ARQUIPÉLAGO

REVISTA DA UNIVERSIDADE DOS AÇORES

8

Filosofia

Ponta Delgada • 2007

RACIONALIDADE DA CIÊNCIA COGNITIVA À FILOSOFIA[1]

*Sofia Miguens**

«Alguns animais pensam e raciocinam; ponderam, testam e rejeitam hipóteses, agem em função de razões, por vezes após deliberarem, *imaginarem consequências e pesarem probabilidades*; eles têm desejos, esperanças e ódios, às vezes por boas razões. Também cometem erros no cálculo, agem contra o seu próprio melhor juízo, ou aceitam doutrinas com base em evidência não adequada. Qualquer um destes feitos, actividades, *acções ou erros é suficiente para mostrar* que um tal animal é um animal racional, pois ser um animal racional é simplesmente isso, ter atitudes proposicionais, não importa quão confusas, injustificadas, ou erróneas estas atitudes possam ser. Proponho que esta é a resposta. A pergunta é: que animais são racionais?» Donald Davidson, *Rational Animals*.

1. Introdução. Racionalidade. Racionalidade prática e racionalidade teórica. Irracionalidade.

O propósito do presente artigo é abordar as questões da racionalidade e da irracionalidade a partir de um projecto de investigação actualmente em curso na Universidade do Porto, cujo título é *Rationality, Belief, Desire*

* Universidade do Porto.

[1] Quero agradecer ao Departamento de Filosofia da Universidade dos Açores a oportunidade que o convite para a participação nas *VIIas Jornadas de Estudos Filosóficos* proporcionou de debater em público alguns dos materiais do Projecto de Investigação adiante referido. Agradeço especialmente ao Doutor Rui Sampaio o generoso acolhimento e a troca de ideias que marcaram a minha primeira visita aos Açores.

II – from cognitive science to philosophy[2]. Analisando conceitos e mapeando alguma da literatura sobre racionalidade pretende-se chegar a uma contribuição para o tema das VIIas Jornadas, *Racionalidade e Valores*.

Antes de mais, de que falamos quando falamos de racionalidade e irracionalidade, pelo menos no nosso projecto? Consideramos que falamos de *agentes*. E o que são agentes? Agentes são entidades dotadas de crenças e desejos, cuja acção é guiada por estes. Crenças e desejos são estado mentais representacionais, 'atitudes proposicionais', tais como por exemplo, respectivamente, 'Esta parede é amarela', '2 +2 são 4', e 'quero ler este livro'. A questão básica da racionalidade teórica – saber *em que devemos acreditar – aparece* quando perguntamos o que são crenças justificadas e o que pode e deve contar como razão para o agente passar a ter, ou deixar de ter determinadas crenças no seu corpo de crenças. A questão básica da racionalidade prática – o *que devemos fazer* – reporta-se à situação em que a partir das suas crenças e desejos os agentes deliberam sobre o curso de acção por que optar, por entre alternativas. A referência aqui é a Teoria da Decisão, uma teoria abstracta da escolha racional, em cujo núcleo estão precisamente agentes que deliberam, i.e. que a partir das suas crenças e preferências avaliam cursos de acção alternativos em termos de consequências. Quanto àquilo sobre que deliberam, já Aristóteles dizia, na *Ética a Nicómaco*, que pelo menos nós, os humanos, não deliberamos nem sobre 'coisas eternas' (como por exemplo verdades matemáticas), nem sobre o que não depende de nós (como termos um corpo de um determinado tipo, que por exemplo não nos permite voar), mas sobre aquilo que no futuro é indeterminado. De acordo com a teoria da decisão racional, um agente racional será aquele que, tendo em consideração a utilidade (a utilidade é uma unidade de 'desejabilidade') dos estados do mundo que podem ser consequência das suas acções e a probabilidade da obtenção desses estados, age de modo a maximizar a utilidade esperada (a 'utilidade esperada' será igual ao produto da utilidade do resultado pela probabilidade da ocorrência deste). Agir racionalmente é, assim, agir de forma a maximizar a utilidade esperada[3].

---

[2] O Projecto POCI/FIL/55555/2004, do qual sou Investigadora-Responsável, decorre no âmbito das actividades do Instituto de Filosofia da Universidade do Porto (Unidade I&D 502 da FCT).

[3] Ainda 'mais fundo' do que esta definição está a ideia de racionalidade instrumental, cuja natureza não vai ser aqui discutida, mas cujo teor é o seguinte: chamamos 'racional' ao agente que mobiliza os meios adequados em ordem à prossecução dos seus fins. Obviamente, no momento em que se avança esta consideração os agentes devem ser já, de alguma forma, dotado de fins.

Em ambos os casos, no caso da racionalidade teórica e no caso da racionalidade prática, existe espaço para a irracionalidade. São casos de irracionalidade teórica os casos em que acreditamos no que não temos razões para acreditar, não acreditamos no que temos razões para acreditar, etc. São casos de irracionalidade prática os casos em que não fazemos o que temos boas razões para fazer e que de acordo com o nosso melhor juízo queremos fazer, os casos em que fazemos aquilo que acreditamos que não devemos fazer, etc. Os exemplos de irracionalidade abundam nas nossas vidas mentais e incluem situações importantes para compreender a nossa natureza, tais como o auto-engano (caso em que escondemos de nós próprios que possuímos razões para acreditar que algo é o caso e não acreditamos) e a fraqueza da vontade ou *akrasia* (caso em que vemos o que seria melhor, e fazemos o pior, i.e. agimos intencionalmente de forma contrária ao nosso melhor juízo). Se saber o que é o 'bem', ou algum bem, para nós, nos movesse necessariamente, não haveria fraqueza da vontade.

