

MESTRADO EM FILOSOFIA
FILOSOFIA CONTEMPORÂNEA

Irracionalidade Motivada e Agência Humana. Uma teoria explicativa da agência a partir da racionalidade e do livre arbítrio - a perspetiva de Alfred Mele.

Isabel Maria Fernandes de Sousa
Guedes

2019



Isabel Maria Fernandes de Sousa Guedes

**Irracionalidade Motivada e Agência Humana.
Uma teoria explicativa da agência a partir da racionalidade e
do livre arbítrio - a perspectiva de Alfred Mele.**

Dissertação realizada no âmbito do Mestrado em Filosofia orientada pela Professora Doutora
Sofia Gabriela Assis de Moraes Miguens Travis

Faculdade de Letras da Universidade do Porto

Julho de 2019

Irracionalidade Motivada e Agência Humana.
Uma teoria explicativa da agência a partir da racionalidade e
do livre arbítrio - a perspetiva de Alfred Mele.

Isabel Maria Fernandes de Sousa Guedes

Dissertação realizada no âmbito do Mestrado em Filosofia orientada pela Professora Doutora
Sofia Gabriela Assis de Moraes Miguens Travis

Membros do Júri

Professor Doutor João Alberto Cardoso Gomes Pinto
Faculdade de Letras - Universidade do Porto

Professora Doutora Sofia Gabriela Assis de Moraes Miguens Travis
Faculdade de Letras - Universidade do Porto

Professor Doutor Mattia Riccardi
Faculdade de Letras - Universidade do Porto

Classificação obtida: 17 valores

Sumário

Declaração de honra	7
Agradecimentos	8
Resumo	10
Abstract	11
Introdução	12
Capítulo 1 – Acrasia – Contextualização	16
Capítulo 2 – O contributo de Donald Davidson	18
2.1. A ação incontinente	18
2.2. Ação intencional e raciocínio prático.....	20
2.3. A relevância e a natureza da expressão <i>All things considered</i>	22
2.4. O princípio da continência.....	24
2.5. Considerações.....	25
Capítulo 3 – Irracionalidade Motivada – A perspectiva de Alfred Mele	29
3.1. Acrasia	31
3.1.1. Distinção de acrasia relativamente a Insanidade e Depressão	31
3.1.2. A fraqueza de vontade.....	34
3.1.3. Os desejos de segunda ordem e o compromisso prático.....	36
3.1.4. Fraqueza e compulsão.....	38
3.1.5. Melhor juízo.....	40
3.1.6. Ação Intencional.....	41
3.1.6.1. Intenções Proximal e Distal e Ação Intencional.....	41
3.1.6.2. Intenções Conscientes.....	42
3.1.6.3. Perspetivas Motivacional e Intelectual na explicação da ação intencional.....	45
3.1.7. Raciocínio prático e raciocínio avaliativo.....	48
3.1.8. O poder motivacional do raciocínio prático.....	51
3.1.9. <i>Self control</i>	54
3.1.9.1. Auto controlo e tentações proximais.....	59
3.1.10. A Pessoa idealmente autocontrolada.....	61
3.1.11. Desejo e motivação.....	62
3.1.11.1. Atitudes que essencialmente englobam motivação para a ação.....	63
3.1.11.2. Força Motivacional.....	65
3.1.12. <i>Core weak-willed action</i>	69
3.1.13. Hipótese modificada da proximidade.....	72

3.1.14. O papel das intenções.....	73
3.1.15. Sentimentos acráticos.....	77
3.1.15.1. <i>Strict Akritic Feelings</i>	79
3.1.16. A Decisão.....	87
3.1.16.1. A decisão prática.....	89
3.1.16.2. Decisões conscientes.....	94
3.2. O auto engano.....	94
3.2.1. Crenças e agência motivacionalmente enviesadas.....	96
3.2.2 Processos Psicológicos envolvidos na <i>Straight Self Deception</i>	98
3.2.2.1. O Paradoxo.....	98
3.2.2.2. Desejos e enviesamentos.....	99
3.2.2.3. <i>Acting as if</i>	103
3.2.3. Condições suficientes para o auto engano, na aquisição de crenças.....	104
3.2.4. Auto engano na retenção de crenças.....	105
3.2.5. <i>Wishful Thinking</i>	106
3.2.6. <i>Twisted Self Deception</i>	108
3.2.6.1. A abordagem da <i>twisted self-deception</i> centrada na Motivação.....	109
3.2.6.2. Abordagem híbrida da <i>twisted self-deception</i> que inclui motivação e emoção.....	110
3.2.7. Considerações.....	115
3.3. <i>Free Will</i>	121
3.3.1 A resposta Libertária.....	125
3.3.1.1. Libertarianismo Modesto.....	126
3.3.1.2. Libertarianismo Suave.....	126
3.3.1.3. <i>Daring Soft Libertarianism</i>	128
3.3.2 A resposta Compatibilista.....	133
3.3.3 O conceito de <i>Free Will</i>	135
Considerações Finais	137
Referências bibliográficas	147

Declaração de honra

Declaro que a presente dissertação/relatório é de minha autoria e não foi utilizado previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (afirmações, ideias, pensamentos) respeitam escrupulosamente as regras da atribuição, e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referência. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

Porto, 19 de julho de 2019

Isabel Maria Fernandes de Sousa Guedes

Agradecimentos

Este trabalho resultou de longo caminho, composto de incertezas, dificuldades e inúmeros desafios. Mas foi um caminho percorrido com o contributo de várias pessoas, indispensáveis na descoberta da melhor direção a seguir em cada momento. Não teria sido possível chegar ao destino sem o apoio, energia, estímulo e motivação de todos quantos caminharam ao meu lado.

Em primeiro lugar presto o meu sincero agradecimento à minha orientadora, Professora Doutora Sofia Miguens, pelo modo como acolheu e apoiou a escolha do tema, pela sua paciência e respeito pelo meu ritmo de trabalho e liberdade de pensamento. Agradeço a sua dedicação e constante disponibilidade. A sua orientação foi sempre muito próxima, de elevado rigor e nível científico, marcada por uma visão atenta, sempre oportuna e saudavelmente crítica, orientação essa, que acompanhou todas as etapas do meu trabalho, ajudando-me a encontrar o caminho. Não posso deixar de salientar também a dimensão pessoal, pois as suas palavras sempre se revestiram de incentivo e motivação, que alavancaram cada etapa deste projeto, oferecendo-me a confiança e segurança de que tanto precisava. Muito obrigada pelo seu empenho inexcedível.

Agradeço ao meu marido, Nuno Amado, pelo amor, apoio incondicional, generosidade e compreensão. Agradeço o seu sorriso constante, mesmo nos fins-de-semana que sacrificou em prol da realização deste meu projeto.

Agradeço à Filipa, a minha querida filha que amo incondicionalmente e que sempre me dirigiu palavras de incentivo e admiração. Espero conseguir compensar os momentos de atenção que lhe devo.

Agradeço à minha amiga de sempre, Branca Saldanha, pelo seu apoio e motivação. Por me ouvir tantas e tantas vezes, com toda a paciência e disposição. Sou grata pela nossa amizade.

Agradeço à minha colega de curso e amiga, Benilde Fonseca, pelo seu acompanhamento ao longo de todo o Mestrado, sempre com palavras que me transmitiam muita tranquilidade. Estou muito grata por ter tido o privilégio de a conhecer.

O meu profundo e sincero agradecimento a todos quantos contribuíram para a concretização deste trabalho, através do estímulo quer intelectual, quer emocional.

Resumo

O presente trabalho analisa o fenómeno da irracionalidade motivada centrando-se na obra do filósofo americano Alfred Mele. Segundo Mele, a maior parte dos casos de acrasia e autoengano derivam da irracionalidade motivada. A perspetiva motivacional de Mele permite ultrapassar a tese da divisão da mente.

Partindo da análise das propostas de Mele, procuro defender uma relação estreita entre os fenómenos da acrasia e do autoengano, sugerindo que a acrasia, vista como irracionalidade prática que se revela nas ações, poderá derivar do auto engano, visto como irracionalidade cognitiva, responsável pela formação de juízos enviesados, que o agente considera nas suas deliberações.

Salientando os fatores contextuais, procuro ainda demonstrar que o auto controlo permite ao agente o ajuste do seu comportamento a contextos práticos, ainda que exista um contexto cognitivo independente do contexto externo.

Palavras-chave: Acrasia, Autoengano, Irracionalidade Motivada, Auto controlo.

Abstract

This work aims to analyze the phenomenon of irrationality focusing on the work of the American philosopher Alfred Mele. According to Mele, most cases of acrasia and self-deception arise from motivated irrationality. Mele's motivational perspective allows us to overcome the division-of-mind thesis.

Starting from the analysis of Mele's proposals, I try to defend a close relationship between the phenomena of acrasia and self-deception, suggesting that acrasia, seen as practical irrationality revealed in actions, may arise from self-deception, seen as cognitive irrationality, responsible for formation of biased judgments, which the agent considers in his deliberations.

Stressing the contextual factors, I also try to demonstrate that self control allows the agent to adjust his behavior to practical contexts, even though there is a cognitive context independent of the external context.

Keywords: Acrasia, Self-deception, Motivated Irrationality, Self-control

Introdução

Desde sempre senti um fascínio pela observação e compreensão do comportamento humano, em concreto, no que se relaciona com a existência de determinados padrões, que marcam o “ser” de cada sujeito, e que se refere talvez àquilo que normalmente ouvimos definir como traço de personalidade. A minha Licenciatura em Psicologia permitiu-me conhecer os padrões considerados patológicos e que, portanto, afetam a dinâmica de vida normal dos sujeitos que as experienciam. Esses padrões vêm identificados e caracterizados no Manual de Diagnóstico e Estatística das Perturbações Mentais (DSM), da *American Psychiatric Association*, que classifica as diversas perturbações mentais, através de critérios desenvolvidos de modo a permitirem o seu diagnóstico fiável¹.

Os critérios de diagnóstico do DSM, identificam sintomas, comportamentos, funções cognitivas, traços de personalidade, sinais físicos e combinações de síndromes e durações, que visam distinguir a variação da vida normal, das respostas transitórias ao *stress*². Ou seja, a experiência dos profissionais ao longo dos anos, permitiu compilar e codificar padrões de comportamento, que abarcam diversas áreas de orientação (biológica, psicodinâmica, cognitiva, comportamental, interpessoal, família/sistemas³), através de um sistema de classificação categorial, considerando os diferentes domínios de sintomas. Por outro lado, as perturbações mentais são definidas, considerando normas e valores culturais, sociais e familiares e, os limites daquilo que se considera normalidade e patologia, variam com as culturas relativamente a tipos específicos de comportamento e portanto, o nível a que uma experiência se torna problemática ou patológica será diferente. Considerar um comportamento como sendo anormal, vai depender das normas culturais internalizadas pelo sujeito e aceites pelas pessoas que o rodeiam⁴.

Com efeito, o DSM descreve a definição de perturbação mental, como contendo os seguintes elementos:

¹ American Psychiatric Association, *DSM-5: Manual de Diagnóstico e Estatística das Perturbações Mentais* (5ª Ed.), Climepsi Editores, Lisboa 2014, p. XLIII.

² American Psychiatric Association, *DSM-5*, op.cit, p. 6.

³ American Psychiatric Association, *DSM-5*, op.cit, p. XLIII.

⁴ American Psychiatric Association, *DSM-5*, op.cit, p. 16.

Uma perturbação mental é uma síndrome caracterizada por perturbação clinicamente significativa, da regulação da emoção, ou do comportamento do sujeito que reflete uma disfunção dos processos psicológicos, biológicos ou do desenvolvimento subjacentes ao funcionamento mental. As perturbações mentais estão geralmente associadas com significativo mal-estar ou défice social, ocupacional, ou noutras atividades importantes⁵.

Mas, ao longo do tempo, surge em mim uma outra inquietação, relativamente a comportamentos que se observam no quotidiano mas que, apesar de não cumprirem os critérios que julgamos racionais, também não se enquadram naquilo que se classifica como patologia. É o caso da ação acrática ou do autoengano.

Estes tipos de irracionalidade não dependem de contextos culturais específicos ou alterações ao nível químico cerebral, nem tão pouco carecem de tratamento/terapia especial. Tratam-se, sim, de processos universais, experienciados por sujeitos normais e que, apesar de não gerarem significativo mal-estar ou défice social, ocupacional, ou noutras atividades importantes, critérios envolvidos na definição de perturbação mental referida supra, afetam, sem dúvida alguma, a vida dos sujeitos, na medida em que constituem situações de agência irracionais.

Afinal, quais os processos mentais que podemos identificar como estando envolvidos nestes casos de irracionalidade? E de que modo podemos explicar estas ações irracionais? O tema em análise neste trabalho, reflete o culminar de muitas questões pessoais e o assunto abordado decorre da curiosidade acerca daquilo que parece transcender a nossa compreensão, tornando-se num desafio que exige, de facto, uma profunda reflexão, ainda que com a consciência que haverá sempre, na compreensão da mente humana, algo que nos escapa. Porque é inegável que o agente humano se contraria a si próprio e este facto reflete-se em pensamentos e atitudes que se constituem paradoxais.

A resposta à minha inquietação e o entusiasmo em aprofundar a temática surge na Unidade Curricular de Filosofia da Ação, com o desenvolvimento de um trabalho acerca da Acrasia, sustentado na perspetiva de Alfred Mele. Este tema e o gosto particular pela forma de exposição de Alfred Mele aumentou o meu interesse, levando-me a alargar as minhas leituras.

⁵ American Psychiatric Association, *DSM-5*, op.cit, p. 22.

Dessas leituras depreendo que Alfred Mele, defende que todo o ser humano é capaz de ações racionais e irracionais, sem que isso o rotule de agente irracional. E explica e analisa esta questão à luz de uma teoria da agência humana, capaz de acomodar ações de todos estes tipos. Explica as ações intencionais à luz do poder do raciocínio prático, juízos avaliativos práticos, racionalidade humana e livre arbítrio. De facto, Alfred Mele perspetiva o ser humano de forma holística e defende que o *self* ou o autocontrolo não se identificam propriamente com a razão. Antes, identificam-se com a pessoa – *person* – concebida de forma ampla⁶. E será este carácter holístico do ser humano, que poderá tornar possível a conceptualização e explicação de instâncias de ação de raiz irracional.

Neste contexto, surgiu o meu interesse de escrever esta dissertação, acerca de um tema que se torna, afinal, numa descoberta do ser humano, como agente dotado de processos mentais que se exteriorizam por ações racionais mas também irracionais. Este trabalho traduz-se na procura de compreensão de conceitos, relações e significados essenciais à compreensão do fenómeno da irracionalidade motivada, através da lente de Alfred Mele. Assim, serão abordadas questões relativas à natureza e explicação da ação, nomeadamente, acrasia e autoengano como sendo instâncias de comportamento irracional, mas outras questões associadas tais como racionalidade prática, intenções, agência, decisão, deliberação, juízo, tomando como fio condutor a obra e as teses de Alfred Mele.

