

O AUTO-DEBATE É POSSÍVEL? DISSOLVENDO ALGUNS DE SEUS SUPOSTOS PARADOXOS*

MARCELO DASCAL

Tel Aviv University
Philosophy Department
P.O.B. 39040, Ramat Aviv
TEL-AVIV 69978
ISRAEL

www.tau.ac.il/humanities/philos/dascal – dascal@post.tau.ac.il

Resumo: O debate consigo mesmo é um fenômeno corriqueiro. Diariamente tomamos decisões – sejam elas importantes ou triviais, teóricas ou práticas – em questões nas quais temos que escolher entre pelo menos duas opções. Para fazê-lo confrontamos uma com a outra seja deliberando pausadamente a respeito dos méritos de cada uma, seja impulsivamente adotando uma delas e descartando as demais. Os auto-debates que mais têm chamado a atenção dos filósofos são aqueles em que pareceria que a racionalidade é violada: do *wishful thinking* ao auto-engano, passando pela *akrasia* (debilidade da vontade) e outras formas aparentemente paradoxais do pensar e do agir. Servindo-se de concepções da racionalidade e do debate desenvolvidas pelo autor, este artigo procura analisar a natureza do auto-debate e seus tipos, e mostrar que a irracionalidade que se costuma atribuir a alguns de seus exemplos típicos só se justifica devido a uma concepção que limita a racionalidade a seu modelo lógico-dedutivo.

Palavras-chave: Auto-debate. *Akrasia*. Vontade. *Wishful thinking*. Auto-engano. Controvérsia. Deliberação. Racionalidade forte. Racionalidade fraca.

IS SELF-DEBATE POSSIBLE? DISSOLVING SOME OF ITS APPARENT PARADOXES

Abstract: The debate with oneself is a common phenomenon. We make decisions every day – important or trivial, theoretical or practical – concerning issues in which we have to choose among at least two options. In order to do that we have to compare these options either by quietly reflecting about the merits of each one, or impulsively adopting one of them and discharging the others. The kinds of self-debate that have most attracted phi-

* Originally published in *Manuscrito*, v. 29, n. 2, p. 319-349, 2006.

losophers' attention are those in which rationality is apparently violated: from wishful thinking to self-deception, including *akrasia* (debility of will) and other apparently paradoxical forms of thinking and of acting. This paper tries to analyze the nature and different kinds of self-debate using the conceptions of rationality and of debate developed by the author, and tries to show that the irrationality that is usually attributed to some of its typical examples is only justified due to a conception that limits rationality to its logical-deductive model.

Key-words: Self-debate. *Akrasia*. Will. Wishful thinking. Self-deception. Controversy. Deliberation. Strong rationality. Weak rationality.

1. INTRODUÇÃO

Cada vez que é preciso escolher entre alternativas de qualquer tipo, nós ficamos em geral divididos entre as opções disponíveis e nos engajamos em um processo de decisão, no qual temos que ponderar sobre as vantagens e desvantagens de cada alternativa. Não é incomum representar esse processo como um julgamento íntimo como se cada alternativa “apelasse” para diferentes interesses ou motivos que residem dentro de nós, que se empenham, por assim dizer, em prevalecer e assim determinar nossa escolha. Quer tal processo se realize consciente ou inconscientemente, por meio de uma deliberação cuidadosa ou sob o impulso do momento, racionalmente ou instintivamente – algo semelhante a um debate parece se desenrolar no interior de nossas mentes entre pelo menos dois contendores (seja o que for o que eles possam efetivamente “ser”) que defendem posições opostas. E isso acontece dúzias de vezes ao dia em nossas vidas. Algumas vezes o debate interno é provocado por circunstâncias externas – completamente informais (ao procurar responder a questão de um estudante, ao escolher uma estrada em uma encruzilhada, ao resolver o que vestir no casamento de sua filha) ou mais formais (ao decidir sobre qual eletrodoméstico comprar, em quem votar, se assinar ou não um contrato, o que responder à pergunta crucial do sacerdote na cerimônia de casamento); outras, o auto-debate faz parte de um processo de pensamento puramente interno (como quando, refletindo, tratamos de verificar se um determinado argumento

efetivamente sustenta uma tese que estamos elaborando ou criticando); algumas vezes nós estamos plenamente conscientes do debate desenca-deado dentro de nós (tal como se passa nas situações formais que mencionamos) e outras, completamente inconscientes dele.

Independentemente de sua forma, esses episódios estão longe de serem marginais em nossas vidas mentais, pois seus resultados determinam em boa parte nosso comportamento, o curso de nossos pensamentos, nossas crenças e desejos, nossos projetos – tanto em questões triviais quanto em questões cruciais. Mas pouco se sabe sobre seu *modus operandi* e nenhuma explicação geral dessa família de eventos mentais existe, tanto quanto eu saiba. Os filósofos, que tradicionalmente têm se interessado por um pequeno conjunto de auto-debates – aqueles que são particularmente desconcertantes porque parecem envolver paradoxos insolúveis, irracionalidade ou conflitos irreconciliáveis entre “faculdades” mentais diferentes ou “partes do eu” – têm produzido uma literatura considerável sobre a fraqueza da vontade, o auto-engano e o eu-dividido. Os psicólogos têm investigado estes e alguns outros casos de auto-debate, como a chamada dissonância cognitiva e o processo de tomada de decisão. Neste artigo, que – além de fontes filosóficas e psicológicas – também se inspira de estudos sobre debates inter-pessoais e sociais,¹ eu procuro fornecer uma análise comparativa dos vários tipos de auto-debate mental que ajuda a dissolver pelo menos uma parte da aura de nebulosidade e de paradoxo que envolve alguns deles.

2. UMA PISTA ARISTOTÉLICA

Podemos começar seguindo uma pista fornecida por Aristóteles. Ele aplica o mesmo termo, *deliberação*, para o debate externo, inter-

¹ Uma comparação detalhada entre debates intra-pessoais e inter-pessoais pode ser encontrada em Dascal (2005b), texto sobre o qual a presente análise está baseada.

pessoal, que se realiza em uma assembléia no curso de uma decisão coletiva, e para o debate interno, intra-pessoal, que transcorre no íntimo de um indivíduo que procura determinar o curso ético apropriado da ação que deve empreender em circunstâncias determinadas.²

A escolha terminológica de Aristóteles não é casual, pois ela reflete uma analogia profunda entre suas concepções desses dois domínios e uma significativa sobreposição entre os dois textos.³ O discurso deliberativo – um dos três tipos de oratória considerados na *Retórica* – aconselha indivíduos ou assembléias sobre que tipo de ação escolher, dadas as circunstâncias. A deliberação – como considerada na *Ética a Nicômaco* – é o processo mental através do qual um indivíduo estabelece suas preferências e decide como proceder. Nos dois domínios o objetivo da deliberação é levar à persuasão racional em favor de uma das opções – no domínio inter-pessoal, à persuasão de outrem; no domínio intra-pessoal, à auto-persuasão. Em ambos os domínios, a deliberação se refere à escolha de uma ação futura dentro de um conjunto de ações possíveis – nenhuma das quais tem sua ocorrência assegurada. Por conseguinte não se está lidando com o necessário – que é o objeto da ciência – mas com o provável. O procedimento racional da deliberação, portanto, não é a de-

² Stuart Hampshire elabora a noção aristotélica de deliberação como pedra angular de sua teoria moral e utiliza a analogia entre o debate externo e o interno nesta passagem exemplar: “A imagem da mente que torna a noção de razão prática substancial é a imagem de uma câmara de conselheiros na qual os interesses dos agentes contrários são representados em torno da mesa, para que cada qual possa falar por si. O mediador, que representa a vontade, pesa os argumentos e a intensidade dos sentimentos por eles transmitidos e então emite uma ordem. A ordem expressa uma decisão e uma intenção de agir de certa forma, a ser executada, e representa a resposta do conselho ao debate nele transcorrido” (Hampshire 1991, p. 51).