Avanço estas observações acerca de irracionalidade porque devo confessar que compreender acusações de irracionalidade endereçadas a agentes, raciocínios, decisões, instituições, constituíu uma das principais motivações dos Projectos sobre racionalidade que temos vindo a levar a cabo.

Resta acrescentar que racionalidade e irracionalidade são estudadas quer na filosofia quer nas várias ciências cognitivas. Ora, à nossa abordagem subjaz a ideia segundo a qual para além dos problemas cognitivos da racionalidade (relativos a raciocínio, decisão, emoções, imputação de mente a outras partes do mundo por determinados agentes, etc), existe um problema especificamente filosófico da racionalidade. E é precisamente por aí que somos conduzidos à questão dos valores.

2. O problema filosófico da racionalidade: racionalidade e princípios, cânones de racionalidade e critérios de correcção.

De forma a substanciar a tese acima, vou introduzir alguma da terminologia que utilizámos no Projecto, bem como alguns exemplos. Falamos de cânones de racionalidade para falar da forma como a lógica, a teoria das probabilidades e a teoria da decisão nos oferecem explicitamente alguns princípios do pensamento e da acção. Mas o facto de determinados princípios normativos do raciocínio e da acção serem estipulados por lógicos, estatísticos e teóricos da decisão não significa que estes sejam necessariamente seguidos

por agentes reais nos seus afazeres cognitivos. Porque é que os agentes não seguiriam os princípios? Uma primeira hipótese é que esses princípios são idealizações que não capturam o funcionamento da racionalidade real (a arquitectura cognitiva dos agentes, o processamento da informação, a capacidade da memória, etc, não são consideradas na formulação de tais princípios, e agentes reais poderiam, por exemplo, desenvolver estratégias diferentes das ideais, precisamente porque não dispõem de recursos perfeitos). Uma outra hipótese tem a ver com a estrutura motivacional do agente, i.e. com a forma como crenças, desejos, intenções, emoções, razões se relacionam para conduzir o agente à acção (o presente artigo acabará por focar sobretudo esta questão). Mas voltemos aos exemplos para percebermos o que são cânones de racionalidade e critérios de correcção. Comecemos com o caso da Lógica

Num livro intitulado *Racionalidade*[4], que serviu de esboço programático para o projecto, utilizei como epígrafe um texto de Lewis Carroll, o autor de *Alice no País das Maravilhas*, que era também um lógico. O texto chama-se "What the Tortoise said to Achilles", foi publicado em 1895 na revista filosófica *Mind* e fala-se aí de Lógica. A lógica procura explicitar as regras pelas quais determinadas quais coisas pensadas / afirmadas, se seguem de outras coisas pensadas / afirmadas; é uma teoria formal, uma teoria da validade dedutiva de argumentos e supostamente diz-nos como pensamentos devem seguir-se a outros pensamentos. O texto diz o seguinte:

–Escrevi! – exclamou Aquiles alegremente, enquanto colocava o lápis no estojo. – E, por fim, chegamos ao final desta corrida ideal! Agora, que aceita A, B, C e D, EVIDENTEMENTE aceita Z.

–Sim? – retorquiu a Tartaruga inocentemente. Vamos deixar isso bem claro. Aceito A, B, C e D. Suponha que, ainda assim, me recuso a aceitar Z?

–Então a lógica pegá-lo-ia pelo pescoço e forçá-lo-ia a aceitar! – respondeu Aquiles triunfante. – A lógica dir-lhe-ia: «Não tem saída. Agora que aceitou A. B, C e D, TEM que aceitar Z!» portanto não tem escolha, não é?

–Qualquer lógica suficientemente boa para me convencer merece ser escrita – disse a Tartaruga – Portanto escreva aí no seu caderno por favor. Chamar-lhe-emos

(E) Se A, B, C e D são verdadeiros, Z tem de ser verdadeiro.

Enquanto não tiver aceite isso, evidentemente não terei de aceitar Z.

---

[4] MIGUENS 2004 (grande parte do material do presente texto provém deste livro).

Qual é o ponto do exemplo? É este: porque é que um determinado agentes cognitivo há-de seguir uma determinada regra da Lógica? Não podemos ter coisas para fazer e não as fazer? O que é isso de dever seguir uma determinada regra? Porquê exactamente essa regra? O texto chama a atenção para o facto de, se eu não tiver aceite a regra, nada me obrigar a seguila. Podemos pensar que a lógica explicita princípios do pensamento, mas o que é que isso significa? Como podemos justificar que devemos usar esses princípios, e que são esses mesmos os princípios a usar? Esta é a questão dos *critérios de correcção*. Poderíamos formular a questão assim: nós humanos, enquanto agentes, não apenas utilizamos princípios para raciocinar e decidir, mas temos que poder justificar por que os utilizamos, e por que utilizamos exactamente esses (para poder por exemplo dizer que uma pessoa está a raciocinar mal, ou a decidir incorrectamente). Há aqui outro ponto importante para a ciência cognitiva: são os pensamentos e não os cérebros que são movidos a argumentos. Na verdade, o ponto de agentes reais – no nosso caso dotados de cérebros –raciocinarem não é a correcção formal mas sim a sobrevivência no mundo. É isso que torna tão difícil em casos específicos que se tornaram célebres na literatura da ciência cognitiva, como o Caso 2 que veremos em seguida, afirmar taxativamente que um agente está a raciocinar mal. O problema é o seguinte: será que o agente está mesmo a raciocinar mal? Não estará a utilizar princípios diferentes?