A partir de Mele, será defendida a posição segundo a qual os comportamentos irracionais fazem parte do quotidiano dos agentes humanos, a par com os comportamentos racionais e é possível explicar estes fenómenos de agência humana, a partir de uma teoria que aborde as diversas dimensões do ser humano, a saber, o raciocínio prático, os fatores internos ao agente, fatores externos/contextuais e livre arbítrio.

Assim, este trabalho passará por ligar a abordagem de Alfred Mele relativamente a cada uma destas dimensões, plasmada ao longo da sua obra, procurando mostrar que formam uma explicação clara e consistente acerca destes tipos de agência.

Tentarei ainda estabelecer uma ligação entre Acrasia e Auto engano, na medida em que parece plausível considerar que ambos os fenómenos têm uma raiz comum, que reside

⁶ MELE, Alfred, *Autonomous Agents*, Oxford University Press, New York 1995, p. 6.

numa irracionalidade motivada, irracionalidade esta, instanciada por agentes racionais, livres e mentalmente sãos.

Capítulo 1 – Acrasia. Contextualização

A Acrasia é um problema de caráter prático. Refere-se a um estado que nos impede de executar certos projetos e intenções, que nos seriam mais favoráveis. É também um problema teórico, central ao trabalho Filosófico, na medida em que constitui um desafio a qualquer teoria da racionalidade. A questão que desde logo se impõe é de que forma, ou, porque razão pode um sujeito escolher, de forma ciente e deliberada, uma opção menos vantajosa?

Alguns Filósofos defenderam que as coisas não são bem aquilo que parecem neste domínio. Sócrates rejeita a crença de que, apesar das pessoas conhecerem aquilo que é melhor fazer, não estão dispostas a fazê-lo, apesar de poderem. Platão já aceita a possibilidade de um sujeito agir contrariamente ao seu melhor juízo, mas refere que quando isto acontece, fazer aquilo que o sujeito julga melhor fazer, não estava ao seu alcance, podendo resultar daqui a ideia de que, aquilo que aparenta ser uma ação decorrente da fraqueza de vontade (*weak-willed*) é, na verdade, uma ação não livre. (Daqui em diante usarei o termo *weak-willed*, por ser mais curto).

Este argumento é defendido ainda por Richard Hare, proponente do prescritivismo e nome importante na história da ética. Este autor apela a uma alegada conexão lógica entre juízo e ação⁷. Para Richard Hare, os juízos morais, servem para guiar a conduta e qualquer caso aparente de acrasia tem que ser aquele em que o agente é incapaz de fazer algo, ou então aquele em que o agente não avalia esse algo como sendo o melhor a fazer, ainda que o faça⁸.

Hugh McCann relaciona o fenómeno da acrasia com o dever moral, na medida em que entende que a acrasia ocorre, quando o agente não tem a motivação para levar a cabo as intenções formadas. A acrasia é, na sua opinião, uma falha na retidão de uma pessoa o que, por sua vez, diz algo acerca do seu caráter⁹. O problema é que, podemos chegar a um julgamento daquilo que é melhor, ou daquilo que é obrigatório e não possuir recursos

⁷ MELE, Alfred, *Backsliding. Understanding Weakness of Will*, Oxford University Press, New York 2012, pp. 1-2.

⁸ HARE, Richard Mervyn, *Freedom and Reason*, Oxford University Press, New York, 1965, pp. 70-85.

⁹ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, Cambridge Scholars Publishing 2013, p. 10.

motivacionais para formar ou executar a intenção apropriada. A segunda opção ocorre quando nos decidimos a fazer o que é suposto, formando uma intenção relevante, mas acabamos tão preocupados com outros motivos, que o valor do primeiro curso de ação sai da nossa atenção e apenas a alternativa inicial rejeitada, obtém suporte para a razão verdadeiramente prática¹⁰. Para o autor, a acrasia tem uma conotação moral, uma vez que a fraqueza de vontade é uma falha de coragem e portanto, vai contra o caráter moral do agente. Existem no entanto instâncias que não envolvem falha na observação do dever moral: um montanhista que desista de escalar uma montanha, por medo, exhibe acrasia sem violar nenhum dever moral¹¹.

Gary Watson e David Pugmire, por seu lado, focam-se na noção de resistência, defendendo que, qualquer agente que pudesse ter resistido à tentação tê-lo-ia feito.

Richard Holton defendeu que é possível que o sujeito exhiba acrasia sem exhibir fraqueza de vontade e vice-versa¹², na medida em que entende a noção de fraqueza de vontade como sendo distinta da noção de Acrasia. Com efeito, para este Filósofo, a fraqueza de vontade acontece quando o agente está demasiado pronto para rever as suas intenções e desenvolve a ideia de que fraqueza de vontade é melhor caracterizada, não como aqueles casos em que as pessoas agem contra o seu melhor juízo, mas sim, nos casos em que os sujeitos falham em agir conforme as suas intenções. Ou seja, para exhibir fraqueza de vontade um agente tem que ter formado uma resolução, que vai rever. Richard Holton oferece ainda a condição suficiente: se alguém avalia prontamente uma resolução, então esse alguém exhibe fraqueza de vontade¹³.

¹⁰ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Pratical Rationality*, op. cit, p. 79.

¹¹ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Pratical Rationality*, op. cit, p. 83.

¹² MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 1-2.

¹³ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 13-18.

Capítulo 2. – O contributo de Donald Davidson

Dedico este capítulo a Donald Davidson, que através de um artigo incisivo publicado em 1970, pretendeu tratar a temática da Acrasia, tese que teve enorme influência na literatura subsequente. Donald Davidson pretendeu defender a possibilidade da fraqueza de vontade ou ação incontinente, definindo-a como um caso de agência defeituoso, cuja natureza nos pode levar até a questionar se é, de facto, inteligível. A sua caracterização de ação incontinente é feita no artigo *How is weakness of the will possible*, sob uma visão geral do raciocínio prático e ação intencional, que desempenham um papel central nos nossos juízos avaliativos¹⁴. Ou seja, para Donald Davidson é fundamental o conceito de raciocínio prático para a explicação da ação acrática e é a partir da racionalidade/irracionalidade que chegará à sua conclusão.

2.1. A ação incontinente

Segundo Donald Davidson, a vontade de um agente é fraca, se o agente age, e age intencionalmente, contra o seu próprio melhor juízo, casos relativamente aos quais se diz que, por vezes, falta ao agente a vontade/força para fazer aquilo que ele julga, ou em algum nível acredita, seria melhor, consideradas todas as coisas. Estas são as ações que Donald Davidson denomina de incontinentes, ou que, ao fazê-las, o agente age incontinentemente. Nesta teoria, Donald Davidson discute a ação incontinente dependente apenas da atitude ou crença do agente e como sendo não independente do conhecimento do agente, de que outro curso de ação seria melhor. Ou seja, se um agente acredita que determinado curso de ação é o melhor, consideradas todas as coisas, ou essa é a coisa certa a fazer ou aquilo que deveria fazer e, mesmo assim, faz outra coisa qualquer, o agente age incontinentemente. Mas o agente age incontinentemente, também porque sabe que tem disponível outro curso de ação, melhor do que aquele que desempenhou, ou seja, juízos comparativos bastam para a incontinência. Donald Davidson caracteriza a ação incontinente ou reveladora de fraqueza de vontade, segundo as seguintes condições:

¹⁴ STROUD, Sarah, «Weakness of Will», *The Stanford Encyclopedia of Philosophy*, (Spring 2014 Edition) Edward N. Zalta (ed.).

Ao fazer x , um agente age incontinentemente se e só se:

- a) o agente faz x intencionalmente
- b) o agente acredita haver uma alternativa de ação y , disponível;
- c) o agente julga, consideradas todas as coisas, que seria melhor fazer y do que fazer x ¹⁵.

Donald Davidson acredita na existência de ações incontinentes e avança a razão pela qual entende justificar a tendência para negar este fenómeno, apresentando dois princípios, que ganham força a partir de uma visão muito persuasiva da natureza da ação intencional e do raciocínio prático o que implicaria, aparentemente, que a incontinência é impossível¹⁶:

P1 – Se um agente quer fazer x mais do que quer fazer y e acredita ser livre de fazer x ou y , então o agente vai intencionalmente fazer x , se fizer x ou y intencionalmente.

P2 – Se um agente julga que será melhor fazer x do que fazer y , então o agente quer fazer x mais do que quer fazer y ¹⁷.

Para Donald Davidson, P2 relaciona juízos daquilo que é melhor fazer com a motivação ou a vontade e acrescenta que P2 se refere a uma forma ligeira de internalismo. Para o Filósofo, as proposições P1 e P2 são atrativas, na medida em que formam uma moldura de raciocínio prático e ação intencional, que atribui um papel motivacional importante aos juízos avaliativos do agente. Mas apesar de atrativa, esta moldura, considerando P1 e P2 em conjunto, implicariam que a ação incontinente não é possível e, por consequência, a ação acrática seria impossível¹⁸. Mas, na realidade P1 e P2 não tornam a ação acrática impossível, uma vez que a sua definição caracteriza o agente que incontinentemente faz b , como assumindo, não que seria melhor fazer a do que fazer b , mas assumindo que seria melhor fazer a do que b , consideradas todas as coisas. É precisamente este último juízo que torna possível a ação acrática, mesmo que P1 e P2 sejam verdade. A expressão *all things considered* marca a diferença entre estes dois juízos, porque o agente pode deter o segundo sem deter o primeiro. A referida expressão tem um papel importante na lógica

¹⁵ DAVIDSON, Donald, *Essays on Actions and Events*, Oxford University Press Inc., New York 2001, pp. 21-22.

¹⁶ STROUD, «Weakness of Will», *art. cit.*

¹⁷ DAVIDSON, *Essays on Actions and Events*, op. cit., p. 23.

¹⁸ Um agente nunca faria intencionalmente b , quando julga que seria melhor fazer a , sendo o agente livre de fazer a ou fazer b .

pela qual poderíamos entender a estrutura do raciocínio prático, na medida em que indica um juízo que é condicional ou relacional, em vez da forma incondicional ou *all-out*¹⁹. Mais adiante, veremos a especificidade desta expressão no quadro teórico de Donald Davidson. Portanto, a atribuição da importância que Donald Davidson dá aos juízos do agente permite inferir, parece-me, o caráter profundamente internalista da sua teoria. Donald Davidson desenha duas formas de incontinência, que se relacionam e tendem até a confundir-se: por um lado o desejo distrai-nos do bem ou força-nos para o mal e por outro lado, a ação incontinente favorece sempre as paixões egoístas que se sobrepõem ao sentido do dever e da moralidade. Nesta matéria Donald Davidson propõe separar a incontinência da moralidade, e defende que existem numerosas ocasiões nas quais o prazer imediato se opõe a princípios, educação ou sentido do dever e, mesmo assim, julgamos ou sabemos que, consideradas todas as coisas, devemos optar pelo prazer^{20, 21}.

2.2. Ação intencional e raciocínio prático

Quando um agente age com intenção, valoriza positivamente uma situação (um fim ou uma ação que satisfaz certas condições). Ou seja, o agente sabe ou percebe que determinada ação, que lhe é livre de fazer, realizará a situação valorizada, motivo pelo qual o agente age. E age, devido ao seu valor ou desejo e a sua crença. Esta visão, segundo Donald Davidson, está próxima de Aristóteles, no que diz respeito à explicação de uma ação, descrevendo as razões que o agente teve para agir e permite iniciar uma discussão acerca do raciocínio prático, o raciocínio que leva à ação. Segundo Aristóteles, existindo o desejo, o agente infere a ação que acredita poderá satisfará esse desejo. Este desejo e esta crença criam as condições para a ação intencional. Esta explicação é obviamente insuficiente para uma análise satisfatória da ação incontinente. Portanto, acrescenta a esta visão o ponto de vista de Tomás de Aquino, que desenha o drama de um homem

¹⁹ STROUD, «Weakness of Will», *art. cit.*

²⁰ Esta ideia foi ilustrada com o seguinte exemplo: já estou deitada, após um dia cansativo, e percebo que me esqueci de lavar os dentes. Considerando a minha saúde, seria melhor levantar-me e ir escová-los, mas talvez não seja tão grave não os escovar apenas por aquele dia e, além disso, se me levantar irei ficar irritada, o que irá alterar a minha noite de sono. Consideradas todas as coisas, concluo que é melhor permanecer deitada, mas ainda assim o meu dever de ter que escovar os dentes vence e eu levanto-me: o meu ato é intencional, apesar de ir contra o meu melhor juízo e, por isso, é incontinente.

²¹ DAVIDSON, *Essays on Actions and Events*, op. cit., pp. 23-30.

incontinente, a quem se apresenta o lado da razão (pelo qual determinado ato não é permitido, por ser ilegal) em contradição com o lado da luxúria (pelo qual esse mesmo ato deve ser desempenhado, por ser prazeroso). Mas para Donald Davidson, este quadro de raciocínio moral, é inadequado para a análise da incontinência uma vez que não consegue dar uma correta interpretação de casos simples de conflito moral²². Esta imagem apenas reflete uma batalha entre dois lados, o da razão e o da luxúria, facto que não considera sequer o processo de pesar considerações ou até mesmo atribuir a culpa ao agente. Por isso, Donald Davidson adiciona um terceiro interveniente, a consciência - *The Will*. É a consciência que decide quem ganha a batalha, que é capaz de julgar a força dos argumentos de ambos os lados em conflito, executar a decisão e assumir as responsabilidades.

Resta resolver o problema de como pode a consciência julgar uma opção como melhor e no entanto, não optar por esse melhor caminho. Donald Davidson diz que não nos basta conhecer as razões de cada lado, temos que perceber como elas emergiram, questão que se relaciona com juízos efetuados pelo lado certo (razão, moralidade, família, patriotismo) ou juízos baseados nas considerações relevantes conhecidas pelo agente, este sim, o juízo pertinente. Centra-se aqui no problema da formação ou natureza do raciocínio prático, problema que reside no facto de existir a assunção tácita de que os princípios morais têm a forma de condicionais universalizados.