³ Aristóteles considera a noção de deliberação suficientemente importante para escrever um tratado especial a ela devotado (*Peri Symboulías*, que está listado no catálogo de Diógenes Laertes). Sobre possíveis vestígios desse tratado na *Retórica*, ver a tradução espanhola e o comentário de Q. Racionero, pp. 99, 214, 222.

monstração, mas sim a avaliação das opções possíveis por meio da comparação entre as vantagens e desvantagens de cada uma delas. Consequentemente, a idéia de deliberar com sucesso comporta uma forma de correção, ainda que a deliberação não produza conhecimento ou verdade; trata-se, para Aristóteles, de uma “correção do pensamento”, que difere da “correção da demonstração” característica da ciência; Aristóteles chama esse processo de “calcular” e entende esse cálculo como algo típico da deliberação.⁴

As propriedades compartilhadas pelos dois tipos de deliberação, creio eu, indicam explicitamente que ela ocupa no esquema de Aristóteles – tanto na versão inter-pessoal da *Retórica* quanto na versão intra-pessoal da *Ética* – a posição que o que chamo de ‘controvérsia’ ocupa em minha tipologia dos debates inter-pessoais.⁵ Assim, a autoridade de Aristóteles talvez justifique (pelo menos heurísticamente) que se tome a deliberação intra-pessoal, tal como descrita por Aristóteles, enquanto exemplo por excelência de um tipo de auto-debate que pode ser chamado de “auto-controvérsia”. Um argumento por analogia parcialmente similar se aplicaria aos outros componentes das tipologias tripartites da ética e retórica aristotélicas. Mas não é necessário para os nossos propósitos aqui prover tal argumento. Uma vez que identificamos a deliberação como o tipo de auto-debate que ocupa a posição intermediária na tríade, podemos assumir que os outros dois tipos – “auto-discussão” e “auto-disputa” – são passíveis de serem definidos por sua respectiva “distância” da delibera-

⁴ Neste parágrafo, eu parafraseio e extrapolo afirmações de Aristóteles que se encontram em *Rhet.* 1357a-b, 1358b, 1359a; *Eth. Nic.* 1112a-1113a, 1139a-b, 1141a, 1142b.

⁵ Em minha tipologia sobre dos debates inter-pessoais proponho três “tipos ideais” (utilizando este conceito de Max Weber): “discussão”, “disputa” e “controvérsia”. Dada a exigüidade do espaço disponível, não é possível dar as definições desses conceitos aqui. O leitor poderá encontrá-las em outros escritos (e.g., Dascal 1998a, 1998b, 2000, 2005c).

ção, medida em termos de como as propriedades típicas desta diferem das propriedades principais da auto-discussão e da auto-disputa.

3. AUTO-ENGANO VS. DEBILIDADE DA VONTADE

A comparação entre auto-engano (*self-deception*) e debilidade da vontade (*akrasia*) é um bom ponto de partida. Análises filosóficas recentes sobre a debilidade da vontade e o auto-engano sublinham uma diferença profunda entre esses dois fenômenos (e.g., Davidson [1969], 1986, pp. 80-81; Searle 2001, pp. 234-236, 2005; e também Rorty 1986, ainda que em menor extensão). *Prima facie*, os dois fenômenos envolvem um tipo de irracionalidade. No entanto, a irracionalidade característica das análises da *akrasia* é de um tipo “ameno” (*soft*), se comparado à irracionalidade “forte” (*hard*) que caracteriza as análises do auto-engano. Mais precisamente, enquanto que um caso típico de debilidade da vontade ou incontinência (e.g., ter decidido fazer uma coisa e ao invés disso se ver fazendo outra coisa) não manifesta inconsistência formal alguma; o auto-engano, ao contrário, envolve a aceitação simultânea de uma crença que *p* e de uma crença que *não p*, o que é logicamente absurdo. Essa diferença no tipo de irracionalidade envolvida fornece a meu ver boa parte da explicação do porque a debilidade da vontade é “um problema menor para a psicologia filosófica” do que o auto-engano (Davidson), do porque a primeira é “menos perigosa” do que o último (Rorty), e do porque é “uma forma de irracionalidade comum e natural”, provindo da natureza mesma da ação racional (Searle). A razão para tal, em minha opinião, é que, sendo menos “forte”, a “irracionalidade” da *akrasia* pode ser tratada em termos de uma concepção plausível de racionalidade como uma forma de *razoabilidade*, enquanto que a irracionalidade forte atribuída ao auto-engano implica, ou a impossibilidade absoluta desse fenômeno, ou então, quando é considerado possível, a necessidade de uma divisão implausível do eu, às vezes em subsistemas independentes, ou no máximo em subsistemas conectados frouxamente.

4. RACIONALIDADE “FORTE” (*HARD*) VS. RACIONALIDADE “FRACA” (*SOFT*)

Por racionalidade ‘forte’ entendo uma concepção de racionalidade que tem a lógica clássica e suas aplicações como modelo fundamental. Esta concepção vê na inconsistência lógica a expressão paradigmática da irracionalidade e considera a certeza como principal objetivo e característica do conhecimento. Como a matemática é a implementação mais bem sucedida desse ideal de racionalidade, a racionalidade forte privilegia aquilo que é considerado como as principais razões desse sucesso. Consequentemente, tal concepção considera como condições da praxis e do pensamento racional, ou como suas manifestações preferidas, parâmetros como: a obediência irrestrita ao princípio de contradição; a utilização de definições precisas formuladas em termos de condições necessárias e suficientes; a argumentação visando a conclusões irrefutáveis, modelada sobre a dedução; a formalização desse procedimento pelo uso de sistemas simbólicos; a quantificação e a computabilidade; a axiomatização dos domínios do conhecimento; e demais parâmetros semelhantes.

Por racionalidade ‘fraca’ entendo, em um sentido amplo, uma concepção de racionalidade que procura levar em conta e desenvolver os meios para lidar com um grande número de situações – tanto teóricas como práticas – em que a incerteza e a imprecisão são a regra. Embora reconhecendo a aplicabilidade e utilidade dos altos padrões da racionalidade forte em certos campos, a concepção em questão se recusa a identificar como “irracional” tudo o que fica aquém desses padrões. Ela lida com a vastíssima área do “razoável”, que fica entre a racionalidade forte e o irracional. O modelo por trás da idéia de racionalidade fraca é o de uma balança em que razões a favor e contra (uma posição, uma teoria, uma ação, etc.) são comparadas e ponderadas, afim de tomar uma decisão, favorável ou contrária. Mas há uma profunda diferença entre “ponderar”

razões e “computar” razões.⁶ Com exceção de um punhado de casos, os pesos das razões não são precisamente quantificáveis e independentes do contexto. Conseqüentemente, “pesá-los” não assegura resultados conclusivos cuja negação implicaria em contradição. A “balança da razão”, contrariamente à dedução, “inclina sem necessitar” – usando a excelente frase de Leibniz.⁷ Não obstante, se a pesagem é apropriadamente realizada, a inclinação que pende para um dos pratos da balança oferece um guia razoável para a tomada de decisão.⁸ A lógica da racionalidade fraca é, portanto, não-monotônica e não pode ser reduzida ao padrão da lógica dedutiva. É a lógica das presunções que justificam racionalmente conclusões sem efetivamente demonstrá-las, da heurística para a solução de problemas e para a geração de hipóteses, da interpretação pragmática, da

⁶ “Rationes non sunt numerandas sed ponderandas” (as razões não estão para ser contadas, mas pesadas), enfatiza Leibniz em diversas ocasiões (esta citação, e.g., encontra-se na carta a Gabriel Wagner, 27.2.1697; in Leibniz 1875-1895, vol. 7, p. 571, traduzida ao inglês e comentada em Leibniz 2006). Para o uso onipresente da idéia e da metáfora de ‘balança da razão’ por Leibniz, ver Leibniz (2006) e Dascal (2005a).

⁷ Novamente uma frase frequentemente empregada por Leibniz. O *locus classicus* é o *Discours de Métaphysique* [1686], par. 13 (in Leibniz 1999, p. 1546).

