas mais amenas. Para a discussão a seguir, é desejável apenas que fique claro que os julgamentos morais envolvem crenças e possuem um sentido descritivo, ainda que eventualmente secundário, de forma que não podemos caracterizá-los como simples manifestações de atitudes conativas.

2. Contribuições da Psicologia Moral

A produção de julgamentos morais envolve o funcionamento de diferentes sistemas cognitivos e ocorre em diferentes etapas, algumas das quais são inacessíveis à introspecção. Dentre os pesquisadores dedicados ao estudo da psicologia moral há praticamente um consenso sobre esse ponto. No entanto, não há acordo sobre a importância relativa de cada uma dessas etapas.

Haidt e Bjorklund acreditam que o julgamento moral é o produto de intuições rápidas e automáticas, as quais são seguidas por um lento processo de raciocínio moral, cuja função principal é encontrar razões para justificar as intuições do indivíduo. O raciocínio do próprio indivíduo raramente teria algum valor causal direto sobre o julgamento moral por ele realizado, exceto

em subculturas específicas como a dos filósofos (2008, p. 193). Segundo, esses autores, nós frequentemente apresentamos argumentos como mero recurso retórico para convencer a nós mesmos de que estamos certos ou para conseguir aliados.

As justificações morais teriam, entretanto, efeito causal quando empregadas em discussões morais travadas com outras pessoas. Há uma série de pesquisas na área de psicologia social que demonstram a tendência de sintonizarmos as nossas atitudes àquelas das pessoas com as quais interagimos (*e.g.*, PENNINGTON & SCHLENKER, 1999; SINCLAIR *et al.*, 2005). Nós sentimos uma necessidade de pertencer a um determinado grupo e os nossos julgamentos morais são moldados por aquilo que os outros membros do grupo acreditam. Além disso, somos péssimos em procurar evidências contrárias às nossas crenças, mas somos muito bons em encontrar falhas nas crenças das pessoas ao nosso redor. Assim, frequentemente só conseguimos enxergar os argumentos contrários às nossas crenças quando outros membros do nosso grupo os explicitam. Por esses motivos, o julgamento moral precisa ser compreendido também como um processo social, não apenas como um ato privado de cognição (HAIDT & BJORKLUND, 2008, p. 181; HAIDT, 2012).

O cérebro avalia e reage incessantemente. Há tecidos nervosos responsáveis por integrar informações oriundas dos ambientes interno e externo e responder a uma questão básica: aproximar-se ou evitar (*approach-avoid*)? No caso humano, essa questão traduz-se na dicotomia bom e mau. Esse tipo de avaliação (bom-mau) é a dimensão mais básica de todos os julgamentos. A mente humana está o tempo todo avaliando e julgando tudo o que ela vê e ouve em termos de bom ou mau, mesmo quando se trata da face de uma pessoa ou de palavras em uma língua estrangeira (HAIDT & BJORKLUND, 2008, p. 186-7).

Aquilo que Haidt chama de intuição moral se refere ao surgimento repentino na consciência de um sentimento valorativo (bom-mau) em relação ao caráter ou às ações de outrem, sem que o sujeito tenha conscientemente seguido etapas de pesquisa, ponderação de evidências ou inferência de conclusões. A intuição moral é um tipo especial de processo automático, no qual sempre está envolvido ao menos algum traço de sentimento valorativo. Na maioria dos casos, essa intuição (*flash of feeling*) conduzirá diretamente à condenação ou à aprovação da conduta em questão¹⁰. Haidt denomina julgamento moral essa experiência consciente de censura ou elogio, que inclui

¹⁰When we refer to moral intuitions, we mean strong, stable, immediate moral beliefs. These moral beliefs are strong insofar as they are held with confidence and resist counter-evidence (although strong enough counterevidence can sometimes overturn them). They are stable in that they are not just temporary whims but last a long time (although there will be times when a person who has a moral intuition does not focus attention on it). They are immediate because they do not arise from any process that goes through intermediate steps of conscious reasoning (although the believer is conscious of the resulting moral belief) (SINNOTT-ARMSTRONG, YOUNG & CUSHMAN, 2010, p. 246).

a crença a respeito da retidão do ato. Para ele, como regra, o julgamento moral apenas confirmará a intuição moral (HAIDT & KESEBIR, 2010, p. 802; HAIDT & BJORKLUND, 2008, p. 188).

Por outro lado, o raciocínio moral é uma atividade mental consciente, caracterizada pela transformação de informações recebidas sobre pessoas e situações com a finalidade de alcançar um julgamento moral. Qualificar o raciocínio moral como um processo consciente significa afirmar que ele é intencional, controlável e exige esforço de quem o realiza (HAIDT, 2001, p. 818). Embora, no âmbito da consciência, quem raciocina moralmente acredite estar tentando chegar a um julgamento, normalmente, segundo Haidt, a decisão já foi tomada intuitivamente, e o raciocínio serve apenas para encontrar justificativas (HAIDT & BJORKLUND, 2008, p. 189).