O caso de Linda é um caso célebre na psicologia cognitiva, frequentemente referido a propósito dos estudos de um economista e de um psicólogo cognitivo, D. Kahneman e A. Tversky, sobre as heurísticas e os enviesamentos (*heurístics and biases*) característicos da nossa forma de pensar e decidir. O caso de Linda diz respeito à forma como pensamos sobre probabilidades, e é uma situação experimental em que uma determinada descrição é fornecida aos sujeitos. A descrição é esta:

'Linda tem 31 anos, é solteira, tem uma forte personalidade e é muito inteligente. Licenciou-se em filosofia. Enquanto estudante interessou-se profundamente por assuntos como a discriminação e a justiça social e também participou em manifestações anti-nuclear'.

Em seguida é feito este pedido aos sujeitos: 'Ordene as seguintes afirmações de acordo com a sua probabilidade, atribuindo o número 1 à mais provável e 8 à menos provável'.

a) Linda é professora primária.
b) Linda trabalha numa livraria e tem aulas de Ioga.
c) Linda participa activamente no movimento feminista.
d) Linda trabalha como psiquiatra na assistência social.
e) Linda é membro da Liga da Mulheres Votantes.
f) Linda é bancária.
g) Linda é agente de seguros.
h) Linda é bancária e participa activamente no movimento feminista.

Num grupo sem qualquer formação em probabilidade e estatística, 89% dos inquiridos indicaram que a afirmação (h) era mais provável que a afirmação (f). Quando a mesma questão é apresentada a sujeitos com preparação académica em estatística – por exemplo estudantes de pós-graduação do programa de ciência da decisão da Stanford Business School – uma percentagem muito próxima de sujeitos (85%) tem a mesma opinião. Sistematicamente acontece que os sujeitos consideram a afirmação 'Linda é bancária e activa no movimento feminista' como mais provável do que a afirmação 'Linda é bancária'. Ora, ninguém pode ser 'bancária e activista no movimento feminista' sem ser simplesmento 'bancária'. Quando se aprende teoria das probabilidades aprende-se que a probabilidade de uma conjunção não pode ser maior do que a de cada uma das proposições que a compõem. O raciocínio comum dos sujeitos é uma violação tão flagrante da teoria das probabilidades, que o facto de ocorrer sistematicamente quer em sujeitos não treinados em estatística quer em sujeitos treinados, se torna interessante. Podemos mesmo dizer que estas pessoas são todas irracionais?

O ponto dos exemplos é o seguinte. Temos um facto: as pessoas constantemente pensam e agem de formas que consideramos racionais e irracionais, e é relativamente simples identificar *cânones de racionalidade*: teorias como a lógica, a teoria das probabilidades, a teoria da decisão oferecem-nos tais cânones. Mas qual é a relação desses cânones formais com a forma como agentes reais pensam e decidem realmente? Os cânones explicitam princípios, e aparentemente nós usamos princípios para pensar e agir, mas se existe um problema filosófico da racionalidade este diz respeito precisamente à *justificação* dos princípios. Para nomear o que está aqui em causa, usamos no nosso projecto o termo que um epistemólogo americano, Alvin Goldman utiliza no livro *Epistemology and Cognition* (1986), o termo *critério de correcção*. Um critério de correcção explicita as razões por que consideramos determinados conjuntos de regras, normas, princípios, como

cânones da racionalidade. Formular um critério de correcção envolve dizer exactamente, por exemplo, por que razões consideramos que teorias como a lógica, a teoria das probabilidades, a teoria da decisão nos dão os princípios da racionalidade, da nossa racionalidade – e considerações cognitivas acerca de agentes reais terão que entrar aí.

3. Motivação para a acção. O que devemos fazer?- Kant e Hume, razão e paixões

O meu terceiro exemplo é a história de Medeia. Através dele aproximamo-nos daquilo de que o resto deste artigo trata (o resto do artigo trata basicamente de razão e paixões, ou mais propriamente, de teorias filosóficas da motivação para a acção, e da comparação das teorias de dois filósofos, D. Hume e I. Kant). Todos devem conhecer a história de Medeia, a personagem da famosa tragédia de Eurípedes com o mesmo nome. O marido, Jasão, abandona Medeia para assegurar o futuro dos dois filhos. Ela é banida de Corinto pelo rei Creonte. Retalia matando o rei, a filha, com quem Jasão tencionava casar, e finalmente mata também os seus próprios filhos. O que faz Medeia agir? Ela agiu de acordo com o seu melhor juízo? Foi possuída por algo? Estava fora de si? Estava divida por emoções em conflito (amor pelos filhos, ódio pelo marido, desejo de vingança)? Teve fraqueza da vontade, *akrasia* (via o que seria melhor fazer, e fazia o pior – na definição do filósofo americano D. Davidson)? Como havemos de lidar com um problema como este (a forma como razão e paixões intervêm na acção de um agente, ou o que é para agentes como nós, *querer*)?