⁸ Deve-se notar que, estando baseada sobre inferências não-dedutivas, a racionalidade fraca oferece apenas uma garantia frágil para suas recomendações, se comparada com a garantia sólida oferecida pela racionalidade forte. No entanto isso não significa que possa ser culpada de um erro cognitivo crônico, análogo ao da debilidade da vontade, denominado por Davidson ‘debilidade da garantia’ e definido do seguinte modo: “uma pessoa possui evidência a favor e contra uma hipótese; com base em toda evidência que tem, ela julga que a hipótese é mais provável do que improvável; apesar disso, essa pessoa não aceita a hipótese” (Davidson 1986, p. 81). Uma pessoa que se vale da racionalidade fraca pode, é claro, incorrer nesse erro, mas apenas se ela deixar de considerar as razões que considera relevantes, errar em atribuir-lhes os pesos apropriados, em compará-las adequadamente e em não admitir que o resultado possa ser superado por uma eventual apresentação de razões ‘mais fortes’.

negociação, e de inúmeros outros procedimentos dos quais fazemos uso na maioria das esferas de nossas vidas.

A lógica da racionalidade fraca é também a lógica da deliberação, um processo usualmente representado em termos do modelo da balança.⁹ É precisamente porque sua racionalidade é de uma variedade fraca que a deliberação pode ser associada com a “auto-controvérsia”, para a qual nenhum procedimento rígido de decisão existe. Aristóteles, como vimos, considera a deliberação como a pedra angular da tomada de decisão racional e da ação racional. É por meio da deliberação que se pode estabelecer preferências conforme as quais se decide agir – quando se procura agir racionalmente. Falhar a esse respeito é ser culpado de *akrasia*. Esse pecado é, portanto, um pecado contra a racionalidade. Lembremo-nos, no entanto, que deliberação para Aristóteles não é demonstração. A demonstração implica em necessidade, enquanto a deliberação prescinde da necessidade: onde as coisas necessariamente acontecem, não existe escolha, e por conseguinte não há lugar para a deliberação.¹⁰

A debilidade da vontade, portanto, pode ser censurada no máximo por cair em uma irracionalidade fraca e não em uma irracionalidade forte. É mais fácil explicar a *akrasia* desta forma do que invocando uma presumida irracionalidade forte, como no caso do auto-engano, e isso por duas razões. Primeiro, como a negação ou rejeição das conclusões das inferên-

⁹ Embora certamente relevante para a distinção entre as racionalidades fraca e forte, a ênfase de Hampshire (1991, p. 52) sobre a dependência de uma decisão deliberativa dos argumentos em seu apoio (em contraste à independência de um cálculo aritmético de argumentos para decidir se está certo ou errado) incide sobre o fato de que a decisão moral ou política é uma decisão para agir *por certas razões*, e não sobre o caráter fraco da inferência o que eu estou enfatizando aqui.

¹⁰ No mesmo viés, Leibniz escreve: “a natureza da vontade requer liberdade, a qual consiste no fato de que a ação voluntária é espontânea e deliberada, i.e., consiste em que ela exclui a necessidade que suprime a deliberação” (*Causa Dei* [1710], in Leibniz 1875-1890, vol. 6, p. 441).

cias “fracas” não implica em contradição, não é um grande escândalo que tal negação ou rejeição aconteça. Antes, é uma questão de rotina que não desafia (a menos que aconteça massivamente) as regras de inferência da racionalidade fraca. Segundo, como a razão principal para dividir o eu é evitar a inconsistência lógica, tal necessidade não se apresenta para o caso da irracionalidade fraca, pois esta no máximo apresenta formas de dissonância cognitiva menos agudas. Tais tipos de dissonância são corriqueiros e o eu costuma tratar deles sem desintegrar-se, da mesma forma que é capaz de considerar razões opostas no curso de um processo deliberativo sem ser, literalmente, despedaçado pela tensão gerada.

5. DEBILIDADE DA VONTADE E AS ‘LACUNAS’ DE SEARLE

A explicação de John Searle (2001, 2005) do caráter rotineiro da *akrasia* – ou seja, de algo não problemático do ponto de vista da psicologia filosófica – toma *prima facie* uma direção bem diferente. Ela se assenta sobre sua preocupação, profundamente anti-determinista, de preservar a liberdade de escolha, característica para ele essencial da ação humana voluntária. Como tal, a ação humana não pode ser, de seu ponto de vista, sujeita a coerção alguma, qualquer que seja sua origem – nem mesmo provindo esta da razão mais fortemente racional (p. ex., uma demonstração matemática) aduzida para executar a ação em questão.

Essa preocupação leva Searle a postular uma série de “lacunas” ou descontinuidades (*gaps*) que separam a deliberação da ação. Entre essas descontinuidades, as principais são as que separam entre a deliberação e a formação de uma intenção ou decisão para agir de acordo com seu resultado e entre a intenção e sua efetiva realização (Searle 2001, pp. 14-15). Essas lacunas são requeridas para que o agente humano possa exercitar sua liberdade. Se não fosse livre para decidir de modo diferente do que lhe fora “recomendado” tinha seu melhor juízo deliberativo ou para agir de modo diferente do que tinha decidido, o próprio juízo seria “causalmente suficiente” para produzir a decisão e esta para empreender a ação,

e ao agente caberia tão somente assistir passivamente ao desdobramento da cadeia causal da deliberação-decisão-ação independentemente de sua intervenção. Segundo Searle, o fato de que nós frequentemente experimentamos esse desdobramento como um suave fluxo não deveria, todavia, nos induzir a acreditar que não existe nele descontinuidade, nem que a intervenção do agente ativo a cada passo da cadeia deliberação-decisão-ação está ausente. Felizmente, o fenômeno da *akrasia* existe para nos lembrar disso. Ele é justamente um “sintoma” do que está acontecendo a cada momento em que nós agimos voluntariamente (Searle 2005, p. 76). A *akrasia* acontece rotineiramente e é fácil de explicar porque tudo o que faz é quebrar a aparência de suavidade, trazendo à luz o que sempre esteve presente de algum modo, embora escondido: as descontinuidades e o livre arbítrio, sendo este supostamente possível somente graças a elas.

O modo de argumentar de Searle ergue uma muralha entre o “Modelo Clássico” (como ele o denomina), que considera estar inteiramente equivocado, e sua própria teoria, na qual as lacunas são escavadas em profundidade e postas em evidência, estabelecendo abismos intransponíveis entre processos mentais como o são a deliberação, a intenção e a decisão. Em sua busca por um modelo formal da razão prática, ele considera exclusivamente o *modelo dedutivo*, e – sobre essa única base – conclui: “é impossível obter uma lógica formal de razões práticas que seja adequada aos fatos da psicologia filosófica” (Searle 2005, p. 56).

Essa atitude sugere uma desconcertante confiança na racionalidade forte como o único modelo e como a única ferramenta disponíveis para a explicação da ação racional, mesmo quando esse modelo parece ser obviamente inadequado para os próprios propósitos de Searle. Alguns aspectos de sua posição que vamos considerar a seguir poderiam ser significativamente beneficiados se fossem tratados nos termos da racionalidade fraca acima definida.

Consideremos o tratamento que Searle propõe para a *akrasia*. Ele apresenta assim a fórmula que resume sua forma básica:

(1) É melhor fazer A e eu decidi fazer A, *mas* eu estou, voluntária e intencionalmente, fazendo B (Searle 2001, p. 235).

Não é surpreendente que tal formulação esteja diretamente conectada com a teoria das lacunas. Segundo Searle, a simples existência da *akrasia* constitui um contra-exemplo decisivo contra os que não aceitam a tese segundo a qual a liberdade que demonstra o acrático quando não faz o que decidiu fazer ocorre *em todos os casos* de ação voluntária. Por conseguinte, teria que existir uma fórmula paralela para a ação racional não-acrática, isto é, aquela em que se faz o que se decide:

(2) É melhor fazer A e eu decidi fazer A *e* estou, voluntária e intencionalmente, fazendo A.

É fundamental, no entanto, notar a diferença reveladora (que enfatizo propositadamente) entre as duas fórmulas: *mas* na primeira vs. *e* na segunda. A interpretação da conjunção adversativa *mas* em geral requer uma delicada operação pragmática, pois o uso dessa conjunção raramente – para não dizer jamais – se refere à mera oposição semântica ou lógica entre dois membros da conjunção (ver Dascal 2003, Capítulo 6).