Haveria dois tipos distintos de processos mentais envolvidos nessas valorações, um deles seria automático e o outro controlado. A maior parte da cognição humana pertence ao primeiro tipo e é semelhante à cognição presente em outros animais. Trata-se de um sistema muito antigo, surgido há milhões de anos, cujo funcionamento permite a solução de inúmeros problemas complexos de forma inconsciente. Ele possibilita, por exemplo, o reconhecimento de pessoas e objetos instantaneamente. Por outro lado, há um tipo de cognição que é único ao ser humano. Ele opera por meio de palavras e, portanto, não pode ser mais antigo do que a linguagem. Ao contrário do processo anterior, a habilidade para raciocinar por meio da linguagem é um processo lento e exige esforço. Haidt emprega a metáfora do elefante que é conduzido por um homem para explicar como esses diferentes tipos de cognição interagem (HAIDT, 2013)¹¹:

Automatic cognition -the elephant- is sometimes called 'hot' cognition because it has the power to motivate us to action. The neurological systems for judgement are connected to the neurological systems for behavior. Controlled cognition, in contrast, is sometimes called 'cool' cognition. It is not connected to behavioral centers in the brain. So please imagine the human mind as a small and somewhat ineffectual rider perched on the back of a large, powerful, and rather smart elephant. The rider can try to steer the elephant, and if the elephant has no particular desire to go one way or the other, it may listen to the rider. But if it has its own desires, it's going to do what it wants to do (HAIDT, 2013).

Há uma série de pesquisas consistentes com a ideia de que muitas vezes avaliamos instantaneamente e depois raciocinamos apenas para justificar a avaliação feita previamente (ZAJONC, 1980; MONAHAN *et al.*, 2000; LUO

¹¹ A tese de que a cognição humana é formada por dois sistemas, um lento e outro rápido, foi consagrada na psicologia por meio dos trabalhos de Kahneman e Tversky sobre heurística e vieses. Para uma descrição detalhada desses sistemas ver Kahneman, 2013.

et al., 2006). Dentre os estudos consistentes com esta ideia, há pesquisas empíricas realizadas através de entrevistas sobre julgamentos morais (HAIDT *et al.*, 1993), de questionários respondidos pela Internet sobre *trolley problems* (HAUSER *et al.*, 2008), de manipulação de julgamentos morais por meio da indução de emoções com a utilização de hipnose e de alterações ambientais (WHEATLEY & HAIDT, 2005; SCHNALL *et al.*, 2008), e de imagens por ressonância magnética funcional (GREENE & HAIDT, 2002).

Talvez o exemplo mais claro de como a consciência fornece racionalizações *post hoc* seja fornecido pelos estudos de Gazzaniga (1985) em pacientes que tiveram o corpo caloso responsável pela conexão dos dois hemisférios cerebrais seccionados. Quando um paciente realizava uma ação em função de um estímulo sobre o hemisfério cerebral direito, por exemplo, levantar e sair andando, o hemisfério esquerdo, responsável pela linguagem, não respondia algo como ‘não sei por que estou fazendo isso’, mas inventava uma razão qualquer, como ‘estou indo buscar um refrigerante’. Segundo Gazzaniga, nosso raciocínio verbal consciente não comanda as nossas ações; ele funciona como uma secretária de imprensa encarregada de inventar explicações convincentes para qualquer coisa que a pessoa faça. Como não temos acesso ao sistema de cognição rápida que desencadeia nossos julgamentos e condutas, só nos restaria empregar o sistema de cognição lenta para adivinhar explicações para os seus resultados (GAZZANIGA, 1985; HAIDT & BJORKLUND, 2008; WAHL, 2014, p. 69-71).

No entanto, como veremos adiante, há casos em que os processos conscientes de raciocínio parecem sim desempenhar um papel causal sobre a produção dos nossos julgamentos morais.

Julgamentos morais exigem crenças e ideias como merecimento, capacidade de agir, desaprovação e transgressão, as quais não podem ser explicadas apenas com referências a mecanismos puramente conativos (JOYCE, 2013, p. 550; JOYCE, 2014, p. 135-137). Somente tendo em vista a existência de elementos cognitivos podemos explicar por que normalmente não fazemos julgamentos morais a respeito de animais. Em razão de não possuírem um tipo específico de capacidade para agir que consideramos essencial para a valoração moral, não condenamos moralmente as ações deles.

Além disso, o raciocínio não serve apenas para racionalizar decisões previamente adotadas. Ainda que, como sustenta Haidt, o raciocínio consciente apenas em poucos casos tenha um papel causal na produção de julgamentos morais, ele pode ser bastante relevante. Uma única decisão tomada por meio desse tipo de raciocínio pode exercer influência duradoura quando ela altera as motivações que orientam a conduta do indivíduo. Por exemplo, muitos indivíduos, por meio de raciocínio consciente chegam à conclusão de que é imoral comer carne e isso acaba produzindo alterações no comportamento que se ramificam ao longo do tempo (PIZARRO & BLOOM, 2003).

Alterações como essa, principalmente por meio de processos de interação social, influenciarão as emoções e as intuições morais que teremos no futuro. Exemplos como esse do vegetarianismo moral demonstram como alterações nos julgamentos morais podem ser provocadas por processos de deliberação e de reflexão (STERELNY, 2012, p. 141). O nojo e a repulsa que vegetarianos sentem por carne é uma consequência, não uma causa, de suas convicções morais (FESSLER *et al.*, 2003). De forma semelhante, pessoas que consideram a homossexualidade repugnante podem tornar-se tolerantes em razão de processos de reflexão e modificar as suas intuições e sentimentos iniciais. Sterelny menciona vários exemplos da história recente que demonstram como as intuições dos indivíduos de uma sociedade podem ter mudanças drásticas em poucos anos. Além da postura em relação à homossexualidade, as respostas à crueldade com animais, aos métodos de educação das crianças, ao papel das mulheres, à obesidade e à depressão sofreram importantes alterações que se refletem nas nossas reações viscerais, não apenas no plano da consciência (STERELNY, 2010, p. 286-7). Casos como esses demonstram a instabilidade e o dinamismo de nossas intuições morais. A mudança de convicção atrelada a uma adequação das práticas cotidianas à nova convicção acaba por alterar os tipos de objeto que provocam repugnância no indivíduo (STERELNY, 2010, p. 282-4 e 293-4)¹².