Donald Davidson propõe usar o padrão do raciocínio prático a partir da ideia central de que um princípio moral, tal como, mentir é *prima facie* errado, não pode ser interpretado coerentemente como um condicional universalmente quantificado, mas antes deve ser reconhecido como significando que um ato sendo mentira, *prima facie*, faz desse ato um ato errado. A ação intencional é gerada diretamente de juízos incondicionais tais como - seria melhor fazer *a* do que fazer *b*. Assim, o raciocínio que pára em juízos condicionais (*prima facie*, *a* é melhor do que *b*, na presença de todas as considerações relevantes conhecidas por nós) «*is practical only in its subject, not in its issue*»²³. Desta forma, podemos conceber logicamente, que a incontinência, para os agentes acráticos, é

²² Refere-se a conflito moral como sendo um caso em que existem boas razões para fazer uma ação e existem boas razões para fazer o contrário ou mesmo, retrair-se de fazer a ação.

²³ DAVIDSON, *Essays on Actions and Events*, op. cit., p. 39.

caracterizada como entender que, consideradas todas as coisas, será melhor fazer *b* do que fazer *a*, mesmo que o agente faça *a* em vez de *b*, e por uma razão. O juízo de que *a* é melhor do que *b*, consideradas todas as coisas, é um juízo relacional ou *prima facie* e, portanto, não pode conflitar logicamente com nenhum juízo incondicional. Daqui se poderia formar a pergunta: mas como pode um agente julgar *a* melhor do que *b*, consideradas todas as coisas e não julgar que *a* é melhor do que *b*²⁴?

A partir desta questão, Donald Davidson, vai explanar acerca da relevância da expressão: consideradas todas as coisas - *All things considered*.

2.3. A relevância e a natureza da expressão *All things considered*.

A expressão *all things considered* refere-se ao que o agente conhece, acredita ou detém, é a soma dos seus princípios relevantes, opiniões, atitudes e desejos²⁵.

Conseguimos perceber o carácter relacional de um juízo *all things considered*, se olharmos para os juízos avaliativos, que desempenham um papel importante na fase inicial do raciocínio prático, na qual consideramos que razões e considerações favorecem fazer *a*, e que razões e considerações que favorecem fazer *b*. Os juízos *prima facie* apontam uma relação na qual à luz de *r* (em que *r* se refere a uma consideração), é *prima facie* melhor fazer *a* do que fazer *b*. Os juízos *prima facie* identificam um aspeto ou perspectiva pelo qual *a* é superior a *b* (*favoring relation*).

Os juízos *prima facie* apresentam três características:

- o juízo *prima facie* não é, ele próprio, uma conclusão acerca da superioridade global de *a*, na medida em que não nos permite uma conclusão avaliativa incondicional a favor de fazer *a*, *prima facie*. O tipo de juízo avaliativo *all-out* tem a forma lógica mais simples de que *a* é melhor do que *b*;
- nenhum juízo *all-out* pode advir de um juízo *prima facie*;
- o facto de um agente ter chegado a um juízo *prima facie*, nem sequer dá a esse agente bases suficientes para desenhar a conclusão *all out* correspondente, considerando que esse

²⁴ DAVIDSON, *Essays on Actions and Events*, op. cit., pp. 31-39.

²⁵ DAVIDSON, *Essays on Actions and Events*, op. cit., p. 40.

juízo *prima facie*, é baseado numa consideração e, se essa consideração é outra, outro juízo *prima facie* pode ser feito de que *b* é melhor do que *a*.

O juízo *prima facie* têm caráter relacional, entre a consideração *r* e fazer *a* (*favoring relation*), a partir de um *prima facie* e da suposição que *r* obtém. Essa relação não nos permite alcançar uma conclusão avaliativa incondicional para fazer *a*, ou seja, nenhum juízo *prima facie* reveste a forma de um condicional material. Para Donald Davidson, um juízo com esta forma lógica *prima facie*, é uma forma apropriada para modelar aquilo que acontece nas primeiras fases do raciocínio prático, nas quais ensaiamos as razões pró e contra as opções que estamos a considerar. Mas nenhum juízo *prima facie*, em separado, compromete o agente a uma conclusão global (*overall*) a favor de *a* ou *b*. Apesar do raciocínio prático partir de juízos *prima facie*, cada um identificando uma consideração/aspecto pela qual uma das opções é superior a outra, o progresso do raciocínio prático precisa de considerar como *a* se compara a *b*, não apenas no que respeita a uma consideração/aspecto, mas à luz de diversas considerações, tidas em conjunto. Portanto, o juízo *all things considered*, é um juízo condicional relacional e não um juízo *all-out* a favor de *a*²⁶. O juízo *all things considered* é um juízo relacional, na sua natureza e distinto do juízo *all-out*, ou seja, é possível fazer um juízo *all things considered* a favor de *a*, que não corresponde ao juízo *all-out* a favor de *a* (tal como o caso da investigação). Esta é a chave, segundo Donald Davidson, para compreender a fraqueza de vontade. P1 e P2, em conjunto, implicam que o agente que chegue a uma conclusão *all-out* a favor de *a*, não vai intencionalmente fazer *b*. Mas um agente incontinente nunca chega à conclusão *all-out*. Para explicar este ponto, é necessário retomar as três características dos juízos *prima facie*. A primeira e segunda características, sustentam a subclasse de juízos *prima facie*, que são um juízo *all things considered*. Esses juízos não são equivalentes, nem implicam logicamente nenhum juízo *all-out*. Assim, o agente incontinente, que falha ao desenhar a conclusão *all-out* que corresponderia à sua conclusão *all things considered* e fazer a ação correspondente, não está a contradizer-se: para Donald Davidson esse agente exhibe um defeito de racionalidade. A terceira

²⁶ Por exemplo, numa investigação de um crime, o facto de uma prova incriminar um sujeito *x*, não quer dizer que esse sujeito seja de facto o culpado. É necessário considerar o conjunto das provas existentes e relacioná-las, de forma a conseguirmos chegar a uma conclusão global (*over all*).

característica, dos juízos *prima facie* não sustentam, normalmente, a subclasse de juízos que são *all things considered*²⁷.

2.4. O Princípio da continência

Ao refletir sobre a falha na incontinência, Donald Davidson defende que o agente não mantém logicamente crenças contraditórias, nem a sua falha é necessariamente uma falha moral. Aquilo que está errado é que, o agente incontinente age e julga irracionalmente²⁸.

O desenhar de um juízo *all things considered* a favor de *a*, dá ao agente bases suficientes para concluir que *a* é melhor, *sans phrase* e, portanto, fazer *a*. A transição de um juízo *all things considered* a favor de *a*, para o correspondente juízo *all-out*, e então fazer *a*, é imposto por um princípio de racionalidade, a que Donald Davidson chama o princípio da continência. Este princípio diz-nos para «*perform the action judged best on the basis of all available relevant reasons*»²⁹. Ora, o agente incontinente viola este princípio tornando-se irracional, porque seguir este princípio é condição para a racionalidade³⁰.

O princípio da continência convida-nos a ações que podemos desempenhar se quisermos, ou seja, deixa os motivos connosco. Aquilo que é difícil é adquirir a virtude da continência, de modo a tornar nosso, o princípio da continência. À questão de qual a razão do agente para fazer *a* quando acredita que, consideradas todas as coisas, seria melhor fazer outra coisa, a resposta será: o agente não tem razão nenhuma, no sentido de que, a razão para fazer *a* existe sim, mas o que falta ao agente é a razão para não deixar prevalecer a sua melhor razão, para não fazer *a*. Na incontinência, o agente não consegue entender-se a si próprio, porque «*he recognizes, in his own intentional behaviour, something essentially surd*»³¹.

Mas a relação entre irracionalidade e incontinência merece algum cuidado. Com efeito, segundo Donald Davidson, a ação incontinente, deve ser ela própria intencional, ou seja, feita por uma razão. O homem incontinente tem a razão para fazer *b* e faz *b* por essa razão.

²⁷ STROUD, «Weakness of Will», art. cit.

²⁸ DAVIDSON, *Essays on Actions and Events*, op. cit., pp. 40-42.

²⁹ DAVIDSON, *Essays on Actions and Events*, op. cit., p. 41.

³⁰ STROUD, «Weakness of Will», art. cit.

³¹ DAVIDSON, *Essays on Actions and Events*, op. cit., p. 42.

Aquilo que lhe falta, e falta segundo a sua própria visão, é a razão suficiente para fazer *b*, dadas todas as considerações que tem em favor de *a*. Ou seja, o agente, segundo a sua própria visão, não tem uma razão adequada para fazer *b* e por isso, não consegue que a sua ação faça sentido. Portanto a ação acrática, na visão de Donald Davidson, é necessariamente irracional e é neste sentido que se trata de uma instância de agência defeituosa e não completamente inteligível³².

2.5. Considerações

Na teoria Davidsoniana, o agente que age acriticamente, fá-lo porque se baseia num juízo condicional, efetuado com base no seu próprio conjunto de considerações, que é menor, e não num juízo incondicional, e portanto, não chega a uma conclusão *all-out*. Assim, parece-se legítimo supor que a tese desenvolvida por Donald Davidson está imbuída de um caráter profundamente internalista. Aliás, ele próprio assim o reconhece, quando refere que o princípio P2 se refere a juízos de valor refletidos em vontades.

Donald Davidson concebe uma avaliação do agente, que é interna ao próprio agente e que se vai relacionar com a ação. Mas parece que esta visão deixa de fora fatores externos que poderão influenciar o juízo decisivo para a ação, bem como fatores psicológicos, que podem não estar acessíveis à consciência na altura em que o agente age, mas que podem constituir a motivação para determinado curso de ação, ainda que de forma inconsciente, ou não consciente nesse momento *t*, da ação. Segundo Donald Davidson agir acriticamente pressupõe que existe uma irracionalidade, dado que não existe consonância entre os princípios, opiniões, atitudes e desejos do agente, elementos constitutivos do seu juízo, consideradas todas as coisas, e aquilo que o agente faz.

O caráter internalista da conceção Davidsoniana parece relacionar as razões apenas com o aspeto ou interpretação interna do agente. Ou seja, parece considerar a representação interna do agente relativamente a um determinado conjunto de crenças e valores, que determinam o seu melhor juízo.

Por outro lado, esta representação, bem como o alcance da expressão *all things considered*, apela a uma consciência plena, segura, das razões que o agente gere

³² STROUD, «Weakness of Will», art. cit.

internamente. Portanto, por um lado, podemos considerar o caso no qual, apesar do agente estar certo do seu conjunto de crenças e valores, consideradas todas as coisas, age contra eles, sem que contudo consiga explicar o motivo, ou seja, age acráticamente. Mas, por outro lado, consideremos diferente caso, no qual, muitas vezes ouvimos alguém dizer, nem sei por que motivo fiz aquilo ou, não sei como fui capaz de fazer aquilo, porque sei que não o devia ter feito, mas ainda bem que fiz. Ou seja, neste exemplo, o carácter acrático de determinada ação, num momento t , parece, subitamente, passar a fazer sentido, o que poderá transformar essa mesma ação numa ação continente, no momento tI . Talvez, nestes casos, o agente chegue à conclusão de que, o seu motivo para agir no momento t se baseou em crenças, valores e desejos que não estavam acessíveis à sua consciência naquele momento, mas que foram, no fundo, as razões pelas quais agiu. Este conjunto de crenças e valores, não conscientes, tornaram essa ação irracional, ou subjetivamente irracional, diria eu. Ou seja, algumas ações acráticas poderão sê-lo, porque não coincidem com as razões racionais ou com o melhor juízo, consideradas todas as coisas, mas apenas à primeira vista.

Se considerarmos o sentido Freudiano nesta análise, podemos conceber, penso eu, que o agente fez a , em desajuste com as razões que habitam o superego, mas o fazer a é congruente com os conteúdos do id, que não estavam acessíveis à consciência, na altura da ação. Por isso é que o agente acrata diz que fez a por razão nenhuma. O agente, de facto, não tinha consciência das verdadeiras razões. Neste sentido, algumas ações acráticas teriam uma explicação subjetiva, que explicaria o conflito interno experienciado pelo agente, precisamente devido às motivações conflitantes entre o id e o superego. Alfred Mele discordaria desta suposição, na medida em que defende a análise da ação acrática, através da perspetiva motivacional, sem precisar de recorrer à tese da divisão da mente, conforme veremos adiante.

Mas outros autores têm diferentes opiniões. Por exemplo, George Ainslie aplica a noção de *picoeconomics*³³ a um modelo do *self* e da vontade – *will*, que lhe permite explicar os

³³ A ideia contida nesta noção é que, da mesma forma que a economia clássica, visa descrever a negociação de recursos limitados entre instituições e da mesma forma que as microeconomias visam descrever a negociação de recursos limitados entre indivíduos, a *picoeconomics* visa descrever as interações para o controlo da capacidade comportamental finita de um indivíduo, dentro desse indivíduo. Tal como nos dois

colapsos da vontade. Defende que, na compreensão do *self* e da vontade é particularmente relevante o facto de os agentes serem seduzidos, frequentemente, por prazeres urgentes relativamente a objetos destrutivos, que interferem com as escolhas de longo prazo estáveis, detidas pelo sujeito. É como se houvesse um *self* que se encontra em alerta constante para o prazer compulsivo, imediato e fortemente motivador e, esse *self* ganha inúmeras vezes, ainda que isso signifique falhar a satisfação de interesses de longo prazo do outro *self*, que é também o *self* do agente, falhando a resposta àquilo com que esse outro *self* estava comprometido.

Se o autor está certo, então, compreender as preferências de agentes como nós, envolve muito mais do que considerar os valores relativos implicados na teoria da escolha racional. A ideia básica aqui contida relaciona-se com a teoria Freudiana do id, ego e superego. Ou seja, considera que existem diferentes partes ou aspetos da personalidade que estão em conflito no indivíduo e esta hipótese deve ser sempre considerada nas escolhas do agente. Este fenómeno é particularmente claro no comportamento de vício³⁴. Também Joshua Knobe refere que a maior parte das nossas decisões fundamentais parecem ser determinadas, em grande parte, por estereótipos não conscientes, heurísticas rápidas e contaminadas e atalhos mentais³⁵.

Para o autor, tem vindo a ser dado demasiado ênfase ao papel da reflexão e da deliberação, em certas áreas da Filosofia. O foco tem vindo a ser a nossa capacidade de nos determos, de refletirmos no valor dos nossos desejos e na aplicação de certos princípios racionais, mas estes processos parecem desempenhar apenas um papel periférico.

Portanto, um estudo apropriado da ação humana precisa de explorar um conjunto mais amplo de fatores: processos não conscientes e automáticos, instantes de emoções, e influências subtis, porém poderosas, que moldam as nossas vidas³⁶.

primeiros casos, as interações para o controlo da capacidade comportamental finita de um indivíduo, dentro desse indivíduo, assemelha-se a uma negociação entre partes.

³⁴ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 14-15.

³⁵ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., p. 135.

³⁶ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., p. 137.