Para encurtar a história: casos de *akrasia* são, sem dúvida, contra-exemplos que nocauteiam a tese segundo a qual a deliberação e a decisão são causalmente suficientes para a execução da ação correspondente; entretanto, *per se* a ocorrência, ainda que freqüente, de tais casos evidentemente não elimina a possibilidade de que, estando de cabeça feita depois de uma deliberação, o agente seja mais propenso a agir de acordo com sua decisão do que a fazer o oposto; em outras palavras, (2) ao invés de (1) é a regra facultativa (*default*); i.e., espera-se que uma pessoa aja de acordo com (2), apesar de não ser *forçado* a agir assim mesmo tendo decidido deliberadamente fazê-lo, como o demonstra (1); isso explica o *mas* em (1): sua função consiste em indicar que acabar fazendo B em vez de A depois de ter decidido fazer A é uma razão suficientemente forte para cancelar, nas circunstâncias descritas, a expectativa expressa pela regra (2) de que A seria a ação desempenhada.

6. DEBILIDADE DA VONTADE ENQUANTO “RE-DELIBERAÇÃO”

A noção de *default* é uma noção fraca, e o que delineamos acima é uma explicação plausível, em termos de racionalidade fraca, da irracionalidade, supostamente forte, que tendemos a atribuir à ação acrática. De modo similar, a impressão de que há um leque infinito de possibilidade cada vez que nós fazemos uma escolha e que temos a liberdade de escolher qualquer uma delas, que Searle descreve como “o modo no qual a *akrasia* caracteristicamente emerge”,¹¹ remete ao fato de que a deliberação não é conclusiva, entre outras coisas por causa de sua sensibilidade a uma multiplicidade de fatores contextuais. Cada escolha possível está associada ao seu conjunto de razões relevantes e aos seus respectivos pesos, de tal maneira que a simples consideração de sua possibilidade está propensa a modificar a inclinação da balança, ao chamar a atenção sobre “outras bases” para a decisão.¹²

Metafisicamente, poder-se-ia perguntar se esse e muitos outros modos pelos quais fatores racionais “fracos” nos *inclinam* na direção de uma ação sem nos *compelir* a fazê-la, deveriam nos levar a considerar esses meios como um tipo inteiramente diferente de causação, por exemplo, como uma “causação intencional”. Se assim o fizéssemos, porém, estaríamos transformando aquilo que é observado como “uma margem de

¹¹ “[C]omo resultado da deliberação nós formamos uma intenção. Mas quando chega o momento de agir há uma gama indefinida de escolhas que se abrem para nós e várias delas são atraentes ou motivadas por razões distintas. Para muitas das ações que fazemos baseados em alguma razão, existem razões em conflito com aquela que nos levou a decidir fazer a ação em questão e que são favoráveis antes a escolher e fazer outra. Algumas vezes agimos em virtude destas outras razões e não de acordo com nossa intenção original. A solução para o problema da *akrasia* é tão simples quanto esta.” (Searle 2005, pp. 76-77)

¹² Isto pode explicar a insistência de Hampshire sobre a dependência de toda decisão deliberativa dos argumentos que a suportam (ver nota 1), ao mesmo tempo que permite a ele não restringir essa dependência aos casos nos quais a decisão é de *agir por determinadas razões*.

folga (*slack*) entre a deliberação e a formação da intenção... e entre a intenção e o empreendimento efetivo da ação” (Searle 2001, p. 231) em uma descontinuidade metafísica, ao invés de aceitar que se trata apenas da forma normal de como razões fracas exercitam seus poderes causais. Tendo tantos motivos para adotar o modelo da racionalidade fraca, é surpreendente que Searle não o tenha feito.¹³ Talvez isto resulte do fato dele simplesmente obliterar com suas próprias mãos a distinção entre a racionalidade fraca e a racionalidade forte, ao argumentar que mesmo a racionalidade mais forte – a da lógica dedutiva – não tem poder de obrigatoriedade: uma inferência pode ser válida, “mas não há nada que force qualquer ser humano a fazer tal inferência” (p. 21). Tal como o Deus de Descartes, que dispunha de uma liberdade situada acima das leis da lógica, a liberdade do agente racional segundo Searle parece ser absolutamente sem limites, pois razão alguma tem o poder seja de obrigá-lo seja de incliná-lo a agir de uma forma ou de outra.

Deixando a esfera metafísica, consideremos agora o acrático na fração de um segundo entre ele concluir uma deliberação com uma decisão e agir contrariamente a essa decisão. Precisamente nesse momento, ele é apenas um *acrático potencial*. Algumas razões contra sua decisão, que ele não havia posto na balança, agora aparecem de estalo em sua mente, talvez devido a uma objeção que vem de uma crítica externa. É aí que o auto-debate do acrático potencial é mais agudo. *Fazer ou não fazer?* Ele tinha decidido parar de fumar; mas ele não tinha levado em conta o peso da compulsão dolorosa por uma tragada em seu cachimbo, exatamente

¹³ A escolha, para a capa do livro *Racionalidade em Ação* de Searle, da obra prima de Vermeer, *Mulher com uma Balança na Mão* é significativa a este respeito. Neste quadro, pinta com sua costumeira precisão, uma cena de pesagem de jóias extremamente delicada. Quando adquiri o livro, pensei que esta escolha tinha sido motivada para indicar que, no texto, o modelo da balança – emblema da racionalidade *soft* – ocupava uma posição central. Esta, porém, está completamente ausente da teoria de Searle neste livro, sendo bloqueada pelo mecanismo *hard* da descontinuidade.

em um momento como esse, quando ele precisa de todos os poderes mentais direcionados para completar e entregar este texto *hoje*. Essa não é uma razão tão forte para reconsiderar a decisão sensata de parar de fumar, mas para admitir uma exceção, somente esta tragada, por favor. Já o tinham advertido quanto ao caráter insidioso de tais súplicas por “somente esta tragada” e tinha recebido instruções sobre como resistir a elas, persistindo na decisão não importa quão persuasivas as súplicas fossem. Obviamente essa pessoa já está engajada, nessa mesma fração de segundo, em uma nova rodada de deliberação que pode resultar em decisão e ação diferentes. Se isto acontece, então ela terá incorrido em *akrasia*.

Quão extensiva e cuidadosa a nova rodada de deliberação será, e o quanto isso pode influenciar seus resultados, depende de certos fatores, entre os quais o quanto a convicção prévia é sólida, o quanto a objeção à decisão prévia é levada a sério, em que medida aprendeu o método de resistir à “tentação”, o quanto o assunto é importante e urgente para o agente, etc. Depende ainda do caráter e do estilo da pessoa. Em uma carta a Leibniz, Thomas Burnett de Kemeny descreve uma atitude deliberativa extrema:

Mr. Locke é conhecido aqui como um homem que medita e pensa mais do que qualquer um sobre o que ele escreve, especialmente em filosofia. Ele considera e reconsidera, ele examina uma idéia de frente, de trás e transversalmente (por assim dizer), e ele vê uma coisa através de todos os pontos de vistas possíveis, de modo a ter sempre vantagem sobre um homem que não pensa tanto quanto ele.¹⁴

Quanto mais você examina um assunto por todos os lados, mais apto você está para questionar crenças, suposições, soluções e métodos arraigados – seus ou dos outros – ou seja, o auto-debate se torna mais parecido com a controvérsia. Isso pode, é claro, levar a uma suspensão indefinida do juízo, e será justificado somente para assuntos importantes que

¹⁴ Carta de Burnett a Leibniz, 3 de maio de 1697 (in Leibniz 1875-1895, vol. 3, p. 198).

não são prementes.¹⁵ A re-deliberação, no entanto, é mais comum do que poderia parecer. Seu resultado nem sempre leva à *akrasia*. Ele pode ser a re-confirmação da posição prévia, a *conversão* para a posição oposta, ou uma nova posição que emerge de uma *reconciliação* entre posições contendoras, capaz de acomodar uma nova configuração de pesos na balança. Em cada caso, a inclinação da balança pode se tornar mais forte ou mais fraca no decurso do processo de re-deliberação.