Nesse sentido, até mesmo Prinz reconhece que o raciocínio consciente pode desempenhar algum papel relevante na produção de julgamentos morais. Prinz acredita que temos valores básicos – não matar, não ser cruel, ser justo, etc. – que emergem diretamente de nossas emoções; e valores derivados desses valores básicos. Não comer carne, *e.g.*, seria um valor derivado. Aqueles que acreditam que comer carne é errado argumentam que essa conduta é uma forma de crueldade. O papel do raciocínio seria identificar qual valor básico se aplica ao caso em questão¹³ (PRINZ, 2012, p. 262-3).

Há, além disso, uma literatura substancial sobre atitudes sociais implícitas — em especial sobre atitudes raciais — que corrobora a tese de que processos de raciocínio consciente podem prevalecer sobre processos

¹²Skorupski também procura enfatizar como considerações teóricas podem modificar os julgamentos morais: '[julgamentos morais] are, in this respect, comparable to verdicts about redness. These latter judgments are indeed accountable initially to perceptual experience. But they are accountable also to criteria of theoretical reason. Those latter, external, criteria are mediated through a context of physical theory, a body of beliefs which can retrospectively override a judgment that something is red, even when that judgment is default-justified by perceptual experience. If, for example, we find that atmospheric conditions, or the relative speed of a body to the perceiver, can affect its apparent colour, then we may withdraw colour judgments which perceptual experience would otherwise have justified by default. At the limit, indeed, we could discover that nothing is red, even though we have sensations of red' (1993, p. 132-4).

¹³Extensive reasoning is often required to figure out which if any of our basic values apply in a given case. Most public policy decisions are like this, and they require extensive reasoning as a result. But once we establish which basic values are applicable, reasoning stops and passion settles the issue' (PRINZ, 2012, p. 262-3).

intuitivos. Já é algo bastante consolidado nessa literatura que uma pessoa pode não manifestar qualquer atitude preconceituosa em testes que envolvem medições explícitas, como questionários, por exemplo, mas apresentar vieses racistas em testes de associação implícita. Na versão mais comum desse teste, figuras de rostos humanos, negros ou brancos, são associadas com palavras consideradas positivas ou negativas. Os participantes tipicamente associam mais facilmente pessoas brancas a palavras positivas e pessoas negras a palavras negativas. Quando elas precisam fazer a associação contrária, isto é, associar pessoas negras a palavras positivas ou pessoas brancas a palavras negativas, elas costumam demorar mais tempo. Esse tipo de efeito acontece mesmo com pessoas que quando entrevistadas não manifestam qualquer atitude racista (*e.g.* GREENWALD *et al.*, 1998; MALLON & NICHOLS, 2010, p. 300-2).

Essa literatura indica que processos controlados conscientemente podem exercer uma importante influência sobre as respostas verbais e sobre o comportamento. Quando esses processos são sobrecarregados, por meio de tarefas que requerem um acréscimo da atividade cognitiva, ou então quando eles são enfraquecidos, *e.g.*, por meio do consumo de álcool, as respostas verbais e o comportamento passam a se alinhar mais diretamente com as atitudes implícitas.

Richeson e Shelton (2003), por exemplo, conduziram um experimento destinado a examinar a influência de interações interracialis no funcionamento cognitivo. Nessa experiência, os participantes brancos tinham que manter contato por certo tempo com outro indivíduo, branco ou negro, e, então, realizar um teste que exigia atenção e esforço cognitivo (teste de Stroop). Antes da experiência, as atitudes raciais automáticas dos participantes foram medidas por testes de associação implícita. Aqueles que apresentaram vieses racistas mais acentuados nessa etapa prévia obtiveram piores resultados no teste cognitivo. No entanto, esse efeito só foi observado naqueles que haviam interagido com um indivíduo negro. Isto é, aqueles participantes que apresentavam vieses racistas acentuados, mas que interagiram somente com o indivíduo branco obtiveram resultados semelhantes àqueles que não manifestavam esses vieses racistas de modo acentuado. De acordo com a interpretação dos autores do experimento, a interação com o indivíduo negro exigia dos participantes mais preconceituosos uma grande utilização de recursos cognitivos para suprimir as respostas preconceituosas automáticas. Essa alta demanda cognitiva acabava exaurindo os recursos necessários para completar o teste e, assim, prejudicava o desempenho nos testes subsequentes¹⁴.

Além disso, uma experiência liderada por Bartholow (BARTHOLOW, DICKTER & SESTIR, 2006) demonstrou que o consumo

¹⁴Em outro estudo, Richeson (RICHESON *et al.*, 2003) apresentou neuroimagens consistentes com essa interpretação.

de álcool prejudica de forma significativa a habilidade para regular vieses implícitos e, deste modo, inibir respostas racistas. Segundo os autores, esse resultado se deve ao enfraquecimento do controle cognitivo provocado pelo álcool¹⁵.

Portanto, há toda uma linha de pesquisa elaborada a partir do paradigma da dupla cognição que fornece muitos indícios de que pelo menos em relação a certos comportamentos considerados moralmente relevantes, processos automáticos acabam cedendo a processos cognitivos mais lentos e reflexivos (MALLON & NICHOLS, 2010, p. 300-2).