Relativamente à acrasia, Joshua Knobe, invoca o cenário proposto por Platão, segundo o qual a mente é composta por partes distintas e, uma dessas partes, a que Platão chamou de *thumos*, serve um tipo de poder que reforça as decisões da razão e rejeita as inclinações dos apetites. Acrescenta ainda que o trabalho empírico recente corrobora esta visão e que, de facto, a mente é composta de partes distintas³⁷.

Mas, voltando a Donald Davidson, o que pretendo sugerir é que, no contexto Davidsoniano, o apuramento das razões que constituem o melhor juízo não é, parece-me, tão facilmente determinável quanto Donald Davidson faz supor. Pelo contrário, as razões de cada agente têm uma plasticidade idiossincrática à vivência de cada agente em particular e poderão variar, penso eu, até em função do contexto existente no momento da ação. Esta ideia parece ser corroborada pelo pensamento de Alfred Mele, quando se refere ao contexto psicológico do agente acrático, conforme veremos mais adiante.

Após Donald Davidson, outros cursos de ação foram explorados, concretamente na inclusão de fatores de natureza externa, que enfatizam os fatores motivacionais, além do juízo avaliativo do agente, bem como a possível divergência resultante da avaliação das opções do agente e a sua motivação para agir.

É nesta linha que Alfred Mele introduz, na sua conceção de acrasia, noções que ainda não haviam sido tratadas e/ou consideradas, tais como motivação e autocontrolo.

O pensamento de Alfred Mele contrasta, a meu ver, com a posição de Donald Davidson, pelo seu carácter externalista, ao incorporar fatores motivacionais, além dos juízos avaliativos do agente acerca das suas opções, admitindo ainda a divergência entre estas duas instâncias.

Para o autor, a acrasia não é de todo uma ilusão e as teorias da ação, o poder dos juízos avaliativos, agência humana, racionalidade, raciocínio prático e motivação, devem acomodar este fenómeno.

Com efeito, ao longo da sua obra, Alfred Mele desenvolve uma teoria da força motivacional dos desejos.

³⁷ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., p. 139.

Capítulo 3 – Irrracionalidade Motivada. A perspectiva de Alfred Mele

A fraqueza de vontade torna-se particularmente fascinante se considerarmos a sua relação com importantes aspetos Filosóficos, incluindo de que forma as nossas ações intencionais devem ser explicadas, o poder do raciocínio prático e juízos avaliativos práticos, racionalidade humana e o livre arbítrio³⁸.

Aristóteles defendeu que a acrasia é um traço de carácter, exibido num comportamento intencional e livre, o qual é contrário ao melhor juízo do agente. A encrateia, segundo Aristóteles (auto controlo, continência, força de vontade) é também um traço de carácter, exibido num comportamento congruente com o melhor juízo do agente, na presença de tentações para agir em sentido contrário. Assim, para Aristóteles, um agente acrático está na condição de ser vencido, mesmo pelos prazeres que a maior parte das pessoas consegue superar, enquanto a pessoa encrática está num estado de superação dos prazeres pelos quais a maior parte das pessoas é derrotado.

Alfred Mele segue a veia Aristotélica, considerando as noções de auto controlo e acrasia numa forma mais abrangente. Para Alfred Mele, aparentemente, podemos exercer o auto controlo na superação das mais diversas tentações, ou sucumbir acráticamente a essas tentações. Existe portanto, um ponto médio entre acrasia e encrateia, não se verificando o requisito de que todas as ações acráticas manifestem acrasia³⁹.

Atualmente, as ações acráticas e encráticas são discutidas como sendo mais do que traços de carácter⁴⁰. De forma a melhor compreendermos o alcance desta noção, parece pertinente invocar a definição oferecida pelo DSM: o traço de personalidade é «a tendência para o individuo se comportar, sentir, perceber e pensar de formas relativamente constantes através do tempo e das situações nas quais o traço se pode manifestar»⁴¹.

³⁸ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 116-117.

³⁹ Uma pessoa mais controlada do que a maior parte das outras pessoas em relação ao consumo de álcool, pode, numa determinada condição, sucumbir a essa tentação. Apesar de não exibir acrasia, no sentido representado por Aristóteles, exibiu um auto controlo imperfeito que resultou numa ação acrática. Da mesma forma, alguém com um traço de carácter acrático pode resistir a uma tentação agindo de forma auto controlada.

⁴⁰ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 3.

⁴¹ American Psychiatric Association, *DSM-5*, op.cit, p. 988.

No fundo, os traços de personalidade são as tendências e/ou apetências que nos caracterizam enquanto pessoa singular e que formam certos padrões de comportamento e de agência. Quando ouvimos alguém dizer: já sabia que ias dizer isso, tu és assim! Essa pessoa está a referir-se a um determinado traço de carácter que identificou na pessoa que somos, pelo que somos. Um agente acrático exhibe esse padrão de comportamento como um traço de carácter próprio.

Ou seja, para Alfred Mele, acrasia é um traço de carácter, que exhibe um certo tipo de falha ou deficiência de força, a força do *self control*.

Mas nem todos os eventos de ação incontinente são, no entanto, a manifestação de um traço acrático, na medida em que é possível, que uma pessoa *self controlled* caia numa tentação, o que não quer dizer que o seu *self control* seja deficiente, mas antes, que o seu poder de *self control* não é perfeito.

No entanto, acrasia e *self control* não podem ser vistos como traços globais de carácter, uma vez que uma pessoa pode resistir a certos tipos de tentação e não a outros⁴².

As ações acráticas são contrárias ao melhor juízo do agente e o melhor juízo do agente contém duas características: os seus conteúdos são proposições declarativas⁴³ e é baseado nos valores e crenças do próprio agente (ou aquilo que o agente toma como sendo as suas crenças e valores).

A literatura não foca, por exemplo, a questão de como um agente livremente e intencionalmente faz *a*, ainda que julgue que, da perspectiva avaliativa dos seus pares, seria melhor não fazer *a*, ou seja, é apresentada a perspectiva avaliativa do agente e fica de fora a perspectiva avaliativa dos pares do agente, que pode ser oposta.

Ora, para Alfred Mele, é necessária uma maior atenção àquilo que é, de facto, o melhor juízo e considera que a definição de Donald Davidson, acerca do que é uma ação acrática, está comprometida, apresentando-nos o seguinte caso: um agente acredita que, *all things considered*, fazer *y* é melhor do que fazer *x*, mas acredita também que, *all things considered*, fazer *x* e *z* também é melhor do que fazer *y*. Se o agente fizer *x* e *z* preenche os requisitos descritos por Donald Davidson, mas no entanto, não está a ter uma ação

⁴² MELE, Alfred, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford University Press, New York 1987, p. 3.

⁴³ Proposições que podem ser verdadeiras ou falsas.

acrática⁴⁴. Este problema ficaria resolvido pela definição de juízo violador da ação acrática, que consiste no facto de um agente não precisar de julgar que um curso de ação *a* é o melhor entre todos os cursos de ação possíveis, mas que *a* é melhor entre as restantes opções. Portanto, melhor como um todo e não em relação a um ou outro aspeto. Se, mesmo assim, o agente livremente e intencionalmente opta por uma das outras alternativas, então está a agir acraticamente.

Assim, para Alfred Mele, um melhor juízo envolvido num juízo violador da ação acrática, é um juízo de que, no todo, é melhor fazer *a*, sendo a expressão - *melhor*, entendida exclusivamente e relativamente às opções do agente. O conteúdo desse juízo pode ser verdadeiro ou falso e o juízo é feito em função das crenças e valores do agente. Quando Alfred Mele se refere ao juízo de alguém de que é melhor fazer *a*, não deve ser inferido que o agente entende que *a* é a melhor de entre todas as possibilidades de ação, e a expressão *melhor* deve ser entendida como sendo o melhor, no geral. A expressão melhor juízo nas discussões filosóficas acerca das ações acráticas é, pelo menos, sugestiva de uma crença à qual chegámos tendo por base uma deliberação consciente⁴⁵.

3.1. A acrasia

3.1.1. Distinção de acrasia relativamente a Insanidade e Depressão

A literatura dá atenção a casos de ação acrática, nos quais, na altura da ação, os agentes estão cientes de que estão a agir contrariamente ao seu melhor juízo. Além disso, os juízos são baseados no raciocínio prático.

Estes casos são denominados, por Alfred Mele, de *core akratic action*, e define-os como sendo ações livres, sãs e intencionais que, o agente não deprimido conscientemente reconhece, no momento da ação, que são contrários ao seu melhor juízo, juízo esse que é baseado no raciocínio prático. Vemos que esta definição apela as noções de sanidade e depressão, às quais o autor prestou alguma atenção.

⁴⁴ Alfred Mele ilustra esta questão com o seguinte exemplo: o agente acha que é melhor correr do que andar de bicicleta por um quilómetro antes do pequeno-almoço, mas também acha melhor andar de bicicleta e nadar um quilómetro antes do pequeno-almoço, do que correr um quilómetro, optando por esta segunda hipótese.

⁴⁵ MELE, *Backsliding. Understanding Weakness of Will*, op. cit. pp. 4-6.

Relativamente à sanidade, Alfred Mele esclarece que apesar de se tomar como garantido, que as ações acráticas não manifestam insanidade, esta explicitação raramente é efetuada. Alguns teóricos sustentam que nenhuma ação que manifeste insanidade é livre, mas nem todos os filósofos que escrevem sobre livre arbítrio se apoiam no requisito de sanidade para a ação livre e outros assumem posições segundo as quais algumas ações que manifestam insanidade, são vistas como sendo livres⁴⁶.

Conforme vimos, a definição de insanidade, considerando esta noção como sinónimo de perturbação mental, consta no DSM como sendo:

Uma perturbação mental é uma síndrome caracterizada por perturbação clinicamente significativa, da regulação da emoção, ou do comportamento do sujeito que reflete uma disfunção dos processos psicológicos, biológicos ou do desenvolvimento subjacentes ao funcionamento mental. As perturbações mentais estão geralmente associadas com significativo mal-estar ou défice social, ocupacional, ou noutras atividades importantes⁴⁷.

Ora, considerando os critérios expostos torna-se, de facto, relevante esclarecer que a acrasia não se caracteriza como um comportamento patológico. As instâncias de acrasia, apesar de afetarem a vida do agente que as efetua, não podem ser associadas ao mal-estar ou défice social, ocupacional ou outros, do agente.

Relativamente à depressão, alguns Filósofos sustentam que se as crenças do agente de que deveriam moralmente fazer a , envolvem necessariamente a motivação para a ⁴⁸.

Com efeito existem duas teses distintas acerca desta afirmação. O cognitivismo, acerca de tais crenças, ou seja, a tese de que tais crenças são atitudes com conteúdo de verdade e nem todas essas crenças estão condenadas a serem falsas. A segunda tese diz respeito a um internalismo robusto acerca das crenças do dever moral da primeira pessoa. Segundo esta tese, necessariamente qualquer crença que um sujeito possua, moralmente para A , engloba motivação para A ⁴⁹.

Para Alfred Mele, não existe teoria moral plausível que implique que crenças morais de dever fazer (entendidas como atitudes com conteúdo de verdade valorizado) impliquem necessariamente motivação para agir em conformidade. Na defesa desta teoria Alfred

⁴⁶ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 7-9.

⁴⁷ American Psychiatric Association, *DSM-5*, op.cit, p. 22.

⁴⁸ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 9.

⁴⁹ MELE, Alfred, *Motivation and Agency*, Oxford University Press, New York 2003, p. 107.

Mele recorre à noção de apatia, no sentido relacionado com depressão clínica. Nestes casos, o agente sabe aquilo que é melhor fazer, apesar de não ter motivação para fazê-lo⁵⁰.

Alfred Mele defende que, considerando a forma que a apatia se poderá aplicar aos seres humanos, o agente que sofre deste problema pode acreditar, no sentido cognitivista, que deveria moralmente fazer *A* sem no entanto ter motivação para *A*. Ou seja, as atitudes que englobam essencialmente motivação para *A*, são distinguíveis das atitudes que contingentemente o fazem: uma atitude de um agente *S* essencialmente engloba motivação para *A*, se e só se, englobar motivação para *A* não apenas na situação atual do sujeito *S*, mas em todos os possíveis cenários nos quais *S* tem essa atitude e contingentemente engloba motivação para *A*, se e só se, apesar de na situação atual de *S*, englobar motivação para *A*, este não for o caso em possíveis cenários nos quais *S* tem essa atitude⁵¹.

Para Alfred Mele, uma teoria moral não efetua exigências legítimas metafísicas ou conceptuais numa teoria da motivação humana e, mais especificamente, assumindo o cognitivismo moral, uma teoria moral plausível, não pode requerer que as crenças de dever moral essencialmente englobem motivação para agir nessa conformidade⁵².

Os casos de ação acrática a que Alfred Mele se refere não incluem depressão⁵³.

Invocando de novo o DSM, as perturbações depressivas, incluem as perturbações perturbação de desregulação do humor disruptivo, perturbação depressiva major, perturbação depressiva persistente, perturbação disfórica pré menstrual, perturbação depressiva induzida por substância/medicamento, perturbação depressiva devida a outra condição médica, perturbação depressiva com outra especificação e perturbação depressiva não especificada. A característica comum a todas estas perturbações é a presença de tristeza, vazio ou humor irritável, acompanhada de alterações somáticas e cognitivas que afetam de forma significativa a capacidade de funcionamento do indivíduo⁵⁴.

⁵⁰ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 9-10.

⁵¹ MELE, *Motivation and Agency*, op. cit., pp. 125-126.

⁵² MELE, *Motivation and Agency*, op. cit., p. 130.

⁵³ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 10.

⁵⁴ American Psychiatric Association, *DSM-5*, op.cit, p. 183.

Mais uma vez, torna-se importante a ressalva de que as ações acráticas tratadas por Alfred Mele não incluem a depressão, pelos motivos expostos para a noção de insanidade.

3.1.2. A fraqueza de vontade

Alfred Mele apresenta a noção de *core weak-willed action* como sendo: «*free, sane, intentional action that, as the nondepressed agent consciously recognizes at the time of action, is contrary to his better judgment based on practical reasoning*»⁵⁵.

Uma característica que a noção de *core akratic action* ou *strict akratic action*⁵⁶ partilha com as noções tradicionais de ação acrática, é o facto de ambas serem contrárias ao melhor juízo do agente. Mas, Alfred Mele defende que as noções tradicionais deveriam acomodar alguns casos de ação acrática, nos quais o agente não age contrariamente ao seu melhor juízo, antes age em conformidade com esse melhor juízo⁵⁷.

Advoga que as concepções tradicionais da ação acrática, bem como o seu inverso, rondam uma espécie de compromisso - *commitment* com a ação, ou seja, um compromisso constituído por um melhor juízo decisivo.