7. TIPOS DE CONVICÇÃO

Mas existem outros modos de reagir à objeção (interior ou exterior) e às ‘novas razões’ que ela apresenta. Se eu estou muito fortemente persuadido de minha decisão, sentindo-me seguro quanto à ela, não é provável que uma objeção – embora bem apoiada por razões convincentes – seja capaz de me fazer duvidar dela. Eu posso então desconsiderar inteiramente a objeção e evitar meter-me em um processo custoso de re-deliberação. Nesse caso, nenhuma auto-controvérsia será provocada e não há perigo de *akrasia*. Minha confiança virtualmente absoluta em minha decisão e sua conseqüente imunização frente às objeções pode ter sua origem em um dos seguintes cenários:

- (i) Não houve deliberação e eu tenho uma firme intuição, uma sensação instintiva de que é isso que eu tenho que fazer.
- (ii) Houve deliberação e – por princípio ou por hábito – eu não mudo de idéia após tomar uma decisão.
- (iii) Houve deliberação e eu estou seguro de que a decisão é correta porque a mesma foi alcançada pela aplicação de um método impecável.

¹⁵ “É desta maneira que os areópagos absolveriam, com efeito, um homem cujo caso fosse tão difícil para ser decidido; eles adiariam a decisão por um longo tempo e se reservariam um século para refletir a respeito” (Leibniz, *Nouveaux essais sur l’entendement humain* [1705], 2.21.23; in Leibniz 1875-1895, vol. 5, p. 167).

No cenário (i), não há deliberação antes da decisão porque eu nunca tive dúvida quanto a esta decisão ser a certa. Minha confiança está baseada na ‘evidência imediata’, proveniente de meu acesso privilegiado ao conhecimento de mim mesmo. Na medida em que não precisa de justificação (portanto, não é uma questão de deliberação) a decisão não admite objeções (portanto, é inquestionável). Já que os outros não têm o acesso privilegiado que eu tenho à fonte de minha evidência, eles estão *a priori* desqualificados para questionar minha certeza. Se as objeções dos outros à minha escolha sub-repticiamente “contagiam” minha mente, eu tenho que ser cuidadoso quanto a essa “voz” traiçoeiramente disfarçada como se fosse minha para negar-me a escutá-la seja qual for seu poder de atração. Se eu aparento discutir contra tal voz dentro de mim, não é porque eu admito que haja um espaço para um debate real sobre minha decisão; eu o faço apenas para propósitos erísticos, para me re-assegurar com uma “vitória” nesse pseudo-debate contra as objeções ao que eu, de qualquer modo, sei que está certo. Eu estou apenas lutando contra o canto das sereias, o estratagema destinado a me desviar do que eu estou seguro que tenho que fazer, utilizando para esse fim um contra-estratagema. Eu poderia também evitar o perigo considerando o “método da tenacidade” discutido por Pierce, isto é, recusando-me a me expor à crítica dos que podem desafiar meu auto-convencimento; fecho-me então em um círculo protegido e “mantenho, sistematicamente, distância de tudo o que possa provocar a mudança de [minhas] opiniões”.¹⁶ Em ambos os casos minhas manobras são exemplos de auto-engano. Aristóteles chama o tipo de pessoa descrita neste cenário de “impulsiva” ou ‘excitável’ e a classifica como alguém que não tem auto-controle, i.e., como um acrático potencial (*Eth. Nic.* 1141a20). Mas as decisões de tal pessoa estão longe de serem voláteis, pois elas só mudarão como o resul-

¹⁶ Em “A fixação da crença”. In: Peirce (1931-1958, vol. 5, p. 235; publicado inicialmente em *Popular Science Monthly*, 12, pp. 1-15, 1877). Para uma discussão crítica ao artigo de Peirce, ver Dascal e Dascal (2004).

tado de um choque frontal com uma poderosa sensação instintiva contrária. Na ausência de tal choque, a pessoa irá se agarrar ferrenhamente à sua decisão, não importa qual seja.

No cenário (ii), a pessoa parece ser o homem capaz de conter-se que está em pleno auto-controle. Ele deliberou e chegou à decisão que se compromete consigo mesmo a realizar e que acata. A diferença entre eles reside no fato de que a pessoa do primeiro tipo é “difícil de convencer e não é facilmente persuadida a mudar suas convicções”, enquanto o homem contido está “disposto, ocasionalmente, a ceder à persuasão”. O segundo resiste à paixão e cede à razão, enquanto que o primeiro resiste à razão. Aristóteles chama as pessoas do segundo cenário de ‘obstinadas’ e lista entre os motivos de sua teimosia “a sensação agradável de vitória em não ser persuadido a mudar de idéia e o desconforto de ter os decretos de sua vontade e desejo soberanos anulados”.¹⁷ Para essas pessoas a decisão sinaliza o ponto final do processo deliberativo. Re-abrir esse processo e aceitar qualquer modificação de seus resultados sinalizaria para eles auto-derrota porque o eu admitiria desse modo sérias imperfeições tanto em sua razão quanto em sua vontade – no desempenho da escolha racional e na capacidade de acatá-la.

No cenário (iii), à diferença de (ii), a base da convicção da pessoa não é um orgulho pessoal, ávido por defender *post factum* a decisão feita, mesmo quando a pessoa se dá conta de que a decisão se choca com a razão. Sua auto-segurança é, antes, a certeza de que o método de deliberação utilizado é apropriado à questão em pauta, rigorosamente conclusivo, e foi corretamente aplicado – de modo que a decisão alcançada não admite qualquer revisão sobre uma base racional. A pessoa pode eventualmente admitir que, apesar de ter verificado cuidadosamente os resultados, por si mesma e/ou com a ajuda de colegas competentes, resta alguma possibilidade mínima de erro. Esse pode ser um erro humano ou uma falha não percebida no procedimento. Uma dúvida sobre os resultados,

¹⁷ As citações neste parágrafo são de *Eth. Nic.* 1151b4-15.

provocada por questionamento interno ou externo, é uma boa ocasião para verificar a correção do procedimento, seu modo de aplicação ou outros possíveis problemas. Isso abre espaço para o que pode parecer uma re-deliberação. Mas a única maneira que esse tipo de pessoa admite para persuadir-se a abandonar sua convicção é ser colocada diante do fato de que um ou mais *erros* no sentido estrito foram cometidos por ela – isto é, erros que poderiam ser corrigidos pela reaplicação do procedimento (sempre que possível, por pesquisadores independentes). Uma vez que a verdade é determinada por tal re-calculagem, nenhuma razão para a dúvida pode permanecer, e tal pessoa, de seu próprio acordo, é obrigada a aceitar o resultado (qualquer que ele seja) e tem nisso o mesmo grau de confiança que ela tinha no resultado prévio – quer este seja confirmado ou corrigido. Ao contrário do cenário (ii), abandonar sua convicção prévia desse modo não é percebido no cenário (iii) como derrota, mas antes como uma vitória do método que se adotou e como uma prova mais de que se está na trilha certa.

8. CONVICÇÃO VS. DELIBERAÇÃO

Cada uma das fontes de convicção em (i), (ii), e (iii) diferem em seu posicionamento frente à deliberação; mas todas elas compartilham sua repugnância por ela. Cada uma tem suas próprias razões para dispensar a deliberação no auto-debate, mas todas estão persuadidas que elas podem operar melhor do que a deliberação. A deliberação é repugnante para elas por causa do tempo que consome e da energia que sua execução apropriada demanda. Frente a isso, cada um dos três tipos de convicção nos cenários examinados supostamente oferece um meio finito e capaz de economizar tempo e energia mental para finalizar qualquer auto-debate. Ademais, a deliberação desperta repugnância nas estratégias baseadas na convicção por sua natureza aberta. Ela pode começar por deliberar sobre um tópico bem delimitado e acabar discutindo o universo inteiro, em razão da ânsia de não desconsiderar, no processo de pesagem,

quaisquer razões relevantes. Em oposição a isso a convicção mantém a discussão dentro do compartimento hermético no qual ela é mais efetiva, e assim evita a disseminação perigosa da dúvida nos assuntos filosóficos e metodológicos gerais, cujas fronteiras não se podem determinar de antemão. E acima de tudo, o grande inimigo da convicção é visto, corretamente, como sendo a racionalidade fraca característica da deliberação. Do ponto de vista da convicção, é possível superar esse inimigo através da adesão ao método da racionalidade dura (cenário iii) ou através da recusa categórica de questionar uma decisão, ainda que esta tenha se realizado pela via da deliberação fraca cujos resultados podem *em princípio* ser questionados (cenário ii), ou ainda abrindo mão, completamente, do pressuposto de que as decisões e crenças confiáveis são aquelas que se baseiam nos processos racionais (fracos ou duros) de tomada de decisão e substituindo-os por intuições, sentimentos e impulsos inexplicáveis (cenário i).