Como mencionei anteriormente, Haidt apresenta inúmeros casos nos quais as pessoas condenam moralmente uma conduta, mas não conseguem encontrar justificativas para o julgamento (HAIDT, 2001). Os resultados apresentados por Haidt são frequentemente empregados por aqueles que pretendem demonstrar que os processos conscientes de reflexão não estão conectados com os julgamentos morais (*e.g.*, DWYER, 2009). Certos dilemas sobre o incesto provocam esse fenômeno ('moral dumbfounding'). Por exemplo, é moralmente aceitável que irmãos por parte de pai, criados separados, mantenham relações sexuais consensuais com a utilização de dois métodos contraceptivos? Muitos respondem a questões como essa afirmando que a conduta é moralmente inaceitável, mas não sabem fundamentar a resposta, apenas dizem que é errado.

No entanto, exemplos como esses não demonstram que o julgamento moral não foi produzido por um processo de raciocínio consciente. Em casos assim é importante fazermos uma distinção entre justificação e explicação. As pesquisas de Haidt demonstram que os participantes frequentemente falham quando são instigados a fornecer uma justificativa plausível para os seus julgamentos. Contudo, essas falhas não implicam necessariamente uma incapacidade para explicar esses mesmos julgamentos. Quando um participante diz que dois irmãos fazerem sexo é errado porque é incesto, ele não oferece uma justificativa sobre por que devemos considerar o incesto errado, mas pode estar fornecendo uma explicação acurada sobre a origem de seu julgamento. As pessoas têm consciência sobre a existência de uma regra que proíbe o incesto. Portanto, quando elas dizem que 'é errado porque é incesto', elas podem estar identificando uma norma acessível à consciência que foi determinante na produção do julgamento (MALLON & NICHOLS, 2011, p. 285).

Ademais, embora as pessoas não sejam capazes de fornecer justificativas para seus julgamentos em algumas situações, elas são perfeitamente capazes em outras. Em um estudo conduzido por Cushman e seus colegas (CUSHMAN *et al.*, 2006), muitos participantes foram capazes de justificar os seus julgamentos com base na distinção entre ação e omissão. O

¹⁵Para uma discussão mais aprofundada sobre esses estudos, ver Nosek *et al.* (2007) e Kelly, Machery & Mallon (2010).

fato mais relevante é que os julgamentos feitos previamente por aqueles que foram capazes de articular esse princípio realmente demonstravam uma maior utilização da distinção entre ação e omissão. Essa ligação entre ter um princípio elaborado conscientemente e utilizá-lo nos julgamentos indica que o raciocínio desempenha algum papel no processo de elaboração de julgamentos morais (CUSHMAN, YOUNG & GREENE, 2010, p. 60).

Ainda, em certas circunstâncias, as pessoas alteram os seus julgamentos para amenizar as inconsistências de suas intuições iniciais (LOMBROZO, 2009) ou quando são incentivadas a raciocinar sobre o assunto (PIZARRO, UHLMANN, & BLOOM, 2003, p. 657-8). Outros estudos de neuroimagem (GREENE *et al.*, 2004) e de tempo de reação (GREENE *et al.*, 2008)¹⁶ associam processos de raciocínio aos julgamentos morais. Pelos menos em algumas oportunidades o compromisso consciente que temos com certas normas morais influencia as nossas deliberações sobre o que é certo ou errado (CUSHMAN, YOUNG & GREENE, 2010, p. 47-8).

Todos esses resultados que indicam a importância do raciocínio para os julgamentos morais contradizem a posição não-cognitivista clássica. Como procurei demonstrar por meio do argumento das expressões embutidas, é difícil explicar a partir de posições não-cognitivistas esse tipo de raciocínio e as discussões morais travadas a partir dele.

As discussões sobre moralidade indicam que algo mais do que emoções está envolvido. Emoções podem desempenhar um papel importante na origem dos julgamentos morais, mas quando dizemos que alguma conduta é errada não queremos dizer apenas que não gostamos dela. Nossos julgamentos envolvem a formação de crenças e o conjunto dessas crenças acaba formando uma teoria normativa. Quando discutimos sobre a moralidade, estamos discutindo sobre o conteúdo e as implicações de uma teoria normativa que é em alguma medida compartilhada dentro de um grupo (NICHOLS, 2008, p. 268).

Como disse no começo desse capítulo, Darwin já percebia que os julgamentos morais não podem ser confundidos com emoções ou com instintos, pois eles requerem uma série de capacidades intelectuais. Não há somente um processo psicológico responsável pela produção de julgamentos morais. Esses julgamentos resultam de interações complexas entre diversos sistemas psicológicos. Portanto, uma abordagem sobre eles não deve se restringir a debater se os julgamentos morais são fruto de processos cognitivos ou de 'affects'; de raciocínio consciente ou de intuições (CUSHMAN, YOUNG & GREENE, 2010, p. 66-8).

¹⁶Nesse estudo Greene e seus colegas verificaram que a manipulação da carga cognitiva dos testes interfere de modo seletivo com os julgamentos morais utilitaristas. Eles concluíram que esse resultado é consistente com a teoria que associa julgamentos morais utilitaristas a processos cognitivos controlados e julgamentos não-utilitaristas a respostas emocionais automáticas.