Já sugeri que uma teoria filosófica da racionalidade deve considerar agentes cognitivos reais e não apenas teorias formais, e é para nos esclarecer acerca de agentes cognitivos reais que a ciência cognitiva, por exemplo com os estudos na psicologia, neurociência e IA sobre raciocínio, decisão, emoções, informação, é útil. Mas a ciência cognitiva não basta, ou não é o caminho mais rápido, quando se trata de conceber os agentes já funcionando globalmente – raciocinando, decidindo, de uma forma que envolve emoções, percepção, mais ou menos conhecimento disponível sobre o que é o caso no mundo, deliberação, e finalmente acção. Parece-me que a história da filosofia pode ajudar-nos aqui. Por essa razão vou apresentar a questão da motivação dos agentes (humanos) para a acção, sem a qual não poderemos abordar a racionalidade dos humanos, em termos de história da filosofia.

Falei até agora de agentes, das suas crenças e desejos. Nós somos agentes assim, a nossa vida mental é constituída por um grande número de crenças e de desejos. Mas ao descrevermos, por esboço, a nossa vida mental em termos de crenças e desejos, não dissemos ainda nada acerca do que nos move enquanto agentes. Noutras palavras, o que é que nos faz, a partir de todas as nossas crenças e desejos, querer num dado momento alguma coisa específica, e não outra, e agir em conformidade?

Filósofos, e também psiquiatras, falam de desejo ou desejos, psicólogos e neurocientistas falam frequentemente de motivação do comportamento. Em contextos de psicologia e neurociência, a motivação é ligada à avaliação e selecção de finalidades (*goals*) do comportamento e aos mecanismos cognitivos, neuroquímicos e hormonais subjacentes a esses processos. A motivação pode ser definida como uma influência moduladora e coordenadora sobre a direcção, vigor e composição do comportamento, com fontes internas, ambientais e sociais. Mas dizer isso não nos esclarece ainda acerca do querer do agente, do teor dos seus desejos, do que o faz decidir e agir. Vou tratar aqui a questão do que move o agente à acção nos termos da filosofia moral, sobretudo porque me interessa saber o que faz os agentes humanos prosseguirem objectivos simbólicos complexos ('valores'). A ideia básica é que uma resposta à questão *o que é que nos faz, a partir de todas as nossas crenças e desejos, querer num dado momento alguma coisa específica, e não outra, e agir em conformidade?* não pode ser dada por exemplo só por alguma história psicológica acerca de motivação, ou pela neurociência das emoções. Porquê? Porque envolve intencionalidade e normatividade, e uma concepção da natureza do agente que tenha estas questões em conta. Temos aqui um caso por excelência para observar a forma como as tarefas da ciência cognitiva e da filosofia se dividem e complementam: podemos perfeitamente admitir que para falar de iniciação dos comportamento de corpos no mundo temos que considerar questões relativas a neurónios motores e funções executivas e iniciadoras do córtex frontal, questões relativas à motivação do comportamento no sentido acima descrito, questões relativas base neurobiológica das emoções[5], etc. Mas podemos considerar ao

[5] As emoções podem ser definidas como estado ou processos psicológicos que funcionam na gestão de fins de agentes. Elas são tipicamente provocadas por avaliações de eventos como relevantes para dados fins: elas serão positivas quando se avança no caminho em ordem ao fim, negativas quando se é de alguma forma impedido de avançar; os agentes de que aqui se fala são animais, humanos e não-humanos. As emoções funcionam ainda como forma de 'colocar como prioritário' algum fim ou plano do agente relativamente à acção ou vida mental em geral. No seu sentido comum, que importa reter, 'emoção' cobre quer o senti-

mesmo tempo que isso não é suficiente para compreender a forma como as ideias de Deus, liberdade, ou verdade, nos podem motivar à acção. Entre a iniciação cortical de movimentos de um corpo e o ser movido à acção por ideias como Deus ou a verdade, ou o interesse próprio, está em causa a passagem do movimento corporal de um corpo humano e daquilo que o controla, à 'acção'. Ora acções não são apenas movimentos de corpos. São no mínimo movimentos descritos de forma mentalista, intencional, evocando crenças e desejos. Essa descrição deve ainda ser feita sem observação, devendo o agente ele próprio poder dar razões do que é feito (estou a pensar nas descrições clássicas da acção intencional feitas por Elizabeth Anscombe em *Intention (1957)*). As conotações mentais e éticas do termo 'acção' estão ausentes do discurso acerca de comportamento, controlo, funções motoras e executivas, etc. Mas o problema é que mesmo depois de se ter aceite alguma análise do que constitui acção e intenção, por contraste com meros movimentos de corpos humanos, não existe uma versão única do que motiva agentes humanos à acção, a particulares acções. Aliás, se a história da filosofia pode aqui ensinar-nos alguma coisa, o que ela nos ensina é que é possível desenvolver e defender de forma razoável concepções mais ou menos racionalistas da nossa estrutura motivacional, daquilo que nos motiva a agir. A oposição entre mais e menos racionalismo na forma de conceber a motivação para a acção pode ser simbolizada por I. Kant e D. Hume. Tanto Kant como Hume se preocuparam em saber o que move um agente à acção. No entanto, as suas posições acerca da natureza da mente levaram-nos a defender propostas totalmente diferentes. Isto não significa que Hume defenda que nós somos irracionais e Kant defenda que somos racionais: de facto, ambos desenvolvem uma teoria do que é sermos racionais.