Portanto, o agente fica racionalmente comprometido com um sentido de ação congruente com aquele melhor juízo, de tal forma que, o facto de fazer algo no sentido contrário o coloca no caminho da irracionalidade, mas, no entanto, uma ação deste tipo seria subjetivamente irracional, porque seria irracional do ponto de vista do agente. Ou seja, tendo o agente o juízo, não pode racionalmente, do seu ponto de vista, ter motivos para não efetuar a ação. E é neste sentido que Alfred Mele distingue dois tipos de compromisso prático: o avaliativo e o executivo⁵⁸.

⁵⁵ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 57.

⁵⁶ Para Alfred Mele, as duas noções são idênticas, devendo-se a diferença apenas a um refinamento das definições. Em *Irrationality* Alfred Mele refere-se a *Strict Akratic action*. No capítulo 3 de *Backsliding*, o termo *core weak-willed action*, substitui o termo *core akratic action*.

⁵⁷ Para ilustrar esta questão apresenta-nos Rocky, que prometeu à mãe que não jogaria futebol, mas que, contrariamente ao seu melhor juízo, aceitou um convite de alguns amigos e decide de facto ir jogar. No entanto, quando chega o momento do jogo, acaba por não aparecer, não por entender que é melhor não jogar, mas porque ficou com medo. Ou seja, ele não jogaria, mesmo que julgasse que seria melhor fazê-lo. O facto de Rocky não ter jogado pode ser entendido como fraqueza de vontade ou acrasia mesmo que achasse que seria melhor não jogar. Com este exemplo pretende argumentar que as concepções tradicionais de ação acrática devem ser latas, de forma a acomodar casos como o de Rocky.

⁵⁸ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 8-16.

O compromisso avaliativo coloca questões tais como, o que seria melhor fazer (do ponto de vista dos meus próprios desejos e crenças), enquanto o compromisso executivo coloca questões de carácter prático, tais como, o que vou fazer agora.

Alfred Mele prefere que as concepções de juízo de ações continentes e incontinentes sejam vistas à luz dos compromissos práticos. Aquilo que se verifica comum nos casos de ação acrática contra um melhor juízo decisivo e na ação acrática de acordo com tal juízo (ação acrática não-ortodoxa⁵⁹) é o compromisso prático, frustrado por uma motivação competitiva (não coerciva). No primeiro caso, a ação acrática tradicional, o compromisso é avaliativo, no segundo caso, ação acrática inortodoxa, ilustrada por Bruce quando desiste de arrombar a casa do vizinho, devido ao seu receio, é um compromisso executivo⁶⁰.

Em suma, uma concepção de ação acrática deve acomodar ambos os compromissos práticos. O compromisso prático avaliativo subjaz à ação acrática tradicional, enquanto uma ação acrática inortodoxa terá o carácter do compromisso prático executivo.

No caso das ações encráticas (contrário da ação acrática), podemos observar um paralelo com a situação anterior, considerando os dois tipos de compromisso prático: numa ação encrática ortodoxa, o agente age com base no comprometimento avaliativo face às tentações. Na ação encrática inortodoxa, o agente falha um comprometimento avaliativo para fazer *A*, mas tem um comprometimento executivo e a sua decisão ou intenção para *A*, que constitui esse comprometimento, prevalece contra a ameaça de ser abalada de tal forma que, se a decisão ou intenção for abalada naquele sentido, o facto do agente se refrear de fazer *A* seria não-ortodoxamente acrático⁶¹.

⁵⁹ Para exemplificar a ação acrática não ortodoxa Alfred Mele retrata um agente, Bruce, que decide arrombar a casa do vizinho, para ser aceite num determinado grupo de amigos, apesar de julgar que, consideradas todas as coisas, seria melhor não o fazer. Mas no último minuto desiste do crime. A desistência pode ter ocorrido porque a) o melhor juízo prevaleceu, b) porque não teve coragem e, neste último caso, exibe fraqueza de vontade. Mas suponha-se que Bruce começou a sentir alguma apreensão e teme morrer nesta situação. Se apesar de julgar ser melhor não participar no crime Bruce vence o seu medo e de facto participa, então autocontrolou-se, venceu o medo e exibiu então exibiu força de vontade. Logo, Alfred Mele conclui que, alguns exercícios de auto controlo não são, aparentemente, levados a cabo para servir um melhor juízo.

Este agente exibe ambos os compromissos práticos: avaliativo e executivo. Estes compromissos conflituam entre si: o compromisso avaliativo deriva do seu melhor juízo que o compromete a não arrombar a casa, mas a sua decisão compromete-o a participar no crime.

⁶⁰ MELE, Alfred, «Akrasia, Self-control, and Second-Order Desires», *Noûs*, 26, nº 3 (1992).

⁶¹ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 16- 18.

3.1.3. Os desejos de segunda ordem e o compromisso prático

Supondo que os compromissos práticos se esgotam em comprometimentos avaliativos e executivos, os desejos, considerados como classe, não constituem compromissos práticos, na medida em que esses desejos não resolvem questões tais como: o que é melhor fazer, ou, o que devo fazer. E esta resolução é essencial para os compromissos práticos.

Para ilustrar esta questão Alfred Mele descreve a situação de Sally, uma fumadora que apesar de genuinamente gostar de fumar, está convencida de que seria melhor não o fazer. Suponhamos que Sally tem um desejo de segunda ordem de não desejar fumar. Suponhamos ainda que possui um desejo de segunda ordem, competitivo, de continuar a desejar fumar. Com efeito, apesar de acreditar que se não desejasse fumar, deixaria de encontrar prazer nessa atividade, Sally também pensa no facto de perder a oportunidade de praticar a atividade de fumar e obter o prazer associado, chegando até a imaginar o que poderia tomar o lugar desse prazer. Portanto, fumar faz parte da rotina de Sally e tem um importante lugar na sua vida, uma vez que fuma não apenas pelo prazer, mas para se acalmar e manter alerta, factos que a ajudam nas suas tarefas diárias. Verifica-se assim que, mesmo ao nível dos desejos de segunda ordem, Sally enfrenta um conflito e o desejo, por si, não constitui um compromisso prático.

Por outro lado, os desejos de segunda ordem também não determinam o que é racional, e portanto não constituem um compromisso avaliativo. Ao decidir fazer *A* ou adquirindo uma intenção para *A*, o sujeito envolve-se num assunto prático, adquirindo um compromisso executivo. O desejo de segunda ordem de Sally, não associa ou compromete a sua mente com o que vai fazer, e deixa, portanto, o assunto por resolver.

Portanto, para Alfred Mele, os desejos de segunda ordem, como uma classe, não são compromissos avaliativos⁶², nem são compromissos executivos, compromissos estes, melhor compreendidos como formando disposições para fazer decisões correspondentes e adquirir intenções também correspondentes⁶³. E, finalmente, os desejos de segunda

⁶² O desejo de Sally de desejar continuar a fumar, por exemplo, não a compromete racionalmente com o desejo de continuar a fumar.

⁶³ Por exemplo, o desejo de um fumador de não desejar fumar, pode motivá-lo a intencionar uma ação para esse fim, motivando-o a decidir o melhor que puder para convencer-se de que já não deseja fumar.

ordem, como uma classe, não resolvem questões do que é melhor fazer ou o que devo fazer, logo, não constituem um compromisso prático. Neste sentido, diferem significativamente de juízos avaliativos decisivos, decisões e intenções.

Por estes motivos, o autor sugere que a concepção que envolve juízos de ações continentais e incontinentais deve ser moldada em termos de compromissos práticos, independentemente da noção de desejos de segunda ordem (ou de ordem superior).

Na ação acrática, Alfred Mele defendeu de que forma intervêm os compromissos práticos avaliativo e executivo, conforme descrito no ponto anterior, mas e no comportamento continente?

As ações continentais que estão em desacordo com o melhor juízo decisivo do agente, contam como sendo uma ação continente, devido ao facto do compromisso executivo ter prevalecido contra motivação de outro tipo.

Em primeiro lugar, Alfred Mele considera que, qualquer concepção de comportamento continente carece de uma concepção de continência apropriada.

Tradicionalmente, a continência é tratada como um traço, constituído pela capacidade de restringir ou conter desejos que ameaçam o agente de seguir a direção conforme o seu melhor juízo, mas, Alfred Mele introduz a noção de auto controlo e descreve três formas de olhar para esta noção: como sinónimo de continência; num sentido mais lato, no qual o auto controlo pode ser exibido, mesmo num comportamento que vá contra aquilo que o agente julga melhor; num sentido mais estreito, de acordo com o qual, o auto controlo é exibido apenas no comportamento motivado por um desejo ou intenção de controlar uma condição motivacional do agente bem como os resultados comportamentais apropriados de tal comportamento.

Neste sentido, parece plausível sustentar que o comportamento intencional, que exhibe auto controlo, é motivado, pelo menos em parte, por atitudes motivacionais de ordem superior.

Assim, esta terceira noção de auto controlo, faz da motivação de ordem superior condição essencial para o comportamento intencional, na medida em que agir continuamente para *t*, requer possuir um desejo relevante de segunda ordem, nessa altura⁶⁴.

⁶⁴ MELE, «Akrasia, Self-control, and Second-Order Desires», art. cit., p. 295.

Parece razoável efetuar um paralelismo entre desejos de segunda ordem e graus de ponderação ou de deliberação (refiro-me ao tempo e esforço que o agente está disposto a investir na reflexão da deliberação, que irá definir a ação a desempenhar), na medida em que os desejos de segunda ordem seriam o elemento que o agente considera importante controlar, ou, usando a terminologia de Alfred Mele, auto controlar-se, na sua deliberação acerca de qual ação intenciona desempenhar. Ou seja, o desejo de segunda ordem, devidamente auto controlado pelo agente, poderá determinar a ação oposta ao desejo de primeira ordem do agente. Este oposto poderia resultar em duas ações encráticas diferentes, ou mesmo em duas ações contrárias: uma ação continente, se o agente considerar o desejo de segunda ordem e auto controlar-se de forma a não agir em função do primeiro, ou uma ação acrática, caso o agente não delibere em função do desejo de segunda ordem, ou seja, não pondere exercer o seu poder de auto controlo. Neste último caso, o sujeito poderá mesmo não investir tempo no esforço de refletir acerca da melhor decisão.

3.1.4. Fraqueza e compulsão

Alfred Mele preocupou-se em distinguir ações derivadas de fraqueza de vontade de ações em que os agentes são obrigados ou compelidos (e portanto, não livremente), a fazer. Recorde-se que Alfred Mele definiu uma ação *core weak-willed*, como sendo livre (e portanto não compelida), sã, intencional, que o agente não deprimido reconhece conscientemente, na altura da ação, ser contrária aos seu melhor juízo, juízo esse baseado no raciocínio prático⁶⁵.

O autor apresenta formas de efetuar a distinção entre fraqueza e compulsão, através de recursos compatibilistas e libertários, mostrando alguns dos recursos de que dispõem os defensores da ação livre.

Alfred Mele recorda Jonh Fisher e Mark Ravizza que defendem a tese de que um agente é moralmente responsável por uma ação, na medida em que cria em si próprio um mecanismo *moderately reasons-responsive* - responsividade a razões moderada.

⁶⁵ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 33.

O mesmo pode ser dito do facto de um agente efetuar livremente uma ação, na medida em que o agente torna o mecanismo dele próprio, ao assumir a responsabilidade pelo que fez.

A *moderate reasons-responsiveness* – responsividade a razões moderada - consiste numa *regular reasons-receptivity* – recetividade a razões moderada - e, pelo menos uma *weak reasons-reactivity* – fraca reatividade à razão - do mecanismo de sequência atual, que conduzirá a uma ação.

A *Reason-receptivity* – recetividade a razões - é a capacidade de reconhecer razões existentes. A *reasons-reactivity* – reatividade às razões - a capacidade de traduzir razões em escolhas e, por conseguinte, num comportamento subsequente.

A definição característica de uma *regular reasons-receptivity* – recetividade a razões regular, envolve a compreensão de um padrão de *reasons-receptivity* – *recetividade a razões*, atuais ou hipotéticas.

O mecanismo de um agente que leva à ação, no mundo atual, é pelo menos um *weak reasons-reactive* – fraca reatividade a razões, uma vez que existe um mundo possível com as mesmas leis, no qual um mecanismo deste tipo é operativo no agente, que diz que existe razão suficiente para fazer de outro modo. E o agente reconhece esta razão e o agente faz de outro modo, por esta razão.

Portanto, pode ser sugerido que a *core weak-willed action* difere da ação compelida, contrária ao juízo do agente, uma vez que, enquanto a ação *weak-willed* provém do mecanismo de *moderately reasons-responsive* - responsividade a razões moderada do próprio agente, a ação compulsiva, não. Apesar de, em ambos os cenários os agentes satisfaçam as condições para uma *regular reasons receptivity* - recetividade a razões regular, o mecanismo do agente que age de forma *weak-willed* é pelo menos *weakly reasons-reactive* – reativo às razões de forma fraca - e o mecanismo do agente compulsivo não é. No caso deste último agente, não existe um mundo possível, no qual faça de outro modo. Pode referir-se ser este precisamente o caso de um desejo irresistível revestido de força compulsiva.⁶⁶.

A explicação de Alfred Mele alinha-se com a definição de compulsão descrita do DSM:

⁶⁶ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 49-50.

Comportamentos repetitivos (por exemplo, lavagem das mãos, ordenação, verificação) ou atos mentais (por exemplo rezar, contar, repetir palavras em silêncio) que o indivíduo se sente compelido a realizar em resposta a uma obsessão, ou de acordo com regras que devem ser aplicadas rigidamente. Os comportamentos ou atos mentais destinam-se a impedir ou reduzir a ansiedade ou o mal-estar, ou a prevenir um evento ou situação receados; porém, estes comportamentos ou atos mentais não estão ligados de forma realista com aquilo que se destinam a neutralizar ou prevenir, e são claramente excessivos⁶⁷.

O mecanismo do agente compulsivo envolve comportamentos para os quais se sente compelido ou obrigado a fazer, de forma a aliviar o seu mal-estar ou ansiedade e não se constitui como um mecanismo de responsividade a razões moderada, mas antes, são comportamentos claramente excessivos e, poderemos inferir, prejudiciais ao seu normal funcionamento quotidiano, motivo pelo qual a compulsão é considerada patológica. O agente compulsivo, de facto, não pode fazer de outro modo.

3.1.5. Melhor Juízo

A noção de melhor juízo do agente deve considerar a sua função de guia da conduta. Alfred Mele reforça aqui a sua conceção de que os conteúdos do melhor juízo são proposições declarativas e que é necessário reter a ideia de que o melhor juízo tem um conteúdo de verdade. Mas, apesar de serem uma guia da conduta, também podem estar sujeitos ao fracasso, nos casos de ações *core weak-willed*. O autor, começa por fazer a distinção entre o melhor juízo do que é melhor fazer agora (proximal) e o melhor juízo do que será melhor fazer depois (distal).