Dentre as diversas capacidades intelectuais envolvidas na moralidade, destaca-se a autoconsciência. Para julgarmos normativamente uma ação que nós mesmos praticamos precisamos ter a capacidade de identificar os motivos que nos levaram a praticá-la e de refletir sobre eles, ou seja, de compará-los com outros motivos que poderiam ter nos motivado a agir de forma diferente. Pensar de maneira normativa exige que tenhamos consciência dos fundamentos de nossas crenças e ações, e isso implica a autoconsciência, pois envolve a capacidade de nos identificarmos como os sujeitos de nossas representações mentais. Um ser sem autoconsciência pode estar consciente da existência de um objeto que ele deseja e agir de acordo com essa informação. No entanto, um ser com autoconsciência, além de estar consciente da existência do objeto, está também consciente do fato de que *ele* deseja o objeto. Ele não pensa apenas no objeto que deseja, mas também nos seus próprios desejos que o inclinam a agir de determinada forma. Esta autoconsciência a respeito dos motivos garante uma distância reflexiva que possibilita ao sujeito questionar os seus próprios motivos (KORSGAARD, 2006, p. 112-116). Através da autoconsciência o indivíduo é capaz de se colocar na posição de espectador de seus próprios desejos. A partir dessa posição ele pode comparar as suas ações no presente às passadas e aprová-las ou desaprová-las. A autoconsciência possibilita a existência de atitudes de segunda ordem, em que o sujeito torna-se capaz de ter desejos sobre os seus próprios desejos e crenças sobre as suas próprias crenças (BAKER, 2013, p. 191-192). Ele pode pensar que teria sido melhor se tivesse agido de acordo com outro desejo (DARWIN, 1871/2004, p. 73-74)¹⁷.

Esse tipo de reflexão requer outras capacidades, como a memória e a linguagem. A capacidade de refletir sobre os motivos que nos levaram a agir de determinada forma depende da capacidade de nos lembrarmos de nossas ações, dos seus motivos e dos resultados produzidos: a lembrança de que uma ação provocou um sentimento de insatisfação pode, *e.g.*, fornecer subsídios para o juízo de que *deveríamos* ter agido de outra forma (DARWIN, 1871, p. 70-72). O agente que dispõe de autoconsciência e de uma capacidade para se recordar de experiências passadas também dispõe da capacidade para pensar normativamente (BAKER, 2013, p. 191-192), ainda que em um sentido puramente prudencial.

A linguagem, por sua vez, também é condição necessária para autoconsciência, tendo em vista que a existência desta depende da capacidade do sujeito ter um conceito de si mesmo - *self-concept* (BAKER, 2013, p. 135). Além de possibilitar a reflexão consciente sobre os motivos de nossas ações e a formulação da crença de que deveríamos ter agido de outra forma, a linguagem permite também que compartilhem essas reflexões e crenças com os outros indivíduos de nosso grupo. Como indica o modelo social-

¹⁷ Sobre a questão do papel da autoconsciência, ver também Baker (2013) e Damasio (2010).

intuicionista desenvolvido por Haidt (2012), esse tipo de interação social é outro importante elemento na conformação da moralidade.

Além da autoconsciência, da linguagem e da memória, a habilidade de atribuir estados mentais a outras pessoas (teoria da mente) e de compreender que esses estados mentais podem ser diferentes daqueles que possuímos também se vincula diretamente à produção de julgamentos morais. Animais em geral são capazes de considerar os comportamentos passados de outros indivíduos a fim de orientar as suas condutas futuras. No entanto, seres humanos não se satisfazem com simples descrições de comportamento. Nós procuramos compreendê-los postulando a existência de uma mente por trás deles; buscamos explicar *por que* o indivíduo se comportou de determinada forma, e essas explicações envolvem suposições sobre crenças e objetivos¹⁸ (FITCH, 2010, p. 135-40).

Não há muitos indícios de que outros animais possam ter essa capacidade. Até mesmo crianças de 3 a 4 anos de idade são incapazes de representar as crenças de outros indivíduos quando essas divergem de suas próprias crenças, como ficou demonstrado no clássico estudo conduzido por Wimmer e Perner (1983).

O tipo de experimento desenvolvido por eles ficou conhecido como teste ‘Sally-Anne’. Esse teste envolve a encenação de uma história com dois personagens, Sally e Anne. Sally esconde uma bola de gude em uma cesta e então deixa a sala. Enquanto ela está ausente, Anne muda a boca de gude de local, colocando-a em uma caixa. Então, quando Sally retorna à sala, pergunta-se à criança que assistiu à encenação ‘Onde Sally procurará a bola de gude?’ ou ‘Onde Sally pensa que está a bola de gude?’. Em testes desse tipo, crianças de até três anos respondem ‘na caixa’ de modo consistente. Apenas crianças um pouco mais velhas, de três ou quatro anos, são capazes de separar o seu próprio conhecimento acerca de onde está a bola de gude das crenças de Sally (BARON-COHEN, LESLIE & FRITH, 1985; FITCH, 2010, p. 136-9).

Assim como ocorre com relação à autoconsciência, a nossa habilidade de atribuir estados mentais também nos torna seres capazes de ter atitudes de segunda ordem. A capacidade que temos de ter crenças sobre crenças é fundamental para explicar muitos dos nossos julgamentos morais, pois, como regra, julgamos moralmente apenas os seres que acreditamos serem capazes de agir intencionalmente: não dizemos que uma telha agiu errado ao descolar-se do telhado e atingir a cabeça de alguém.

¹⁸Quando analisamos as condutas de outras pessoas empregamos o que Dennett chama de ‘intentional stance’: ‘Here is how it works: first you decide to treat the object whose behavior is to be predicted as a rational agent; then you figure out what beliefs that agent ought to have, given its place in the world and its purpose. Then you figure out what desires it ought to have, on the same considerations, and finally you predict that this rational agent will act to further its goals in the light of its beliefs. A little practical reasoning from the chosen set of beliefs and desires will in many — but not all — instances yield a decision about what the agent ought to do; that is what you predict the agent will do’ (1989, p. 17).

Embora os julgamentos morais envolvam todas essas capacidades mencionadas, os autores que se propuseram a investigar a evolução da moralidade até o momento não foram muito além de uma investigação acerca da evolução da cooperação e de atitudes pró-sociais. É provável que muitos dos aspectos relacionados ao que Darwin chamava de instintos sociais e das emoções que lhes são subjacentes tenham sido selecionados em razão de favorecerem comportamentos cooperativos.