---

mento interno quer a expressão comportamental. As emoções têm assim uma dupla face exprimir / sentir. Encontram-se teorias da natureza das emoções em autores que vão desde Aristóteles, Descartes e Espinosa a Charles Darwin e Sigmund Freud (estes dois últimos autores são especialmente importantes, historicamente, por terem considerado as emoções fundamentais naquilo a que hoje se chamaria a *mente naturalizada*). O trabalho neurológico e filosófico de António Damásio tem dado grande importância à questão das emoções, e precisamente, ao seu peso naquilo a que chamamos 'racionalidade'. Damásio distingue emoções primárias (medo / ansiedade, fúria / raiva, tristeza, felicidade, nojo – é esta de resto a lista comum nos manuais de psicologia) de emoções sociais tais como simpatia, compaixão, embaraço, vergonha, culpa, orgulho, ciúme, inveja, gratidão, admiração, espanto, indignação, desprezo. Em *O Erro de Descartes* (1994) Damásio analisa o papel das emoções no comportamento de decisão, em *O Sentimento de Si* (1999) analisa o papel das emoções (e sentimentos, que considera virem depois) na constituição da identidade própria (self, eu, identidade pessoal). Em *Ao Encontro de Espinosa* (2003) analisa directamente a natureza e significado de emoções e sentimentos.

Acontece que Hume e Kant defendem ideias profundamente diferentes acerca daquilo que motiva os humanos à acção. Hume e Kant constituem uma referência na literatura filosófica contemporânea. Podemos mesmo falar de duas linhagens, uma linhagem sentimentalista e uma linhagem racionalista, exemplificadas respectivamente por Hume e por Kant, na forma de conceber a motivação para a acção. Estas linhagens são perceptíveis nomeadamente na filosofia moral e na filosofia política.

De acordo com D. Hume, a razão é uma escrava das paixões (ou desejos), não tem por si própria qualquer força motivacional, só por si não nos move a fazer coisa alguma. Sem o suporte das paixões (desejos) a razão é incapaz de nos fazer preferir a ou b, ter um ou outro objectivo. É este o contexto da afirmação do *Tratado da Natureza Humana* (1739) segundo a qual não é contrário à razão eu preferir a destruição do mundo inteiro a um arranhãozinho no meu dedo (Livro II, Parte III, Secção III). É devido ao seu carácter inerte, à sua impotência para nos fazer tender para a ou b ou c, para nos dar fins ou objectivos, que a razão é e deve ser, segundo Hume, 'a escrava das paixões'; ela não deve nem pode pretender outro ofício senão o de servir e obedecer (*Tratado*, Livro II, Parte III, Secção III). Hume define razão como a descoberta da verdade e falsidade; estas consistem num acordo ou desacordo com a realidade. Representar que tal e tal é o caso, calcular as possíveis consequências de acções são incumbências da razão. Sob o título de paixões Hume analisa no *Tratado da Natureza Humana*, sobretudo no Livro II que lhes é precisamente dedicado, por exemplo o desejo, a aversão, o medo, o amor, o ódio, a tristeza, a alegria, a ambição, o orgulho, a humildade, a vaidade, a generosidade, etc. Não é a dimensão representativa da mente que nos conduz à acção, apenas as paixões ou desejos nos movem. Para compreender melhor o que Hume está a dizer acerca do querer humano, podemos comparar a sua posição com a imagem antitética, que pertence a um imaginário filosófico e psicológico intemporal, e que nos faz pensar na Razão como um cocheiro, que controla os desvairados 'cavalos' das paixões e dos impulsos irracionais. Esta imagem corresponde ao clássico dualismo acerca da nossa natureza mental a que podemos chamar 'Razão versus paixões' ou 'Apolo versus Diónisos'. Ora, de um ponto de vista humeano, quando temos aí uma pessoa a que queremos chamar sensata e razoável, não devemos concebê-la como funcionando (motivacionalmente, dinamicamente) da maneira retratada. À imagem do agente como um decisor racional, que controla as suas próprias paixões insensatas, Hume contrapõe a seguinte ideia: todos os trajectos motivacionais, todos os impulsos ou tendências do

querer, são determinados por paixões, preocupações ou cuidados. Noutras palavras, para querermos, para que o nosso querer tenda, temos que nos preocupar, importar, ser movidos. Ora, nós só podemos ser movidos por desejos. A razão ela própria não inclina a nada, apenas nos informa acerca das situações e permite prever resultados possíveis, mais ou menos prováveis, das nossas acções. De um ponto de vista humeano, um agente totalmente racional não faria em última análise absolutamente nada, não seria movido ou motivado para a acção. O que acontece pelo contrário com as pessoas, considerando casos não patológicos de pessoas que vão tendo tendência para agir[6], é que em qualquer situação em que vem a existir um querer conducente à acção existem paixões ou interesses que 'contrabalançam' outras paixões ou interesses, sendo que alguma vem a dominar as restantes, num dado momento, sem no entanto as extinguir. A própria ética, segundo Hume, funda-se em sentimentos e não em 'razão pura'. Assim, de um ponto de vista humeano, uma pessoa sensata ou razoável é assim aquela que é movida por paixões de alguma forma 'acertadas' e não a impossível pessoa 'movida puramente pela razão'.