Analogamente, existem as intenções proximal e distal. Se considerarmos um agente que está a deliberar acerca do que fazer agora, então um melhor juízo é, por definição, um juízo consciente do efeito de que é melhor fazer *a*, agora, emparelhado com a intenção consciente proximal para *a*. Mas, a presença de uma intenção consciente proximal não assegura a sua execução com sucesso ou sequer a tentativa de execução⁶⁸. Por outro lado, consideremos o exemplo formulado por Alfred Mele: se Alex concebe uma intenção proximal consciente de se cortar no âmbito de uma cerimónia, mas, por breves instantes,

⁶⁷ American Psychiatric Association, *DSM-5*, op.cit, p. 976.

⁶⁸ Por exemplo, tenho a intenção de atender o telefone que está a tocar neste momento, mas distraio-me com algo que me desviou a atenção e acabo por não atender o telefone.

sente uma aversão a esse ato, e pára de o executar intencionalmente, então, significa que a intenção inicial foi suplantada por uma intenção efetiva de parar. Mas se Alex não consegue, intencionalmente, parar de se cortar enquanto a intenção proximal consciente subsiste, então, Alex não consegue intencionalmente parar de se cortar, enquanto a sua intenção proximal consciente de se cortar, persiste.

Assim, e considerando a definição de melhor juízo que está em questão, se neste caso Alex julga ser melhor cortar-se, nesse momento, o comportamento intencional contrário ao seu melhor juízo não está disponível para Alex, enquanto o juízo permanece, porque a presença continuada do juízo, requer uma presença continuada de uma intenção proximal consciente de se cortar.

Ou seja, fica identificado um tipo específico de melhor juízo, constituído de tal forma que, enquanto persiste, não permite um comportamento *weak-willed* contra esse melhor juízo, porque não está em aberto para o agente.

Um melhor juízo apenas o é, relativamente ao que fazer agora, adicionando a correspondente intenção consciente⁶⁹.

Em *Irrationality*, Alfred Mele usa o termo *decisive better judgement*, como sendo o juízo feito com base na consideração da totalidade dos desejos e crenças pertinentes do agente. Trata-se do juízo que, na opinião do agente, constitui razão suficiente para a ação. Este melhor juízo decisivo difere do *better judgement*, efetuado apenas do ponto de vista de algum segmento do sistema de valores do agente, o qual não é considerado por ele, como dominante/principal⁷⁰.

3.1.6. Ação Intencional

3.1.6.1. Intenções Proximal e Distal e Ação Intencional

Para Alfred Mele, algumas das nossas intenções referem-se a um futuro não imediato e outras não. Na terça-feira, posso ter a intenção de ir buscar a Ângela ao aeroporto no sábado e posso ter a intenção de ligar ao Nick, agora. A primeira intenção estima uma ação dois dias, futuramente – intenção distal. A última intenção trata-se do que fazer, agora – intenção proximal.

⁶⁹ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 57-60.

⁷⁰ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 6.

As ações proximais incluem ainda intenções de continuar a fazer algo que estou a executar agora e intenções para começar a fazer *A*. Algumas intenções incluem os aspetos proximal e distal – intenção mista. Por exemplo, o Al pode ter a intenção de correr 1 km sem parar, a partir de agora, e estima que a ação vá durar aproximadamente seis minutos. Esta intenção especifica algo que o agente vai fazer agora e também mais tarde.

Pelo menos algumas intenções destes três tipos, são intenções ocorrentes. A minha intenção formada na terça-feira de ir buscar a Ângela ao aeroporto no sábado, levou-me a tomar uma nota no calendário, e essa foi a intenção ocorrida nesse momento. Além disso, em nenhum momento eu estive consciente dessa intenção, porque foi uma intenção ocorrente. Se a minha intenção proximal de telefonar ao Nick me levou a marcar o seu número de telefone, então trata-se de uma intenção ocorrente nesse momento. A intenção mista do Al, que o levou a iniciar de imediato a sua corrida, foi também uma intenção ocorrente no momento.

Existem perspectivas Filosóficas competitivas acerca da relação exata entre itens tais como intenções e estados cerebrais. Sobre este assunto Alfred Mele levanta a seguinte hipótese, acerca das ações intencionais, que entende ser coerente com uma grande variedade de visões fisicalistas desta relação: H - sempre que os seres humanos produzem uma ação intencional manifesta⁷¹, pelo menos um dos seguintes aspetos desempenha um papel causal na sua produção: alguma intenção, a aquisição ou persistência de alguma intenção, o correlato físico⁷² de um ou mais destes itens⁷³.

3.1.6.2. Intenções Conscientes

Para se considerar que as ações ocorrentes desempenham um papel na produção das ações intencionais, devem as ações ocorrentes ser conscientes para cada um dos itens: intenções ocorrentes, aquisição de tais intenções, ou correlatos físicos de alguma destas coisas? Pode um agente ter, num tempo *t*, uma intenção ocorrente para *A* sem estar

⁷¹ Ação manifesta – *overt action* – significa uma ação que envolve essencialmente movimento corporal periférico.

⁷² A relação de correlato físico não é uma relação causal. Existem outras opções, incluindo a relação de identidade ou vários tipos de superveniência.

⁷³ MELE Alfred, *Effective Intentions. The Power of Conscious Will*. Oxford University Press. New York 2009, pp. 10-13.

consciente da mesma, ou parte dela, nesse momento t e sem estar consciente, no momento t , que tenciona fazer A ?

Alfred Mele defende que sim, é possível.

O autor faz a distinção entre consciência tipo 1 - *awareness* e consciência preditiva – *predictive awareness* - tipo 2. Se eu for a caminhar numa rua cheia de pessoas e a pessoa que vai à frente parar de repente, eu fico consciente de que vou chocar com essa pessoa, ou seja, eu prevejo o choque entre mim e a pessoa. Com este exemplo, Alfred Mele elucida o tipo 2, a consciência preditiva, baseada na observação de eventos externos. E a minha consciência – *awareness* – de que vou chocar com aquela pessoa é aquilo a que podemos chamar de consciência preditiva – *predictive awareness*.

Noutro cenário, imaginemos que cheguei a casa e preparo-me para abrir a porta de entrada. Estou pois, consciente que estou prestes a abrir a porta.

Fazendo uma leitura de intenção consciente, o facto de eu ter uma intenção proximal de abrir a porta, em conjunto com a consciência tipo 1 – *awareness* – de que estou prestes a abrir a porta, é suficiente para eu ter uma intenção proximal consciente e esta situação difere da consciência preditiva ou tipo 2.

Alfred Mele defende que é implausível considerar que todas as intenções proximais são conscientes.

Uma condição disjuntiva alegada suficiente, para um agente ter a intenção proximal consciente para A , pode ser construída a partir de várias condições suficientes: o agente estar consciente da sua intenção proximal para A ou alguma parte do seu conteúdo; o agente estar consciente de que tenciona fazer A agora; o agente ter uma intenção proximal para A juntamente com o tipo 1 de consciência – *awareness* - de que está prestes a fazer A .

Outra disjunção pode ainda ser considerada: uma intenção proximal para A pode ser tratada como uma intenção proximal consciente, num momento, se nesse momento, o agente conscientemente faz a decisão proximal para A ⁷⁴.

⁷⁴ MELE, *Effective Intentions*, op. cit., pp. 26-28.

Ora, parece razoável efetuar um paralelismo com a visão de António Damásio, quando este refere que, em muitos momentos da nossa vida, as nossas ações são controladas por processos não conscientes, por exemplo, quando conduzimos um carro.

Para António Damásio, a influência do processamento não consciente nas nossas ações, é uma realidade positiva, pois quando, por exemplo, caminhamos para casa, podemos estar a pensar em algum problema que tenhamos que resolver e não concentrados no caminho que estamos a fazer. Por outro lado, os processos não conscientes podem ser moldados pelo comportamento consciente e, por este motivo, defende que somos responsáveis pelas nossas ações, ainda que executemos firmemente essa ação de forma não consciente. Isto, porque os processos não conscientes vão sendo adquiridos ao longo da nossa formação, uma vez que, ao longo da infância e adolescência vamos transferindo o controlo consciente para processos inconscientes, criando um espaço cerebral, não consciente, capaz de agir em função das intenções e objetivos conscientes. Assim, os mecanismos não conscientes libertam a mente consciente para que se ocupe da análise de situações e planeamento futuro.

No entanto, também é verdade que os resultados da deliberação consciente são limitados por predisposições não conscientes, quer ao nível biológico quer ao nível de condicionamento social e, neste caso, o controlo não consciente da ação é um importante aspeto a considerar.

No entanto, a maior parte das decisões importantes são assumidas muito antes da sua execução na mente consciente e, portanto, passíveis do controlo consciente capaz de minimizar os efeitos das predisposições não conscientes.

António Damásio salvaguarda a possibilidade de que as interações dos mecanismos conscientes e inconscientes podem desajustar-se perante os apetites e os desejos. E neste caso, um mecanismo de negação bem-sucedida pode ser a chave para a resolução desse desajuste, mas a capacidade de usar eficazmente esse mecanismo exige uma preparação consciente demorada⁷⁵.

Considerarmos o alinhamento dos mecanismos inconscientes, com as intenções proximais conscientes, parece sugerir um contexto de harmonia da mente/corpo de cada

⁷⁵ DAMÁSIO, António, *O Livro da Consciência*, Circulo de Leitores, 2010, pp. 331-336.

agente, uma vez que, a consciência preditiva de que nos fala Alfred Mele, considerada como um mecanismo, é desencadeada precisamente no sentido de tornar a nossa vida mais prática, bem como parece suportar o comportamento e a intenção racionais, sendo esta racionalidade o suporte dos melhores juízos do agente.

Ao mesmo tempo, os mecanismos favorecem a poupança dos recursos cognitivos, que podem ser usados em tarefas de complexidade superior e que requerem o nosso raciocínio consciente para tomadas de decisão e planeamento futuro.

Mas fica também ressalvada a possibilidade do desajuste entre estes mecanismos, facto que parece abrir a porta para a possibilidade da irracionalidade motivada, em concreto, da ação acrática.

3.1.6.3. Perspetivas Motivacional e Intelectual na explicação da ação Intencional

Uma grande fonte de interesse relativamente à ocorrência das *strict akratic actions* reside no facto deste fenómeno levantar questões difíceis, acerca da relação entre pensamento e ação, relação essa, de primordial importância para qualquer teoria explicativa do comportamento intencional, capaz de harmonizar pensamento e papel explicativo.

E esta questão torna-se complicada perante a existência de duas perspetivas para a explicação da ação intencional, nos campos da Filosofia da Ação, Psicologia Moral e no pensamento comum: a perspetiva motivacional e a perspetiva intelectual⁷⁶.

A primeira relaciona-se com o facto de que aquilo que o agente faz quando age, está estreitamente relacionado com aquilo que o motiva mais fortemente na altura. Nesta conceção, e no caso das ações intencionais, conhecer o porquê do agente estar naquelas condições motivacionais seria importante, para perceber porque agiu de determinada forma. Apesar de algumas vezes ser assumido que a relação entre o que os agentes estão mais motivados a fazer em determinada altura e o que tentam fazer nessa altura é determinista, a noção de força motivacional não tem que pressupor determinismo (nem determinismo global, nem determinismo local, acerca do funcionamento interno do agente) e, nesta linha, há autores libertários inclinados a pensar que alguns desejos terão

⁷⁶ MELE, *Autonomous Agents*, op.cit., p. 16.

mais poder motivacional do que outros. Mas talvez um agente consiga resistir a um desejo mais forte, agindo em função de um desejo menos forte e talvez a relação entre desejos e ações seja indeterminista de tal forma, que apenas exista a probabilidade de um agente agir em função do desejo mais forte, entre dois desejos concorrentes.

A perspetiva motivacional aplica-se a todos os agentes, independentemente da espécie biológica, enquanto a perspetiva intelectual, é aplicável apenas aos humanos, possuidores de intelecto prático. A perspetiva intelectual consegue pesar opções e fazer juízos acerca do que é ideal, melhor, ou suficientemente bom para fazer. E a ideia central à perspetiva intelectual é precisamente o facto de estes juízos possuírem um papel significativo na explicação de algumas ações intencionais.

No âmbito da ação humana intencional, uma forma de combinar estas perspetivas é considerar que, nos agentes intelectuais, a força motivacional e o juízo avaliativo andam par a par⁷⁷.

Para Alfred Mele as *core weak-willed actions* são melhor explicadas pela perspetiva motivacional: um agente age de forma *weak-willed*, se está mais fortemente motivado para isso naquele momento. Na perspetiva intelectual estas ocorrências *weak-willed* parecem mais difíceis de explicar.

Alfred Mele entende que, uma combinação apropriada das duas perspetivas tem que mencionar alegados intermediários entre motivação e juízo, por um lado, e a ação intencional por outro lado, em particular decisão (ou escolha) e intenção.

As perspetivas motivacional e intelectual, na explicação da ação intencional humana, convergem não apenas na ação intencional evidente ou óbvia, mas também na intenção e decisão, na medida em que decidir fazer *A*, é realizar uma ação mental momentânea de formar a intenção para *A*.

Portanto, propõe uma visão focada no melhor juízo, produzido pelo raciocínio prático avaliativo⁷⁸.

Conforme vimos, Alfred Mele apresenta a sua própria conceção de melhor juízo, bem como explica o seu lugar na produção da ação. Se, conforme é geralmente aceite, os

⁷⁷ O juízo determina causalmente a força motivacional, ou a força motivacional determina causalmente o juízo ou juízo e força motivacional possuem uma causa comum.

⁷⁸ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 60-62.

julgamentos avaliativos de agentes auto controlados e autónomos desempenham um papel preponderante no seu comportamento, então vamos, com certeza, querer entender esse papel⁷⁹.

Consideremos que um agente detém conscientemente, num momento t , um juízo de que é melhor A , no momento t , se e só se, no momento t , o agente conscientemente detém o juízo de que é melhor fazer A no momento t e, simultaneamente, o agente conscientemente tenciona fazer A no momento t . Assim, se num momento t , um agente conscientemente detém um julgamento de que é melhor A , imediatamente, o agente tenciona fazer A imediatamente. O agente tem uma intenção para o presente ilusório, ou seja, uma intenção proximal.

No entanto, as intenções proximais não garantem a sua execução bem-sucedida. O melhor juízo é apenas o melhor juízo, adicionado à intenção correspondente.

Resta portanto perceber se é possível conceber um melhor juízo que esteja ligado estreitamente com intenção⁸⁰. Ou seja, a questão que se coloca é: que influência tem o melhor juízo na intenção?

Se nos focarmos nos juízos, teremos que trabalhar duas noções: os conteúdos dos juízos e a atitude avaliadora desses conteúdos.