De modo semelhante, outras capacidades indispensáveis para a produção de julgamentos morais como a capacidade para a linguagem, para a autoconsciência, etc. podem ser adaptações surgidas para auxiliar na resolução de problemas enfrentados no Pleistoceno. Cada uma dessas capacidades pode ter sido selecionada por razões distintas, pois, embora elas sejam empregadas na formulação de julgamentos morais, elas não são dedicadas apenas a essa tarefa.

O grande problema das principais teorias adaptacionistas é que nenhuma delas demonstrou que a capacidade para fazer julgamentos morais depende de algo como uma adaptação dedicada à moralidade. O simples fato de a moralidade cumprir algumas funções não implica que ela seja uma adaptação biológica.

Quando percebemos todas as capacidades intelectuais que temos, as nossas emoções, a necessidade de vivermos em grupos, etc., não espanta que tenhamos uma capacidade para fazer julgamentos morais; ou seja, não está claro que a produção de julgamentos morais dependa de uma adaptação específica. Talvez nenhum desses elementos necessários ao surgimento da moralidade tenha sido selecionado originariamente em razão de favorecê-la. Assim, ela pode ter surgido como um efeito secundário de nossos 'instintos sociais' e de nossas 'capacidades intelectuais'¹⁹. No entanto, não podemos excluir a possibilidade de que ela seja uma adaptação secundária, caso os elementos que a compõem tenham sido modificados pela seleção natural em razão de a forma de eles interagirem ter desempenhado algum impacto na aptidão de indivíduos ou de grupos. Por enquanto, não há suporte empírico para qualquer conclusão definitiva a esse respeito.

Responder se a moralidade é uma adaptação envolveria investigar como e quando essas capacidades intelectuais surgiram, quais as pressões seletivas que atuaram nesse processo e se essas pressões foram suficientes para modificar a capacidade para fazer julgamentos morais. Um relato completo a respeito da evolução da moralidade envolveria também uma história sobre a evolução das capacidades intelectuais subjacentes à moralidade.

¹⁹Francisco Ayala, por exemplo, considera que a capacidade intelectual foi selecionada em razão de favorecer a construção e a utilização de ferramentas. A capacidade para fazer julgamentos morais seria um efeito secundário desta capacidade (2010, p.5)

Conclusão

As posições não-cognitivistas clássicas negam que os julgamentos morais expressem crenças e tenham conteúdo descritivo. Contudo, argumentos como o das expressões embutidas demonstram que para que essas posições sejam consistentes, elas precisam negar também a possibilidade de raciocinarmos moralmente de acordo com as regras da lógica. As posições não-cognitivistas atuais buscam se adequar a críticas como essa e admitem que os julgamentos morais envolvem, parcialmente, crenças e um sentido descritivo, embora sustentem que a função principal desses julgamentos é manifestar atitudes de aprovação ou de reprovação. Por sua vez, os estudos psicológicos mencionados corroboram a tese de que o raciocínio consciente e as crenças que possuímos influenciam nossos julgamentos morais, de modo que eles não podem ser reduzidos a simples manifestações de atitudes conativas.

Portanto, como Darwin já percebia ao fazer a distinção entre ‘social instincts’ e ‘intellectual powers’, não podemos confundir altruísmo com moralidade. Os julgamentos morais não se restringem à manifestação de instintos sociais, pois eles requerem uma série de capacidades intelectuais; enquanto o altruísmo e até mesmo punições altruístas podem estar presentes independentemente da presença dessas capacidades. No entanto, muitas das abordagens evolutivas acerca da moralidade falham ao não distinguir esses aspectos e restringirem-se a uma investigação acerca da evolução da cooperação e de atitudes pró-sociais.

Referências

AYALA, F. J. Colloquium paper: the difference of being human: morality. *Proceedings of the National Academy of Sciences of the United States of America*, PMID: 20445091, v. 107 Suppl, p. 9015–9022, 2010.

AYER, A. J. *Language, Truth and Logic*. 2. ed. Dover Publications Inc., 2002.

BAKER, L. R. *Naturalism and the First-Person Perspective*. 1. ed. Oxford; New York: Oxford University Press, 2013.

BARON-COHEN, S.; LESLIE, A. M.; FRITH, U. Does the autistic child have a ‘theory of mind’? *Cognition*, v. 21, n. 1, p. 37–46, out. 1985.

BARTHOLOW, B. D.; DICKTER, C. L.; SESTIR, M. A. Stereotype activation and control of race bias: cognitive control of inhibition and its impairment by alcohol. *Journal of Personality and Social Psychology*, PMID: 16536651, v. 90, n. 2, p. 272–287, fev. 2006.

CARNAP, R. *Philosophy and Logical Syntax*. New York: AmsPrInc, 1979.

CUSHMAN, F.; YOUNG, L.; GREENE, J. D. Multi-system moral psychology. In: DORIS, J. M. (Org.). *The handbook of moral psychology*. Oxford; New York: Oxford University Press, 2010. p. 47–71.

CUSHMAN, F.; YOUNG, L.; HAUSER, M. The role of conscious reasoning and intuition in moral judgment: testing three principles of harm. *Psychological Science*, PMID: 17201791, v. 17, n. 12, p. 1082–1089, dez. 2006.

DAMASIO, A. R. *Self Comes to Mind: Constructing the Conscious Brain*. Pantheon Books, 2010.

DARWIN, C. *The descent of man*. Overland Park: Digireads.com, 2004.

DE WAAL, F. *Good natured: The origins of right and wrong in humans and other animals*. Cambridge: Harvard University Press, 1996.