Repare-se que a estas duas imagens antitéticas de pessoa sensata e racional, a da vitória e a do equilíbrio pontual, correspondem duas imagens muito diferentes das pessoas que estamos dispostos a considerar 'irracionais'. Na primeira imagem uma pessoa irracional é aquela na qual o condutor, i.e. a Razão, perdeu o controlo das paixões. Mas de um ponto de vista humeano não podemos sequer afirmar com boa consciência de uma pessoa que consideramos por exemplo maligna, malévola ou perversa que ela é pura e simplesmente irracional e que qualquer humano racional deve ver isso que nós vemos, ao considerar as suas acções. Apenas podemos afirmar que o curso de acção dessa pessoa é guiado pelo tipo errado de paixão. Mas isso é um defeito de sentimento, e não podemos simplesmente considerar que é um facto que essa pessoa não superou a sujeição aos (baixos...) desejos, o que lhe teria permitido auto-controlar-se racionalmente.

Por tudo isto Hume é classificado como um autor sentimentalista, por oposição a racionalista. Note-se que embora uma teoria humeana da natureza mental não nos impeça de criticar moralmente as outras pessoas, ela não nos permite fazê-lo evocando a Razão. Se seguirmos este ponto de vista, deveremos considerar que o que acontece quando emitimos juízos

---

[6] É evidentemente relevante aqui considerar que existem 'patologias do querer', como a depressão, a abulia, etc.

morais é exprimir aprovação ou reprovação. Por mais vilipendiada que a posição humeana seja, nomeadamente devido ao que parece ser a sua natural associação com o relativismo moral, é preciso notar que a tese central de Hume segundo a qual apenas uma paixão contraria, equilibra ou suporta outra paixão não é assim tão invulgar na história do pensamento ético. Por exemplo Espinosa (1632-1677), na *Ética*, defendeu algo de semelhante, nomeadamente que um afecto só pode ser removido se se lhe opuser outro afecto. Esta é uma boa razão para a 'simbólica' escolha, feita por A. Damásio, da figura de Espinosa por oposição à figura de Descartes, numa sua obra recente.

O sentimentalismo de tipo humeano não é a teoria da motivação para a acção mais conveniente em áreas da filosofia prática tais como a ética e a política. E há razões óbvias para isso. Por exemplo um filósofo como Kant, cuja influência nas teorias éticas e políticas é absolutamente incontornável - é possível considerar que a filosofia politica contemporânea tem por referência a obra de J. Rawls, e o pensamento de Rawls é profundamente marcado por Kant - defendeu uma teoria acerca da estrutura da motivação totalmente oposta à humeana e fê-lo por razões de peso, nomeadamente relacionadas com intenções a que poderíamos chamar democráticas. De facto, a teoria kantiana da motivação para a acção visa colocar todos os agentes como iguais entre si na sua motivação (racional) para a acção. Precisamente porque Kant é um autor central no pensamento ético e político contemporâneo vou tomá-lo como referência de uma concepção racionalista da motivação para a acção.

Na perspectiva de Kant temos então o seguinte: os desejos constituem em última análise uma interferência (indesejada...) na passagem do agente à acção, uma espécie de escudo psicológico incómodo que além do mais 'esconde' do próprio agente a natureza da sua (possivelmente racional) motivação para agir. A Razão é para Kant o único motor apropriado, correcto, da acção. É sua convicção que por princípio os seres humanos podem agir de forma totalmente racional (ou pelo menos espera-se que sim, embora a questão – saber se estão ou não a agir de forma totalmente racional – possa ser, em última análise, para os sujeitos humanos eles próprios, obscura). Gostaria que se partisse da seguinte 'intuição' que podemos atribuir a Kant: se se quer fundamentar o estatuto das leis e da obrigação à qual estas submetem o querer de agentes: coisas contaminadas e sensíveis como desejos não podem ser considerados como a fonte de todo o processo. A obrigação é um conceito moral fundamental para Kant, e Kant concebe-a como uma

motivação pela Razão, a Razão maiusculada, aquela que seria comum a todos os homens, e os faria querer voluntariamente a Lei, uma vez libertos dos desejos.

Mas por que razão necessitará Kant de afastar os desejos para ter uma boa teoria da motivação para a acção? A razão é que Kant considera todos os desejos como sendo, precisamente, 'sensíveis e contaminados' e fá-lo em grande parte porque os vê como sendo não apenas individuais mas também egoístas, i.e. desejos de prazer pessoal. Saber se isto é correcto, se de facto nós somos criaturas tais que todos os nossos desejos são desejos de prazer pessoal, saber se é o interesse próprio que nos move sempre, é fundamental se estamos interessados em compreender a estrutura do nosso querer (saber onde fica o eu e o interesse próprio é muito importante para compreender a estrutura do querer). Para Kant, os desejos são basicamente desejo de prazer pessoal e é precisamente isso que há de errado com eles. É por isso que para Kant uma coisa é a motivação por desejos, outra coisa totalmente diferente e muito mais importante, é a motivação pela Razão, que merece o nome mais nobre de 'Vontade'. Kant pensava que as leis, concebida como obrigação racional para o querer é, não podem ser criaturas do apetite, se estão lá para constrangê-lo. A Lei tem que ser independente da contingência dos desejos dos agentes. Assim, para Kant, não importa por exemplo que as inclinações empíricas dos agentes, os seus desejos, sejam generosos ou não, não importa a diferença entre o desejo de fazer algo de amável a alguém e o desejo de matar essa pessoa. Ambos são inclinações subjectivas, logo não podem ser fonte de Lei. Um sentimento que inclina (por exemplo que inclina à simpatia para com outrem, um caso que é tão importante para autores como Adam Smith (1723-1790) e David Hume) é uma mola motivacional de natureza totalmente diferente da boa mola motivacional que é a Razão.