9. O EU PODE ENGANAR A SI MESMO?

Para completar o círculo de nossa investigação sobre os tipos de auto-debate, retomemos agora o auto-engano que, embora talvez não seja um fenômeno tão rotineiro quanto a debilidade da vontade, é certamente conhecido. O auto-engano parece ser uma jogada útil na “auto-disputa” – o tipo mais implacável de auto-debate no qual tudo o que importa é a vitória. Apesar de ser aparentemente fácil concebê-lo nesses termos, não é de todo evidente que o auto-engano possa, de fato, ser concebido como uma jogada no auto-debate.

Sob uma análise mais atenta, o auto-engano não consiste somente em uma situação na qual, de algum modo, alguém acredita, ao mesmo tempo, que uma proposição é verdadeira e que a mesma proposição é falsa; ele é, além disso, uma situação na qual – nos casos mais flagrantes – uma das crenças contraditórias do agente funciona como fator causal que instaura a outra crença e até mesmo a “sustenta” (Davidson 1986, p. 89).

O auto-engano, portanto, manifesta uma forma flagrante de inconsistência, uma irracionalidade tão aberrante que levaria a perguntar se *um único* “eu” poderia ser seu *locus*, e nesse caso, que tipo de eu deveria ser para ter tal capacidade. Não é surpreendente que explicações filosóficas recentes do auto-engano (e.g., Davidson 1986; Rorty 1986; Pears 1986; Searle 2001, pp. 235-236) tenham chegado à conclusão de que, para que o auto-engano seja possível é necessário aceitar algum tipo de divisão do eu.

Suponhamos, para fins da argumentação que se segue, que o eu seja dividido em, no mínimo, dois sistemas – sejam eles permanentes (como para Pears e Searle), tais que evoluem e variem em sua configuração (como para Rorty), ou que tenham seus limites traçados conforme as circunstâncias (como para Davidson) – sistemas esses que debatem entre si. Ora, uma vez que se divide o eu conforme qualquer um desses modos, chega-se a uma situação claramente equivalente à do debate externo. Isto porque se cada uma das crenças contraditórias *p* e *não-p* de um auto-enganador (ou de outro tipo de auto-debatedor) é mantida por *um* dos *dois* contendores, não existe inconsistência lógica em nenhuma deles *per se*, pois cada um mantém apenas uma dessas crenças.

Metonimicamente, a consciência de um dos contendores de que seu oponente sustenta uma crença que é inconsistente com a sua, sugere, é claro, um problema íntimo para ele: ele está certo e o oponente está errado ou vice-versa? Mas esse problema não consiste em descobrir uma inconsistência *no interior dele mesmo*, pois a situação não é tal que ele, João, acredita que *p* e *não-p* nem que João acredita que *p* e João acredita que *não-p*. Da mesma forma que no caso de um debate inter-pessoal cada crença oposta é mantida por um dos contendores, João ou Antônio, assim também no caso do debate intra-pessoal cada uma das crenças em jogo é mantida por um dos dois subsistemas de João, J_1 e J_2 , nenhum dos quais pode ser censurado por contradizer a *si mesmo*, nem, *a fortiori*, por incorrer em auto-engano. Auto-contradição – e auto-engano – ocorrem apenas se os dois subsistemas estão co-presentes e interagem a um nível de integração do eu no qual a contradição é explicitamente representada e, portan-

to, tem que ser tomada em conta e, se possível, resolvida. Na medida em que os subsistemas são “mini-eus” independentes que operam estritamente à parte e não efetivamente integrados em um eu comum, como parece ocorrer em casos de “múltipla personalidade” (tecnicamente denominada hoje em dia “Identidade Dissociativa”), não emerge nenhuma auto-contradição. O auto-engano pode existir, então, sem desafiar a consistência do eu, porque os *componentes* do eu que se encontram em conflito lógico estão, por assim dizer, distribuídos entre diferentes *constituintes* do eu.

Mas essa solução parece dar lugar a mais problemas do que resolver aqueles que se propunha solucionar. Primeiro, se o eu já não é unitário ou integrado, então qual é o objetivo de preservar “sua” consistência lógica ou “sua” racionalidade? Desistir de seu caráter integrado parece sugerir que se estaria *ipso facto* desistindo das exigências a serem preenchidas para que o eu fosse dotado de qualquer membro da família de “propriedades coerentes” à qual a consistência lógica certamente pertence. Talvez então o que os proponentes desta explicação estão tratando de preservar é a consistência de algum componente do eu que por alguma razão privilegiam. Mas qual deles e porque?

Sem dúvida, a consistência lógica de cada subsistema do eu dividido que participa do auto-debate é preservada, na hipótese que estamos explorando, graças à distribuição entre eles dos elementos logicamente discordantes entre si. Mas seguramente o objetivo do exercício não era preservar a consistência dos “mini-eus” postulados pela hipótese. Seria então talvez o objetivo dar conta da coerência de um componente do eu universalmente considerado de fundamental importância, por exemplo a “consciência”? Mas, visto que esse presumido componente de ordem superior do eu integra as produções conflituosas dos mini-eus e “está ciente” delas, a consciência explicitamente contém os elementos cognitivos em discórdia lógica e, portanto, a inconsistência a ser evitada pela manobra da divisão do eu reaparece, e com maior força. Em última análise, a solução proposta pela hipótese da divisão do eu falha clamorosa-

mente como forma de dar conta do auto-engano, porque a explicação que oferece é obtida à custa de privar de significado o “auto” em “auto-engano”.

Pela mesma razão essa solução descarta a possibilidade de considerar o auto-engano enquanto um exemplo genuíno de *auto-debate*, i.e., enquanto um debate que ocorre entre uma entidade e si mesma, uma entidade que argumenta agonisticamente consigo mesma, exercendo os papéis dos dois contendores – o de proponente e o de oponente. Dado o que sabemos sobre casos de auto-engano, se ele fosse um auto-debate genuíno, poderia certamente ser considerado como um excelente exemplo do tipo ideal ‘auto-disputa’. No entanto, na explicação acima discutida para o problema de sua irracionalidade, o auto-engano é de fato uma disputa entre subsistemas independentes, ou seja, uma disputa ‘inter-sistemas’, paralela a uma disputa inter-pessoas, para a qual o prefixo ‘auto’ nada acrescenta. Podemos efetivamente imaginar o auto-engano como uma disputa implacável entre duas entidades virtualmente autônomas (Dr. Jekyll e Mr. Hyde, talvez), cada qual absolutamente convencida de estar certa e agindo, por conseguinte, como se estivesse autorizada a lutar sem restrição por seus próprios interesses, valendo-se de qualquer estratégia enganador disponível para conseguir vencer, a qualquer preço, seu opositor. Neste cenário, o ‘eu’ está eximido de inconsistência ao não exercer outro papel nessa disputa além do de uma fina membrana recolhendo e reunindo sob seu manto unificador os subsistemas em conflito, manto que funciona como a arena neutra em que transcorre a batalha.