DENNETT, D. C. *The Intentional Stance*. Rev. ed. Cambridge, Mass.: A Bradford Book, 1989.

DWYER, S. Moral dumbfounding and the linguistic analogy: Methodological implications for the study of moral judgment. *Mind and Language*, v. 24, n. 3, p. 274–296, 2009.

FESSLER, D. M. T. *et al.* Disgust sensitivity and meat consumption: A test of an emotivist account of moral vegetarianism. *Appetite*, PMID: 12880619, v. 41, n. 1, p. 31–41, 2003.

FITCH, W. T. *The Evolution of Language*. 1. ed. Cambridge ; New York: Cambridge University Press, 2010.

FITZPATRICK, W. Morality and Evolutionary Biology. In: ZALTA, E. N. (Org.). *The Stanford Encyclopedia of Philosophy*. Spring 2016 ed. Disponível em: <<http://plato.stanford.edu/archives/spr2016/entries/morality-biology/>>. Acesso em: 20 jul. 2016.

GAZZANIGA, M. S. *The Social Brain: Discovering the Networks of the Mind*. 1. ed. New York: Basic Books, 1985.

GEACH, P. T. Assertion. *Philosophical Review*, v. 74, n. 4, p. 449–465, 1965.

GIBBARD, A. *Thinking How to Live*. Cambridge, Mass: Harvard University Press, 2003.

GREENE, J. D. *et al.* Cognitive Load Selectively Interferes with Utilitarian Moral Judgment. *Cognition*, PMID: 18158145, v. 107, n. 3, p. 1144–1154, jun. 2008.

GREENE, J. D. *et al.* The neural bases of cognitive conflict and control in moral judgment. *Neuron*, PMID: 15473975, v. 44, n. 2, p. 389–400, 14 out. 2004.

GREENE, J.; HAIDT, J. How (and where) does moral judgment work? *Trends in Cognitive Sciences*, PMID: 12475712, v. 6, n. 12, p. 517–523, 1 dez. 2002.

GREENWALD, A. G.; MCGHEE, D. E.; SCHWARTZ, J. L. Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology*, PMID: 9654756, v. 74, n. 6, p. 1464–1480, jun. 1998.

HAIDT, J. Moral Psychology and the Law: How Intuitions Drive Reasoning, Judgment, and the Search for Evidence. *Alabama Law Review*, v. 64, p. 867–903, 2013.

HAIDT, J. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, PMID: 11212634, v. 108, n. 4, p. 814–834, 2001.

HAIDT, J. *The righteous mind: why good people are divided by politics and religion*. 1. ed. New York: Pantheon Books, 2012.

HAIDT, J.; JOSEPH, C. The Moral Mind: How Five Sets of Innate Intuitions Guide the Development of Many Culture-Specific Virtues, and Perhaps Even Modules. In: CARRUTHERS, P.; LAURENCE, S.; STICH, S. (Org.). *The Innate Mind*. New York: Oxford University Press, 2008. v. 3.

HAIDT, J.; KESEBIR, S. Morality. In: FISKE, S.; GILBERT, D.; LINDZEY, G. (Org.). *Handbook of Social Psychology*. 5. ed. Hoboken, NJ: Wiley, 2010. p. 797–832.

HAIDT, J.; KOLLER, S. H.; DIAS, M. G. Affect, culture, and morality, or is it wrong to eat your dog? *Journal of Personality and Social Psychology*, PMID: 8229648, v. 65, n. 4, p. 613–628, out. 1993.

HARE, R. M. *The Language of Morals*. London; New York: Oxford University Press, 1991.

HAUSER, M.; YOUNG, L.; CUSHMAN, F. Reviving Rawls' linguistic analogy: Operative principles and the causal structure of moral actions. In: SINNOTT-ARMSTRONG, W. (Org.). *Moral Psychology, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge: MIT press, 2008. v. 2. p. 107–143.

JOYCE, R. Evolution and Moral Naturalism. In: CLARK, K. J. (Org.). *The Blackwell Companion to Naturalism*. John Wiley & Sons, Inc, 2016. p. 369–385. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/9781118657775.ch26/summary>>. Acesso em: 20 jul. 2016.

JOYCE, R. Moral Anti-Realism. In: ZALTA, E. N. (Org.). *The Stanford Encyclopedia of Philosophy*. Fall 2015 ed. Disponível em: <<http://plato.stanford.edu/archives/fall2015/entries/moral-anti-realism/>>. Acesso em: 20 jul. 2016.

JOYCE, R. *The evolution of morality*. Cambridge: MIT Press, 2006.

JOYCE, R. The many moral nativisms. In: STERELNY, K. *et al.* (Org.). *Cooperation and its Evolution*. MIT press, 2013.

JOYCE, R. The origins of moral judgment. *Behaviour*, v. 151, n. 2–3, p. 261–278, 2014.

KAHNEMAN, D. *Thinking, Fast and Slow*. New York: Farrar Straus Giroux, 2013.

KELLY, D.; MACHERY, E.; MALLON, R. Race and Racial Cognition. In: DORIS, J. M. (Org.). *The Moral Psychology Handbook*. Oxford; New York: Oxford University Press, 2010.

KORSGAARD, C. Morality and the distinctiveness of human action. In: MACEDO, S.; OBER, J. (Org.). *Primates and philosophers: How morality evolved*. Princeton: Princeton University Press, 2006. p. 98–119.

KORSGAARD, C. M. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. 1. ed. Oxford; New York: Oxford University Press, 2008.

LOMBROZO, T. The Role of Moral Commitments in Moral Judgment. *Cognitive Science*, v. 33, n. 2, p. 273–286, 1 mar. 2009.