Dada a divergência das suas concepções acerca de desejos e 'Razão', Hume e Kant apresentam-nos imagens muito diferentes da deliberação (vale a pena notar que se está aqui a analisar coisas muito diferentes daquelas que são consideradas por exemplo pela teoria da decisão racional – esta diz apenas que a partir do seu *ranking* de preferências, o agente delibera e escolhe de forma a maximizar a utilidade esperada). Mas o que é que está a acontecer no próprio agente? Imaginemos, uma pessoa, um agente que se prepara para agir. Segundo Hume aquilo que nela encontramos nesse momento são paixões e inclinações. Aquilo que determina a decisão é a resolução pontual dos conflitos sempre existentes entre desejos. A deliberação faz-se portanto a partir de ou por sobre este conflito de desejos: aqueles que

perdem podem ser definitivamente afastados, ou apenas provisoriamente silenciados. Consideremos agora a deliberação de um ponto de vista kantiano. Numa pessoa encontramos, como Hume também admite, os desejos. No entanto, para Kant, existe algo cuja natureza é totalmente diferente da natureza dos desejos: trata-se da Razão, que é livre e que está sempre pronta a impedir que as coisas corram mal por causa dos desejos. De facto, o ideal num agente kantiano puro seria que os desejos nem sequer existissem: a condição ideal do querer para Kant é a independência total relativamente a desejos e inclinações. Ser livre de desejos e inclinações é segundo Kant o desejo universal de todo o ser racional, já que desejos e inclinações são um peso, uma escravidão, um afastamento do agente relativamente à sua verdadeira natureza. O que é importante na determinação do querer é a Razão, que legisla. Esta legislação é a 'fonte da normatividade', para usar a expressão de Christine Koorsgaard, uma filósofa de Harvard. O agente kantiano é, assim, uma imagem do auto-controlo racional absoluto. De um ponto de vista kantiano, o nosso destino moral está exclusivamente nas mãos da Razão. Só a rectidão, a capacidade de agir segundo princípios, conta. *No feelings* é a 'lei', aliás, é a única situação que possibilita a Lei. Além do mais, só essa situação pode ser 'democrática', só ela pode tornar-nos iguais uns aos outros, livres de inclinações subjectivas. Estranhamente isso significa, se se pensar duas vezes, que inclinações parciais da vida moral tais como o amor ou a amizade dificilmente encontram lugar na imagem racionalista kantiana do querer. Mas isto significa por outro lado que para Kant a moralidade pode ser fundada na Razão, o que a des-subjectiviza, des-sensibiliza, universaliza, tornando-a propícia a sustentar a Lei de uma forma que a 'mera psicologia' seria – pelo menos do ponto de vista de um kantiano – incapaz de assegurar. Só para terminar esta passagem pela história da filosofia queria referir a diferença que estas questões fazem quando se passa da motivação individual para a acção a concepções de sociabilidade e de coordenação das acções de indivíduos, e portanto a teorias da sociedade.

O que é que a tradição humeana em filosofia moral tem a contrapropor à forma racionalista de conceber a motivação para a acção, e, especificamente, à forma racionalista de conceber a obrigação e a Lei como 'motivadoras' para agentes morais humanos e laço da coesão social? Porque é que agentes humanos se sustentam em sociedade? Um dos bons analistas desta questão é Simon Blackburn, um filósofo neo-humeano de Cambridge[7]. De

[7] BLACBURN 1998.

acordo com Blackburn, a tradição humeana em filosofia moral envolve antes de mais uma ideia razoavelmente optimista da natureza humana. Nomeadamente, nessa tradição parte-se do princípio de que existe nos humanos um mecanismo de simpatia natural, pelo qual as dores e alegrias dos outros 'ressoam' em cada um de nós, e em função do qual naturalmente elogiamos e promovemos acções que promovem prazeres e evitam dores. Mesmo que este mecanismo funcione antes de mais em grupos muito restritos, por exemplo ligados por laços de parentesco, ele é um primeiro indício do tipo de laços que de alguma forma virão a sustentar sociedades. A ideia é que a partir de um mínimo de virtudes naturais serão possíveis numa sociedade virtudes muito mais artificiais (para falar a linguagem de Hume), por exemplo a justiça, sem que seja necessário vir a falar de algum pacto ou contrato racional explícito. Hume, nomeadamente, vê o referido mecanismo de simpatia em funcionamento na forma como os humanos consideram os sentimentos dos outros nos juízos que fazem acerca de si próprios, chegando a afirmar no *Tratado* que as mentes dos homens são como 'espelhos' umas das doutras. É este mecanismo que virá a possibilitar a coordenação entre as acções dos indivíduos através de sentimentos morais (tais como a vergonha, a culpa, o desejo de ser louvado e admirado, etc), sem evocações racionalistas e calculadoras de um contrato. Hume exprime esta situação no *Tratado da Natureza Humana* (Livro II, Parte II, Secção II) e na *Investigação acerca dos Princípios da Moral* (Ap. III) com a imagem dos remadores que remam coordenadamente por interesse comum, sem qualquer promessa ou contrato explícito entre si (Hume considera por exemplo que é assim que a propriedade é respeitada, que o ouro e a prata são feitos medidas de troca e que os significados das palavras da linguagem são fixadas).