10. UM EU “LEVEMENTE-FRACIONADO”?

Obviamente o papel do eu deveria ser mais substancial do que o de um tênue container para servir de base para uma teoria realmente explicativa pelo menos dos “*auto-debates*” genuínos. Se os auto-debates efetivamente exercem – assim como eu acredito haver boas razões para acreditar que exercem – um papel central em nossa vida mental, o eu (e

seus componentes eventuais) não pode deixar de estar envolvido neles e ser assim fortemente moldado por eles. Mas é isso possível sem dividir o eu?

Talvez Davidson esteja certo ao afirmar que o auto-engano não pode realizar seu objetivo de induzir em nós uma crença que contraria o nosso melhor juízo servindo-se do tipo de causalidade que opera por meio de *razões*; isto porque “nada pode ser visto como uma boa razão para deixar de raciocinar segundo nosso melhor padrão de racionalidade” (Davidson 1986, p. 92).¹⁸ Se é assim, “o melhor padrão de racionalidade de alguém” funciona como um critério fundamental de demarcação, separando “razões” de todo o resto dos fatores que podem influir nas crenças (e outros estados mentais) de uma pessoa. A ênfase sobre a *inconsistência* que, segundo Davidson, é o que torna o auto-engano irracional, sugere que o critério em questão é um critério “forte”. Isto implica em excluir as influências “fracas”, não monotônicas, especialmente – mas não só – aquelas de caráter não cognitivo, tais como hábitos, gostos e inclinações. É certo que, segundo a distinção categorial de Davidson entre razões e causas, tais influências são causas e não razões. Mas para que possam se tornar membros do privilegiado clube das razões basta que cruzem o limiar sutil da consciência e que se possam “representar como razões internas”, empregando a linguagem de Searle (2001, p. 115), isto é, que os agentes sejam delas conscientes. Sabemos, todavia, que as representações mentais podem ser vagas, obscuras e comportar muitos elementos implícitos, além de moverem-se com facilidade de fora para dentro da consciência e vice-versa. O limiar proposto parece muito tênue e impreciso, portanto, para estabelecer uma demarcação forte – e ainda menos uma descontinuidade intransponível como o são as lacunas de Searle.

Existem outras razões para encarar a presumida divisão do eu como não do tipo *hard*, mas sim do tipo *soft*. Por exemplo, o fato de David-

¹⁸ Sobre a distinção entre ‘razões’ e ‘causas’, ver Davidson (1963).

son atribuir um caráter provisório e *ad hoc* ao fracionamento do eu em sua explicação do auto-engano, sugere que a divisão em questão é uma demarcação que, para ele, é “fraca” por natureza. Pense-se, por exemplo, no papel que o “princípio de total evidência”, uma peça central do “melhor modelo de racionalidade” segundo Davidson, exerce nessa demarcação. Suponhamos que esse princípio é enfraquecido, passando a exigir-se de quem age, que o faça sob a condição de *total evidência disponível para ele*. Como essa condição de disponibilidade da evidência poderia ser definida – conforme o próprio Davidson (1986, p. 81n) pergunta? Segundo ele, para que a irracionalidade do auto-engano seja possível é preciso que seja dado um passo *causal* (i.e., que não envolva razões) na geração do mesmo, e ele propõe que esse passo consiste no “exílio temporário ou isolamento” (das razões) da exigência da evidência total. Mas um princípio que pode ser suspenso temporariamente, admitindo assim exceções, não é na verdade uma regra presuntiva, cuja violação ocasional está prevista na própria regra e não é incompatível com a manutenção da regra? Obviamente tais violações rotineiras não criam quaisquer fronteiras ou discontinuidades permanentes. No máximo elas revelam a necessidade, pragmaticamente motivada, de divisões fracas e provisórias traçadas no interior do eu para tornar possível o entendimento de fenômenos – por exemplo o auto-engano – que de outra maneira nem mesmo a mais caridosa versão do princípio da caridade poderia salvar das garras da racionalidade forte.

Davidson não está só entre os que sugerem a possibilidade de manter alguma forma de partição do eu, porém mitigando-a. Outros são mais explícitos sobre essa possibilidade e menos preocupados quanto à eventual falha do fracionamento no seu suposto papel explanatório se a mitigação vai longe demais. Um caso em questão é o de George Ainslie (1982, 1986), que sugere um modo de abrandar o que, quem sabe, seja o mais complexo fracionamento do eu normal jamais proposto – a tríade freudiana *id*, *ego* e *superego*, com suas várias subdivisões. Ainslie distingue entre duas funções diferentes do que é usualmente reunido sob a

designação de “mecanismos de defesa do eu”: “(1) evitar as percepções desagradáveis per se” e “(2) evitar os impulsos, isto é, comportamentos que procuram recompensas que podem levar a percepções desagradáveis” (Ainslie 1982, p. 737). A primeira tarefa tem a ver com as percepções correntes cuja evitação gera recompensas imediatas; a segunda está relacionada aos interesses de longo alcance do eu e requer mecanismos capazes de abrir mão de uma recompensa imediata que possa levar a um sofrimento futuro em benefício de uma recompensa postergada. Mecanismos desse tipo incluem “dispositivos de comprometimento prévio”, que são “de longa vista, mas que devem reter o controle do comportamento a longo prazo se é que servem para obter as recompensas que, afinal, são sua *raison d'être*” (p. 746).

Ainslie analisa uma variedade de tais dispositivos, mas o mais interessante para nossos propósitos aqui é o ‘agrupamento de escolhas’ (*bunching of choices*). Esse agrupamento consiste em reunir escolhas conforme seu valor diagnóstico para quem escolhe, isto é, segundo se uma escolha é percebida como um “precedente para fazer futuras escolhas preferindo elementos do mesmo conjunto” ou agrupamento (p. 772). Em nossa terminologia, essa percepção atribui à escolha que tem um valor diagnóstico do tipo definido por Ainslie um peso suplementar no processo de deliberação. Suponhamos que eu tenha decidido, após uma cuidadosa deliberação, parar de fumar. Mas eu sinto uma compulsão intensa para fumar neste exato momento, aqui e agora. Se eu me rendo e violo minha regra auto-imposta, isso assinalará para mim que eu posso fazer o mesmo no futuro. Esse peso adicional tem que ser colocado na balança quando estou por decidir se vou ou não sucumbir à compulsão momentânea de fumar, porque ignorar esse peso pode me levar, contra meu melhor juízo, ao esfacelamento da regra que eu adotei ao seguir meu melhor juízo.

O “agrupamento de escolhas”, portanto, funciona como um mecanismo razoável – isto é, *soft* – de autocontrole que pode ajudar a evitar a incontinência. No entanto, ele pode ser durificado, transformando-se assim no mecanismo *hard* do comportamento compulsivo que proíbe

qualquer exceção à regra particular em jogo (parar de fumar), assim como a qualquer outra regra – que é precisamente a estratégia do superego (Elster 1986, p. 22). Esta estratégia coloca o ego em uma posição difícil: assim como ele tem que se defender da busca imediata do prazer pelo id, o ego tem também que se defender da imposição de regras duras impostas pelo superego. Para realizar essa defesa, de acordo com a teoria de Ainslie, é preciso apelar para um tipo de barganha interna que afrouxa a rigidez dos agrupamentos de escolhas, ao mesmo tempo em que não permite que a pessoa “cruze repetidamente os limites para fazer exceções” à vontade. Isto se obtém estabelecendo “linhas luminosas” que instruem o ego a respeito de quando as circunstâncias são propícias para justificar uma exceção (Ainslie 1982, p. 772). Independentemente de se esta explicação do conflito ego-superego leva à necessidade de admiti-los como entidades autônomas dentro do eu, ela sem dúvida mostra como esse conflito envolve vários dos tipos de auto-debate discutidos neste ensaio. Além disso, ela também mostra claramente como a racionalidade fraca funciona como um componente essencial para escapar dos dilemas e paradoxos que um fracionamento duro do eu gera.