Alfred Mele distingue a crença do sujeito de que é melhor fazer A , do correspondente juízo, da seguinte forma: a crença é o género e o melhor juízo é a variedade. Especificamente, o juízo de que fazer A é melhor, corresponde à crença tipo B de que seria melhor fazer A : a crença tipo B , por definição, é uma crença adquirida de forma não desviante, como uma conclusão do raciocínio prático avaliativo da primeira pessoa, e esse raciocínio é entendido como um processo inferencial, que envolve premissas avaliativas, guiadas pela motivação de estabelecer o que fazer.

Este melhor juízo difere de outros juízos de que é melhor fazer A , essencialmente, devido à sua etiologia. A motivação que guia a inferência prática avaliativa é também a motivação para intencionar de acordo com a conclusão avaliativa.

⁷⁹ MELE, *Autonomous Agents*, op. cit., p. 15.

⁸⁰ MELE, *Autonomous Agents*, op. cit., pp. 19-20.

Assim, todos os melhores juízos, no sentido em questão, têm uma dimensão motivacional, fundamentada na etiologia dos juízos.

Importa destacar a característica importante dos melhores juízos serem não desviantes e derivados da inferência avaliativa prática, devido à qual, estão preparados para funcionar na produção das intenções.

Além disso, os melhores juízos comprometem o agente racionalmente para a ação. Por isso, a ação acrática contra tais juízos seria vista, na perspectiva do agente, como sendo irracional.

Claro que nada do que foi dito implica que a formação do melhor juízo assegura a formação ou aquisição da correspondente intenção. A condição motivacional do agente, que chega ao melhor juízo a favor de fazer *A*, parcialmente com base na sua avaliação dos objetos de desejos pertinentes pode ser tal, que frustra a transição do melhor juízo para a intenção. Porque, a força motivacional dos desejos do agente pode estar desalinhada com a sua avaliação dos itens desejados.

Um agente normal, na ausência de oposição, ao julgar que é *simpliciter* melhor fazer *A*, adquire diretamente a intenção de *A*, não sendo necessário nenhum esforço de auto controlo. Quando um agente acriticamente falha o intencionar de acordo com o seu juízo decisivo, então encontrou oposição significativa, ou seja, algo bloqueou a transição que deveria ter sido direta.

Vejamos de que forma esta situação se concretiza⁸¹.

3.1.7. Raciocínio prático e raciocínio avaliativo

Para Alfred Mele, o melhor juízo é produzido pelo raciocínio prático avaliativo, considerando que este processo é de natureza cognitiva, que envolve premissas de avaliação e é em parte guiado pela motivação de decidir o que fazer. E esta motivação dispõe o agente para ter uma intenção de acordo com a conclusão avaliativa racional, sendo que esta disposição suporta o propósito do raciocínio prático avaliativo, que é chegar a uma resolução satisfatória para um problema prático.

⁸¹ MELE, *Autonomous Agents*, op. cit., pp. 25-27.

O caminho para passar de um melhor juízo (produzido pela razão prática avaliativa) para uma intenção (congruente com esse juízo), é um caminho *by default*, ou pré programado, em analogia ao *software* de um computador. Ou seja, na ausência de determinadas condições, por exemplo, fortes desejos opostos, o funcionamento normal dos seres humanos, leva a que, julgar ser melhor *A*, pode sistematicamente levar a uma intenção para *A*.

Existe portanto, uma relação direta entre um melhor juízo de *a* e a intenção para *a*. Em casos simples, nos quais não se verifica oposição motivacional, ou esta se verifica em pequeno grau, esta transição do juízo para a intenção é fácil, porque não é necessário nenhum esforço ou exercício de *self control*, para a produção da intenção. O oposto verifica-se, quando o agente encontra oposição significativa, e, neste caso, serão necessários esforços adicionais.

Portanto, quando encontramos uma falha *weak-willed* para ter uma intenção de acordo com o melhor juízo do agente, significa que se verificou algum tipo de oposição, que bloqueou a transição existente *by default*⁸².

A aquisição de uma intenção para *A* é, portanto, a condição *default* na formação de um juízo de que é melhor *A*.

Este caminho por *default* não é o único existente entre melhores juízos e a intenção resultante, conforme veremos mais adiante.

O procedimento *by default* ajuda a explicar a transição normal do melhor juízo para a intenção, porque neste caso não é necessária intervenção consciente nem exercícios de autocontrole, por parte do agente.

Por vezes, a transição de julgar melhor *A* é facilmente efetuada para decidir que *A*, mas porque ao decidir que *A* o sujeito forma uma intenção para *A*, decidir que *A* não é uma etapa intermédia entre julgar e adquirir a correspondente intenção. Uma transição suave do julgamento para a decisão é a transição suave do julgamento para a intenção. O inverso é falso, porque decidir é um modo de aquisição de intenção, modo esse que é ativo. Mas as intenções podem ser adquiridas passivamente, ou seja, são frequentemente adquiridas

⁸² MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 57-65.

independentemente de uma ação ou formação de intenção⁸³. Os requisitos para decidir, são mais fortes do que para intencionar.

Adquirir uma intenção com base num melhor juízo, constitui um progresso genuíno em direção à ação. O procedimento por *default* deve ser entendido como constituindo uma disposição para estabelecer cursos de ação julgados melhores, imediatamente após a formação do julgamento. Uma intenção para *A* também inclui um plano para fazer *A*.⁸⁴

Alfred Mele distingue três tipos de casos, nos quais o melhor juízo do agente enfrenta uma oposição, por motivação competitiva:

- o processo por *default*, que gera sem qualquer problema, um *judgement-matching intention*, mesmo na presença de oposição;
- um *judgement-matching intention*, é formado mesmo que o caminho por *default* para a intenção seja bloqueado por uma oposição
- a oposição motivacional bloqueia o caminho por *default* para a intenção e configura a produção de uma intenção *weak-willed*.

Assim, o exercício de auto controlo tem sempre um lugar na explicação dos porquês, quando o caminho por *default* do melhor juízo para a intenção é bloqueado e o agente por vezes intenciona de acordo com os melhores juízos e outras vezes não⁸⁵.

Considerando os casos referidos supra, quando existe uma falha acrática de intencionar de acordo com um determinado juízo, algo bloqueou a normal transição do melhor juízo para a intenção.

Na ação acrática, apesar do agente julgar que *a* é melhor, pode estar mais motivado para não fazer *a*, uma vez que os melhores juízos são formados, em parte significativa, com base na avaliação do agente, acerca da sua razão para a ação. E essa avaliação não tem necessariamente que estar alinhada com a força motivacional dos itens avaliados. Assim, o auto controlo é um elemento importante na questão da formação ou não formação de intenções, quando o caminho *by default* é bloqueado⁸⁶.

⁸³ Quando destranquei intencionalmente a porta do escritório hoje de manhã eu tencionei fazê-lo. Mas como este procedimento é um hábito normal, não foi necessária nenhuma decisão de destrancar a porta.

⁸⁴ MELE, Alfred, *Springs of Action. Understanding Intentional Behavior*, Oxford University Press Inc., New York 1992, pp. 230-234.

⁸⁵ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 65-66.

⁸⁶ MELE, *Springs of Action. Understanding Intentional Behavior*, op. cit., pp. 231-233.

Em suma, Alfred Mele considera que o mecanismo *judgement-matching intention* se verifica normalmente, quando se processa por *default*, casos nos quais nenhum esforço adicional ou de *self control* do agente é necessário. Portanto, para Alfred Mele, o *self control* permite explicar o motivo pelo qual, quando o caminho por *default*, do *better judgement* para a intenção é bloqueado, o agente por vezes intenciona com base no melhor juízo e outras vezes não, pois agir em conformidade com o seu melhor juízo dependerá dos esforços de *self control* do agente⁸⁷.

Evidentemente que sabemos que nem sempre as coisas ocorrem de forma suave (por *default*), porque podemos ser capazes de identificar o melhor e, devido à influência de desejos resistentes, intencionamos o pior. Esta situação mostra, não que os melhores juízos não desempenham um papel na etiologia das intenções e comportamento intencional, mas sim que, nos seres humanos, a formação de melhores juízos não assegura a formação da intenção correspondente. Ou seja, os melhores juízos decisivos podem ser derrotados por motivação oposta.

Nesta mesma linha, estes melhores juízos podem ser, em princípio, sustentados através do exercício de auto controlo. Em face de motivação que ameace tornar inefetivo o melhor julgamento, o agente pode tentar intencionar e agir de acordo com aquilo que julga melhor⁸⁸.

3.1.8. O Poder Motivacional do raciocínio prático

Alfred Mele explora a natureza da motivação, bem como o seu lugar na explicação do comportamento intencional em geral e também a aquisição de estados mentais tais como desejos, intenções e crenças. A sua visão da agência humana tem como foco principal a motivação⁸⁹.

Contrapõe a teoria da motivação antecedente (AMT), teoria do mecanismo cognitivo (CET) e visão das razões motivacionalmente fundamentadas (MGRV).

A AMT afirma que, nos seres humanos, toda a motivação produzida involuntariamente pelo raciocínio prático, que emita uma crença favorecedora de um curso de ação, deriva,

⁸⁷ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 66.

⁸⁸ MELE, *Autonomous Agents*, op. cit., pp. 29-30.

⁸⁹ MELE, *Motivation and Agency*, op. cit., p. 3.

em parte, da *motivation-encompassing attitudes*⁹⁰, já presente no agente antes de ter adquirido a crença.

A CET afirma que, nos seres humanos, alguns casos de raciocínio prático avaliativo, que tenham ou emitam uma crença favorecedora de um curso de ação, involuntariamente produzem motivação, que não deriva de todo de motivação antecedente.

Apesar destas teorias poderem assumir várias formas, ambas entendem as crenças avaliativas de uma forma cognitivista, na medida em que sustentam que essas crenças avaliativas são atitudes com conteúdo de verdade e que nem todas as crenças avaliativas estão condenadas a serem falsas, como deveriam ser, se fosse verdadeira uma teoria do erro, acerca das mesmas.

Alfred Mele vai articular e defender uma versão própria da AMT, menos restritiva. A versão mais simples corresponde a um instrumentalismo puro, segundo o qual o raciocínio gera motivação no sujeito, apenas identificando meios para fins desejados antecedentemente.

A AMT (teoria acerca das capacidades motivacionais do nosso raciocínio prático) carece ainda de ser distinguida da MGRV – visão das razões motivacionalmente fundamentadas - *motivationally grounded reasons view*.

A MGRV é uma visão relativa às condições segundo as quais são garantidas afirmações normativas do seguinte tipo: existe uma razão para *S* fazer *A*; *A* é a coisa racional para *S* fazer. A MGRV é uma tese acerca da questão normativa, ao passo que a AMT se preocupa com a questão causal.

A MGRV afirma que razões alegadamente normativas para a ação, são realmente razões para agir, apenas se forem devidamente sustentadas em motivos que o agente possui. A AMT não se compromete com esta premissa.

Na visão da motivação antecedente, a produção de nova motivação através do raciocínio, envolve motivação já existente antes do agente alcançar a conclusão cognitiva, motivação essa, presente no conjunto motivacional do agente. Alfred Mele rejeita a ideia de que a

⁹⁰ Atitudes (ou estados mentais atitudinais) que são, constituem ou incluem motivação. (MELE, *Motivation and Agency*, op. cit., p. 14). A versão curta do termo é: *antecedent motivation*.

base motivacional total de uma *motivation-encompassing attitude* determina sempre a força motivacional da atitude.

Portanto, uma formulação plausível da AMT deve acomodar a ideia de que a base motivacional total de uma *motivation-encompassing attitude*, nem sempre determina a força motivacional da atitude, uma vez que existe larga evidência de que os modos de representação têm uma influência significativa na força motivacional. Por outro lado, uma AMT plausível, não pode implicar, por motivos conceptuais, que a motivação antecedente deve (*must*) estar envolvida na geração da motivação pelo raciocínio. Talvez, ocasionalmente, outros fatores façam o trabalho da motivação antecedente. Ou seja, a ideia é a seguinte:

AMT Actual human beings and the real world are such that the capacity of justificatory reasoning to generate in a human being, S, a motivation-encompassing attitude whose object is S's A-ing and to do this in virtue of that reasoning's providing S with what he takes to be justification for A-ing is contingente on S's having antecedent motivation capable of being at least a partial positive motivational base for the generated motivation—encompassing attitude; and whenever such reasoning does generate a motivation-encompassing attitude of this kind in us in the way identified, antecedent motivation is in the positive motivational base of the generated attitude⁹¹.*

Alfred Mele entende que pode defender-se, para todas as partes, que a ação intencional é motivada. Os teóricos da motivação antecedente e do mecanismo cognitivo distinguem-se pelas suas afirmações respetivas, acerca de como o raciocínio prático gera a motivação nos seres humanos⁹².

Em suma, é muito plausível que todo o raciocínio prático seja motivado e, pelo menos, alguns dos desejos que motivam tal raciocínio estão localizados na base motivacional positiva dos desejos involuntariamente produzidos pelo raciocínio prático. É plausível, por exemplo, que o raciocínio prático motivado, parcialmente, por um desejo genérico de fazer aquilo que for melhor, dadas as circunstâncias é tal, que um desejo de *A* associado à conclusão do sujeito de que é melhor fazer *A*, tenha uma base motivacional parcial no desejo genérico.

⁹¹ MELE, *Motivation and Agency*, op. cit., p. 90.

⁹² MELE, *Motivation and Agency*, op. cit., pp. 89-91.

A AMT é compatível com abordagens empíricas que favorecem a explicação do comportamento motivado, de acordo com as quais (colocando de parte desejos intrínsecos para a ação) a motivação, em cursos de ação específicos, é produzida por combinações de motivação antecedente e crenças.

Pode ser entendido que, num agente racional, o juízo deliberativo de que um facto de qualquer tipo é razão para fazer *A*, é suficiente para dar ao agente a *motivation-encompassing attitude* que tem como objeto o facto de o sujeito fazer *A*. A combinação desde juízo com o facto de o agente ser racional, é causalmente ou conceptualmente suficiente para isto.

A AMT* de Alfred Mele, faz parte de uma teoria geral da etiologia do comportamento intencional, que também oferece uma explicação da produção involuntária de motivação, pelo raciocínio prático.

Por outro lado, a AMT* não insiste no facto de que o raciocínio pode gerar involuntariamente motivação, apenas se isso beneficiar da existência de alguma *motivation-encompassing attitudes*, não justificáveis pela razão. Isto é, a AMT* não especifica qual o tipo de motivação na qual se sustenta a justificação racional, quando involuntariamente gera motivação ao produzir uma crença. A AMT* defende sim que, nos seres humanos, existe alguma base motivacional antecedente para a motivação gerada⁹³.