Adam Smith, por sua vez, na sua obra sobre moral e paixões (*The Theory of Moral Sentiments*, 1759), defende as seguintes ideias acerca da natureza e funcionamento dos sentimentos morais. Existe nos humanos uma voz interior, um homem dentro do peito, que os 'controla' e faz fazer (aquilo a que normalmente se chama a 'consciência moral'). Ora esse escrutinador das acções individuais, esse espectador imparcial como lhe chama Smith, não tem nenhuma origem sobrenatural ou divina. O homem dentro do peito não é nenhum estranho: a sua jurisdição, o seu poder de nos fazer fazer, funda-se no desejo que os indivíduos têm de serem louvados, na sua aversão a serem inculpados e recriminados, no seu desejo de possuírem as qualidades que admiram nos outros e no seu temor de possuírem os traços que neles desprezam. O observador dentro de cada um, supostamente estrutural em

todos os indivíduos humanos, está assim ligado à capacidade de ver a acção própria sob uma luz 'geral', i.e. como ilustrando determinadas qualidades que quando encontradas nas outras pessoas despertam amor, admiração, desprezo, ódio, etc. Repare-se que o que está aqui em potência é um mecanismo de ajustamento mútuo das acções através de sentimentos morais individuais. Isto significa por exemplo que sentimentos morais tais como a culpa e a vergonha de alguma forma servem para coordenar os humanos entre si, na medida em que ligam as nossas atitudes para connosco com aquilo que os outros pensam de nós. Simon Blackburn chama a atenção no seu livro *Ruling Passions* para a forma como todo este 'mecanismo' põe em relevo algo de essencial para os assuntos morais e políticos, algo que já Hume, antes de Smith, chamava a nossa capacidade de assumir o 'ponto de vista comum'. O que é esse ponto de vista comum? Trata-se de algo de diferente da razão kantiana, e para explicar de que se trata, Hume chama a atenção para o contraste entre avaliações feitas em função do interesse próprio (em que por exemplo considero alguém meu inimigo) e avaliações nas quais espero que os outros concorram (como quando considero essa mesma pessoa viciosa ou depravada). Neste último caso há uma espécie de 'subida' para o ponto de vista comum, que se relaciona com a capacidade de avaliar de uma maneira imparcial. Esta capacidade não é aliás específica da ética (também podemos ser capazes de julgar que um homem que detestamos tem uma magnífica voz ou é um grande pintor, ou um hábil advogado). A ideia de auto-escrutínio, comum a Hume e a Smith e fundamental para despoletar este mecanismo de coordenação, não é, nota Blackburn, de todo nova, e o próprio Hume aponta como predecessores os Estóicos, Cícero e Calvino. O que é inovador na tradição Hume-Smith é mais propriamente a ligação entre o auto-escrutínio, a internalização do olhar dos outros como mecanismo de ajustamento dos indivíduos entre si através dos sentimentos morais e a sustentação da sociedade enquanto sociedade.

4. Conclusão: o que procuramos quando procuramos uma teoria da racionalidade.

Não vou pronunciar-me aqui a favor de Hume ou a favor de Kant no que diz respeito à motivação dos agentes para a acção; o meu propósito era apenas mapear conceitos e problemas em torno de racionalidade e irracionalidade de agentes, e a partir de Hume e Kant expor uma alternativa acerca

do que constitui racionalidade na acção. Mas queria ainda deixar uma ideia acerca das relações entre filosofia e ciência cognitiva no que respeita à questão da racionalidade. Falei de Hume e Kant a propósito da motivação dos agentes para a acção porque não me parece que o campo da filosofia moral, que é o campo em que Hume e Kant se 'enfrentam', esteja 'tomado' pela ciência cognitiva. Esta é uma tarefa para filósofos, é parte de uma teoria filosófica da racionalidade, e parece-me essencial para pensarmos em questões como a que está aqui a ser discutida da racionalidade e dos valores.

Só para terminar recapitulando e organizando: O que nos satisfaria então como teoria filosófica da racionalidade? Penso que ela deve envolver pelo menos: (i) uma *descrição ou caracterização dos factores em jogo nas ocasiões em que agentes passam de determinadas crenças para outras crenças, adicionam ou eliminam crenças do seu corpo de crenças, ou optam, a partir de um conjunto de crenças e desejos, por um curso de acção por entre várias alternativas*, (ii) *um conjunto de hipóteses acerca da forma como decidimos entre critérios de correcção quando falamos da justificação ou racionalidade de crenças e acções*, (iii) *um conjunto de hipóteses acerca das razões por que queremos saber (se de facto queremos) se as nossas crenças são verdadeiras e os nossos raciocínios e acções racionais.*

Penso que a questão de fundo destas Jornadas, a questão da racionalidade e dos valores, faz parte, pela sua própria natureza, de tal investigação.

REFERÊNCIAS:

Anscombe, *Intention*, 2000 (ed. orig. 1957), Cambridge MA, Harvard University Press.

Blackburn, Simon, 1998, *Ruling Passions – a theory of practical reasoning*, Oxford, Oxford University Press.

Carroll, Lewis, 1895, «What the tortoise said to Achilles», *Mind*, 4, pp.278-280.

Goldman, Alvin, 1986, *Epistemology and Cognition*, Cambridge MA; Harvard University Press.

Hume, David, 2001, *Tratado da Natureza Humana*, Lisboa, Fundação Calouste Gulbenkian (tradução: Serafim da Silva Fontes; prefácio e revisão técnica: João Paulo Monteiro).

Miguens, Sofia, 2004, *Racionalidade*, Porto, Campo das Letras.

Smith, Adam, 1976, *The Theory of Moral Sentiments*, Oxford, Oxford University Press.