LUO, Q. *et al.* The neural basis of implicit moral attitude—an IAT study using event-related fMRI. *NeuroImage*, PMID: 16418007, v. 30, n. 4, p. 1449–1457, 1 maio 2006.

MACKIE, J. L. *Ethics: Inventing Right and Wrong*. New ed. Penguin, 1990.

MALLON, R.; NICHOLS, S. Dual Processes and Moral Rules. *Emotion Review*, v. 3, n. 3, p. 284–285, 7 jan. 2011.

MALLON, R.; NICHOLS, S. Rules. In: DORIS, J. M. (Org.). *The Moral Psychology Handbook*. 1. ed. Oxford; New York: Oxford University Press, 2010.

MAMELI, M. Meat made us moral: a hypothesis on the nature and evolution of moral judgment. *Biology & Philosophy*, v. 28, n. 6, p. 903–931, 18 set. 2013.

MONAHAN, J. L.; MURPHY, S. T.; ZAJONC, R. B. Subliminal mere exposure: specific, general, and diffuse effects. *Psychological Science*, PMID: 11202490, v. 11, n. 6, p. 462–466, nov. 2000.

NICHOLS, S. Sentimentalism naturalized. In: SINNOTT-ARMSTRONG, W. (Org.). *Moral Psychology, The Cognitive Science of Morality: Intuition and Diversity*. Cambridge: MIT press, 2008. v. 2. p. 255–274.

NOSEK, B. A.; GREENWALD, A. G.; BANAJI, M. R. The Implicit Association Test at Age 7: A Methodological and Conceptual Review. *Social psychology and the unconscious: The automaticity of higher mental processes*. Frontiers of social psychology. New York, NY, US: Psychology Press, 2007. p. 265–292.

PENNINGTON, J.; SCHLENKER, B. R. Accountability for Consequential Decisions: Justifying Ethical Judgments to Audiences. *Personality and Social Psychology Bulletin*, v. 25, n. 9, p. 1067–1081, 1 set. 1999.

PIZARRO, D. A.; BLOOM, P. The intelligence of the moral intuitions: comment on Haidt (2001). *Psychological Review*, PMID: 12529062, v. 110, n. 1, p. 193–198, jan. 2003.

PIZARRO, D. A.; UHLMANN, E.; BLOOM, P. Causal deviance and the attribution of moral responsibility. *Journal of Experimental Social Psychology*, v. 39, n. 6, p. 653–660, 2003.

PRINZ, J. *Beyond human nature: how culture and experience shape our lives*. Allen Lane, an imprint of Penguin Books, 2012.

PRINZ, J. Where Do Morals Come From?—A Plea for a Cultural Approach. In: CHRISTEN, M. *et al.* (Org.). *Empirically Informed Ethics: morality between facts and norms*. Berlin: Springer, 2013.

RICHESON, J. A. *et al.* An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience*, v. 6, n. 12, p. 1323–1328, dez. 2003.

RICHESON, J. A.; SHELTON, J. N. When Prejudice Does Not Pay Effects of Interracial Contact on Executive Function. *Psychological Science*, PMID: 12741756, v. 14, n. 3, p. 287–290, 1 maio 2003.

RITCHIE, J. *Understanding naturalism*. Stocksfield: Acumen Pub. Ltd., 2008.

SAYRE-MCCORD, G. Metaethics. In: ZALTA, E. N. (Org.). *The Stanford Encyclopedia of Philosophy*. Summer 2014 ed. Disponível em: <<http://plato.stanford.edu/archives/sum2014/entries/metaethics/>>. Acesso em: 20 jul. 2016.

SCHNALL, S. *et al.* Disgust as embodied moral judgment. *Personality and social psychology bulletin*, PMID: 18505801, v. 34, n. 8, p. 1096–1109, 2008.

SINCLAIR, S. *et al.* Social tuning of automatic racial attitudes: the role of affiliative motivation. *Journal of Personality and Social Psychology*, PMID: 16287420, v. 89, n. 4, p. 583–592, out. 2005.

SINNOTT-ARMSTRONG, W.; YOUNG, L.; CUSHMAN, F. Moral intuitions as heuristics. In: DORIS, J. M. (Org.). *The moral psychology handbook*. 1. ed. Oxford ; New York: Oxford University Press, 2010. p. 246–272.

SKORUPSKI, J. The Definition of Morality. *Royal Institute of Philosophy Supplements*, v. 35, p. 121–144, set. 1993.

STERELNY, K. Moral nativism: A sceptical response. *Mind & Language*, v. 25, n. 3, p. 279–297, 2010.

STERELNY, K. *The evolved apprentice: how evolution made humans unique*. MIT Press, 2012.

WAHL, K. *Wie kommt die Moral in den Kopf?: Von der Werteerziehung zur Persönlichkeitsförderung*. Berlin: Springer Spektrum, 2014.

WARNEKEN, F.; TOMASELLO, M. Altruistic Helping in Human Infants and Young Chimpanzees. *Science*, PMID: 16513986, v. 311, n. 5765, p. 1301–1303, 3 mar. 2006.

WHEATLEY, T.; HAIDT, J. Hypnotic disgust makes moral judgments more severe. *Psychological Science*, PMID: 16181440, v. 16, n. 10, p. 780–784, 2005.

WIMMER, H.; PERNER, J. Beliefs about beliefs: representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, PMID: 6681741, v. 13, n. 1, p. 103–128, jan. 1983.

ZAJONC, R. B. Feeling and thinking: Preferences need no inferences. *American Psychologist*, v. 35, n. 2, p. 151–175, 1980.

Email: rogervargasrex@yahoo.com.br

Recebido: Setembro/2016
Aprovado: Dezembro/2016