Em todo caso, se conseguirmos “suavizar” de uma maneira ou de outra o fracionamento do eu, então as suas postuladas partes não serão, de fato, tão separadas como as personalidades “Dr. Jekyll” e “Mr. Hyde”. Ademais, o eu adquiriria nessas condições alguma substância própria. Primeiro, porque cada parte presumida não poderia mais ser constituída por um conjunto fixo de elementos, pois – como vimos – o que pertence em algum momento a um deles pode, em outro, tornar-se normalmente um membro de outro; ou seja, esses elementos se tornariam entidades multifuncionais que “pertencem” *de facto* ao eu e não a uma das suas supostas partes; eles são elementos que o eu pode *usar* para propósitos diferentes, assim como palavras podem ser usadas tanto para pensar quanto para comunicar; o fracionamento do eu passaria então a ser, por conseguinte, de caráter ‘onto-pragmático’ e não de caráter ontológico. Segundo, porque mesmo aqueles que insistem que uma fronteira estrita

separa o domínio das razões do das causas, como Davidson e Searle, em última análise admitem que essas partes do eu interagem *causalmente*, “uma vez que a razão não tem jurisdição do outro lado da fronteira” (Davidson 1986, p. 92). Juntas, a mobilidade dos elementos e a interação causal entre as presumidas partes do eu requerem bases compartilhadas suficientes para serem vistas como pertencentes a uma mesma “personalidade” substancial.¹⁹

Será que o abrandamento das linhas que fracionam o “múltiplo eu” acaba transformando a partição do eu uma “simples metáfora”, um simples “modo de dizer”, um mero “instrumento pragmático”, por não corresponder a algo “estável”, “incorporado”, psicológica ou ontologicamente “real” na pessoa? Para responder a essa questão teríamos que embarcar em uma nova investigação sobre o que permitiria descrever os subsistemas do eu (assim como o próprio eu) como “reais”. Por exemplo, seria sua identidade funcional suficiente para isso, ou se deveria exigir também correlatos cerebrais, módulos neurais independentes encarregados de desempenhar essas funções? Não por menosprezar as possíveis

¹⁹ Além disso, a possibilidade ‘trans-fronteiras’ da interação causal pode ser tudo que é requerido para que a intervenção de um fator ‘externo’, isto é, não envolvido diretamente na auto-disputa, possa ser causalmente efetiva para ‘dissolvê-la’ ou para reduzir seu nível de conflito, atuando assim como o fazem fatores externos como uma intervenção judicial nos conflitos inter-pessoais ou inter-grupais. Imaginemos algum tipo de mecanismo de auto-regulação como algo análogo a um observador ou o analista que interpreta os objetivos das partes em conflito e intervém na esperança de dissolver ou reduzir a tensão ou as dis-funções causadas pelo conflito, ou o análogo interno kantiano – esse árbitro ou o legislador interno que não toma partido por nenhum dos lados nas disputas metafísicas entre as ‘seitas dogmáticas’, mas apenas estuda esses combates com o fim de determinar os limites da razão pura (Kant 1971, Aviii, A423, B452; Dascal 2000, para mais detalhes). Especulações como esta sugerem que as partes do eu, ainda quando envolvidas em disputas acirradas, compartilham, no fim das contas, o interesse comum de preservar sua condição de funcionamento, um bem estar mínimo capaz normalmente de disciplinar até mesmo a auto-disputa.

contribuições da neurociência para reformular e talvez ajudar a resolver o velho problema psico-físico, devemos porém parar por aqui. Provisoriamente, podemos dizer que a resposta a esta e outras questões do gênero dependerá sem dúvida, entre outras coisas, de como encararemos a distinção entre o literal e o metafórico – como uma descontinuidade irreduzível ou como uma diferença gradual.

REFERÊNCIAS

- AINSLIE, G. “A behavioral economic approach to the defense mechanisms: Freud’s energy theory revisited”. *Social Science Information*, 21(6): 735-779, 1982.
- . “Beyond microeconomics. Conflict among interests in a multiple self as a determinant of value”. In: J. Elster (ed). *The Multiple Self*. Cambridge: Cambridge University Press, pp. 133-175, 1986.
- ARISTÓTELES. *Nicomachean Ethics*. Trad. H. Rackham, introd. S. Watt. Ware, Hertfordshire: Wordsworth, 1996.
- . *Retórica*. Trad., introd. e notas de Q. Racionero. Madrid: Gredos, 1990.
- DASCAL, M. “The study of controversies and the theory and history of science”. *Science in Context*, 11(2): 147-154, 1998a.
- . “Types of polemics and types of polemical moves”. In S. Čmejková, J. Hoffmannová, O. Müllerová, J. Světlá (eds.). *Dialogue Analysis VI* (= Proceedings of the 6th Conference, Prague 1996), vol. 1. Tübingen: Max Niemeyer, pp. 15-33, 1998b.
- . “Controversies and epistemology”. In: T.Y. Cao (ed.). *Philosophy of Science* (= Vol. 10 of *Proceedings of the Twentieth World Congress of Philosophy*). Philadelphia: Philosophers Index Inc., pp. 159-192, 2000.

- . *Interpretation and Understanding*. Amsterdam: John Benjamins, 2003. [Trad. brasileira: *Interpretação e Compreensão*. São Leopoldo: Editora Unisinos, 2007].
- . “The balance of reason”. In: D. Vanderveken (ed.). *Logic, Thought and Action*. Dordrecht: Springer, pp. 27-47, 2005a.
- . “Debating with myself and debating with others”. In: P. Barrotta, M. Dascal (eds.). *Controversies and Subjectivity*. Amsterdam: John Benjamins, pp. 31-73, 2005b.
- . “A dialética na construção do saber científico”. In: A.C. Regner, L. Rohden (eds.). *A Filosofia e a Ciência Redesenham Horizontes*. São Leopoldo: Editora Unisinos, pp. 15-31, 2005c.
- DASCAL, M., DASCAL, V. “A des-fixação da crença”. In: F. Gil, P. Livet, J.P. Cabral (eds.). *O Processo da Crença*. Lisboa: Gradiva, pp. 321-353, 2004.
- DAVIDSON, D. “Actions, reasons, and causes” [1963]. Repr. in D. Davidson (2001), pp. 3-19.
- . “How weakness of the will is possible?” [1969]. Repr. in D. Davidson (2001), pp. 21-42.
- . “Deception and division”. In: J. Elster (ed.). *The Multiple Self*. Cambridge: Cambridge University Press, pp. 79-92, 1986.
- . *Essays on Actions and Events*, 2nd ed. New York: Oxford University Press, 2001.
- ELSTER, J. “Introduction”. In: J. Elster (ed.). *The Multiple Self*. Cambridge: Cambridge University Press, pp. 1-34, 1986.
- . (ed.). *The Multiple Self*. Cambridge: Cambridge University Press, 1986.

- HAMPSHIRE, S. *Innocence and Experience*. Cambridge, Mass.: Harvard University Press, 1991.
- KANT, I. *Kritik der reinen Vernunft*. Leipzig: Reclam, 1971 [A = 1st ed., 1781; B = 2nd ed., 1787].
- LEIBNIZ, G.W. *Die Philosophischen Schriften* [1875-1890]. Ed. by C.J. Gerhardt. Berlin [repr. Hildesheim: Georg Olms, 1965].
- . *Sämtliche Schriften und Briefe*, serie 6, vol. 4, ed. Leibniz-Forschungstelle der Universität Münster. Berlin: Akademie Verlag, 1999.
- . *The Art of Controversies*. Ed. by M. Dascal. Dordrecht: Springer, 2006.
- PEARS, D. “The goals and strategies of self-deception”. In: J. Elster (ed.). *The Multiple Self*. Cambridge: Cambridge University Press, pp. 59-77, 1986.
- PEIRCE, C.S. *Collected Papers of Charles Sanders Peirce*. Ed. by C. Hartshorne. Cambridge, Mass.: Harvard University Press, 1931-1958.
- RORTY, A.O. “Self-deception, *akrasia* and irrationality”. In: J. Elster (ed.). *The Multiple Self*. Cambridge: Cambridge University Press, pp. 115-131, 1986.
- SEARLE, J.R. *Rationality in Action*. Cambridge, Mass.: The MIT Press, 2001.
- . “Desire, deliberation and action”. In: D. Vanderveken (ed.). *Logic, Thought and Action*. Dordrecht: Springer, pp. 49-78, 2005.