3.1.9. *Self control*

Segundo Alfred Mele, o *self control* é exibido na vertente motivacional e na vertente avaliativa.

Numa primeira aproximação ao conceito, define *self-control* como a habilidade de controlar a motivação que é contrária ao melhor juízo. Ou seja, a habilidade de prevenir que essa motivação resulte num comportamento contrário ao melhor juízo decisivo do agente. O melhor juízo em questão, inclui juízos acerca de conduta avaliativa apropriada, bem como juízos acerca de ações externas, de modo que o exercício de *self-control* é uma

⁹³ MELE, *Motivation and Agency*, op. cit., pp. 99-101.

resposta à motivação contrária ao melhor juízo do agente, quer esse juízo seja detido explicitamente ou implicitamente.

Um agente *self-controlled* é alguém que está apropriadamente motivado para conduzir-se a ele próprio, conforme julgue melhor, e tem a habilidade de controlar a motivação para o contrário. Mas esta afirmação não equivale a dizer que o agente agirá sempre continuamente, nem podemos supor que um agente que exhibe *self-control* numa área também seja *self-controlled* em outras⁹⁴.

Portanto, a concepção de um agente *self controlled* corresponde a alguém que, caracteristicamente, é guiado pelos seus melhores juízos, mesmo diante de forte motivação competitiva e não entendem a emoção, a paixão e outros estados semelhantes, como não tendo lugar no eu do autocontrole. Em alguns casos, os melhores juízos refletem mesmo o nosso *ranking* avaliativo de emoções e apetites competitivos.

Para Alfred Mele, se tanto o autocontrole como a fraqueza de vontade podem ser manifestos no nosso comportamento, então também podem estar envolvidos em comportamentos mentais. Os sujeitos podem aceitar princípios relativos à aceitação, rejeição e modificação das suas crenças, emoções, valores, preferências ou desejos. A aceitação desses princípios pode, por vezes, assumir a forma de juízos. As ações mentais relevantes dos sujeitos autocontrolados estará de acordo com os princípios que aceitam⁹⁵.

Infere-se aqui que a perspectiva de Alfred Mele, responsabiliza o agente no seu uso do auto controle. Cada agente pode e deve refletir acerca dos seus princípios, considerando a modificação e revisão de crenças e desejos, fazendo-o de forma ainda mais rigorosa, se estiverem em causa esferas nas quais esse sujeito tem consciência do seu menor poder de auto controle. Portanto, apesar do agente estar sujeito aos efeitos de estados tais como a emoção ou a paixão, esse agente arca com a responsabilidade de ajustar esses efeitos, de forma a manifestar comportamentos concordantes com os seus melhores juízos, ajuste esse a ser efetuado através do exercício do auto controle.

Existe ampla evidência de que a motivação exerce, frequentemente, uma influência enviesada naquilo em que acreditamos, tal como nos casos de auto engano. Mas está

⁹⁴ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 50-60.

⁹⁵ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., p. 92.

também reconhecido que temos algum controlo sobre as emoções particulares que resultam na ação. No entanto, também existe lugar para o autocontrolo ao trazer as nossas emoções em alinhamento com os melhores juízos relevantes. Por exemplo, podemos conter uma simpatia exagerada por uma personagem de um filme, se nos lembrarmos a nós próprios que é apenas uma personagem, sendo este um exemplo de auto controlo interno. Ou alguém pode vencer um momento de tristeza, telefonando por exemplo para um irmão, sendo este um caso de autocontrolo externo.

Por outro lado o auto controlo pode ser regional (uma vez que um sujeito pode ser auto controlado numa área da sua vida e *weak-willed* em outra área) ou global (autocontrolo em todas as áreas da sua vida). O autocontrolo surge ainda em grau, na medida em que alguns indivíduos parecem ser mais auto controlados do que outros.

Em agentes normais a capacidade de auto controlo não é equivalente à força física. Cada sujeito aprende a resistir às tentações, prometendo a si próprios recompensas, ou imaginando os efeitos nefastos de uma conduta imprudente. Ou seja, a capacidade de auto controlo inclui uma variedade de competências e considerável conhecimento acerca de quais as competências a usar em situações particulares⁹⁶.

Parece pertinente, neste ponto, recordar que George Ainslie também se refere aos esforços de auto controlo do agente e a um sistema de recompensas.

Reconhece que existem apenas graus de motivação e graus de deliberação, ou seja, não classifica as ações como sendo simplesmente racionais ou irracionais. O autor acredita que o teste para a racionalidade se relaciona com o facto da escolha do agente ser congruente com a gama de interesses do *self*. Na verdade, prefere focar-se na motivação que, entende, implica a recompensa, que se torna necessária para resolver a competição entre os membros da classe de processos que podem ser substituídos um pelo outro.

O seu trabalho experimental demonstrou o fenómeno da reversão preferencial em agentes que obtêm benefícios imediatos, fenómeno este que explica como sendo um desconto hiperbólico de recompensas em perspetiva – *hyperbolic discounting of prospective rewards*, o que significa que a valorização do agente das recompensas em perspetiva é inversamente proporcional à sua espera. O termo *picoeconomics* (já referido em momento

⁹⁶ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 92-94.

anterior) é usado para descrever a tendência que as pessoas têm, para mostrar uma preferência mais forte por uma recompensa imediata, em alternativa a benefícios a longo prazo, tendência esta que aumenta, quanto mais próximas estiverem ambas as situações, do momento presente. Ou seja, perante duas recompensas similares, em comparação, os agentes humanos demonstram preferência por aquela que chega mais cedo, descontando portanto o valor da recompensa mais tardia.

Relativamente à acrasia, a ideia do desconto hiperbólico de recompensas em perspetiva, é sugerido um modelo do *self* como sendo uma população de processos em busca de recompensas.

Os interesses de curto e longo prazo do agente são vistos como processos de procura de recompensas, em competição entre si, de acordo com uma regra de maximização de recompensas esperadas, descontadas em cada momento de escolha.

E esta situação é combinada com um modelo de negociação inter-temporal da vontade. No contexto desta competição, a vontade tem uma função crucial: a de manter a preferência do agente pelo maior e mais tardio em vez do mais pequeno e menos tardio⁹⁷. Deste modo, o poder da vontade – *willpower*⁹⁸ – torna-se numa função necessária, na medida em que permite que o agente mantenha a preferência pelas alternativas maiores e mais tardias ao longo do tempo. Para George Ainslie, o modelo de desconto hiperbólico oferece um mecanismo: a negociação inter-temporal entre estados motivacionais sucessivos. E este mecanismo explica as diversas propriedades do sujeito, sem ser necessário recorrer à noção de um ego homuncular⁹⁹.

Assim, para George Ainslie o auto controlo, é um controlo de impulsos. Evita, ou tenta evitar, preferências temporárias por bens mais pequenos e mais próximos, em vez dos maiores e mais tardios, ou, males maiores e mais tardios em vez dos mais pequenos e mais próximos.

⁹⁷ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 14-16.

⁹⁸ Parece ser possível estabelecer um paralelismo entre a noção de *willpower* e auto controlo, como sendo termos com idêntico significado.

⁹⁹ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., p. 99.

Portanto, qualquer auto controle que sirva os nossos interesses de longo alcance, nunca será irracional¹⁰⁰. Por outro lado, a forma como os interesses de médio prazo, controlam os interesses de curto prazo, mas derrotam os interesses de longo prazo, pode ser considerado irracional.

Portanto, na visão de George Ainslie, a acrasia consiste na falha do agente em agir em função dos seus interesses de longo prazo, para proteger um interesse oposto, de curto prazo. Ou, por outras palavras o agente acrático, não consegue proteger os interesses de longo prazo, através da sua vontade¹⁰¹.

Considerando que para os interesses de curto prazo, o sucesso implica evitar compromissos, mas, por outro lado, implica também, comprometer-se a não agir em função de outros interesses de curta duração incompatíveis e ainda poderá implicar desobedecer a alguns interesses de longo alcance, George Ainslie avança quatro táticas de pré compromisso que podem influenciar escolhas futuras:

- Compromisso extrafísico. Tornar uma escolha futura, fisicamente impossível ou procurar incentivos adicionais que possam influenciar o *self* futuro. Por exemplo, alguém que quer emagrecer toma comprimidos inibidores do apetite;
- Manipulação da atenção. Tentar evitar informação que possa alterar a nossa decisão. Se sabemos de antemão que está disponível uma recompensa sedutora, podemos evitar pensar sobre isso.
- Preparação da emoção. O sujeito pode cultivar ou inibir o processo emocional. Estes processos podem alterar de que forma a expectativa da recompensa influencia as escolhas, pelo menos num futuro próximo.

¹⁰⁰ Desta definição George Ainslie exclui dois casos de auto controle: 1. esforços de auto controle exagerados, ou mal executados, devido a erro ou falta de competências (por exemplo, seguir as regras dos nossos pais, pode não servir as nossas próprias necessidades). Em algum momento, todos nós cometemos erros na escolha dos controles a fazer e, por vezes, esses erros podem resultar em patologia (por exemplo, uma dieta pode ser levada tão longe, que resulta numa anorexia nervosa). 2. Esforços ostensivos de auto controle, podem servir interesses de curto ou médio prazo (por exemplo, os anoréticos podem achar que possuem um arma contra os seus pais, ou uma sensação de poder. E estas recompensas podem ter por base um projeto racional de auto controle. Um vício pode estar mascarado de auto controle). Sobrevalorizar a necessidade de auto controle, a falta de capacidade em exercê-lo e os impulsos em exercê-lo como um disfarce, são alvos comuns da psicoterapia.

¹⁰¹ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 112-114.

- Regras pessoais¹⁰².

Parece possível estabelecer uma analogia de significado entre as táticas referidas por George Ainslie e as competências referidas por Alfred Mele, na medida em que, devidamente aplicadas em situações particulares, conferem a probabilidade do agente executar ações racionais. Estas competências terão necessariamente de estar sustentadas no exercício do auto controlo do agente, ajustado a cada situação da vida.

Para George Ainslie, a negociação inter-temporal é a nossa vontade – *will*. E a nossa vontade, é ameaçada por racionalizações que permitem exceções à escolha disponível. A vontade é mais estável se recorrer a critérios: por exemplo, uma regra pessoal de nunca beber, é mais estável do que ter uma regra de beber apenas dois copos de vinho, pois esta ultima regra está sujeita a definição¹⁰³.

Mas voltando a Alfred Mele, a acrasia é, portanto, uma falha de *self-control*, não implicando, no entanto, a sua total ausência. Tanto a acrasia como a encrateia, residem num *continuum* entre dois extremos: por um lado um agente que falha totalmente no *self-control* e encontra-se totalmente sem motivação para agir conforme julga melhor e, no extremo oposto, está a pessoa perfeitamente *self-controlled*. Um agente acrático pode ser identificado, segundo Alfred Mele como abaixo do normal na sua posse de auto controlo ou na sua motivação para exercer o *self-control* ou em ambos os casos¹⁰⁴.

3.1.9.1. Auto controlo e tentações proximais

Se um desejo não é irresistível, compete a nós próprios decidir se agimos de acordo com o mesmo. Um outro elemento consiste na ideia de que relativamente poucos dos nossos desejos são irresistíveis. E estas duas ideias, em conjunto, sugerem que compete a nós próprios decidirmos, pelo menos num sentido compatibilista, segundo qual dos dois ou mais desejos competitivos, vamos agir.

Considerando a perspetiva motivacional do auto controlo, exercer o autocontrolo ao agir segundo um agente julga melhor, não é problemático. Por exemplo, um fumador que esteja agora mais fortemente motivado para ter presente, que amanhã não vai fumar, do

¹⁰² AINSLIE, George, *Break-Down of Will*, Cambridge University Press, Cambridge 2001., pp. 74-78.

¹⁰³ AINSLIE, *Break-Down of Will*, op.cit., p. 104.

¹⁰⁴ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 50-60.

que não ter esta ideia presente, pode levar o sujeito a executar uma estratégia apropriada: conduzir até um sítio remoto e deixar os cigarros para trás.

Mas consideremos um fumador que está mais motivado para fumar um cigarro, neste momento, do que está motivado para não o fazer. E suponhamos que, ainda assim, este agente julga melhor não o fazer. Pode este agente exercer o autocontrolo para não fumar, agora?

Os exercícios intencionais de autocontrolo para fazer *A*, são impossibilitados pelas tentações proximais preponderantes: querer *B* agora mais, do que querer *A* agora (ou *A* mais tarde), além disso, o agente também quer *B* mais agora, do que quer exercer autocontrolo agora para fazer *A* e portanto, o agente não vai exercer intencionalmente o autocontrolo para este fim. Este cenário sugere que somos impotentes contra a motivação proximal preponderante, que colide com o nosso melhor juízo, ou seja, que essa motivação é irresistível, mas Alfred Mele não concorda com esta sugestão. Se vamos exercer o autocontrolo, então é melhor fazê-lo previamente¹⁰⁵.

Assim, nos casos de tentação proximal preponderante, o agente detém o juízo decisivo de que é melhor fazer *A* agora, do que fazer *B* agora, deseja mais fazer *A* agora, e reconhece que não pode fazer ambos.

Para Alfred Mele, uma tentação proximal, sendo preponderante, não a torna irresistível. Em muitos casos o agente que julga melhor fazer *A* do que *B*, mas deseja mais *B*, pode também querer mais fazer *A* e *C* agora, do que quer fazer *B* agora, e age de acordo com esses desejos compostos. Por exemplo, no caso dos substitutos de cerveja não alcoólica, um agente que julgue melhor evitar beber agora, do que beber agora e cujo desejo proximal de beber seja mais forte do que o desejo proximal de evitar beber, pode ter ainda um desejo proximal, mais forte, de desempenhar a ação composta de evitar beber, enquanto reduz o desconforto do desejo insatisfeito, bebendo uma cerveja sem álcool. O agente impediu a tentação, ao agir em função de um desejo composto. Por outro lado, uma tentação proximal preponderante, pode permitir um exercício de autocontrolo que suporte a ação de acordo com o melhor juízo do agente. Por exemplo, o Ian pretende

¹⁰⁵ MELE, *Autonomous Agents*, op. cit., pp. 41-42.

terminar de pintar um quarto durante o dia de hoje, mas decide fazer uma pausa para lanche. Durante essa pausa liga a televisão e fica tentado a continuar a ver um programa que lhe agradou, em vez de continuar a trabalhar. Ao querer mais continuar a ver televisão do que voltar ao trabalho, o Ian pode, no entanto, exercer o autocontrolo com sucesso. Os desejos não possuem pesos imutáveis e no caso do Ian, a sua tentativa de autocontrolo vai ter o resultado de que o desejo proximal mais forte é o de voltar ao trabalho, do que permanecer em frente à televisão.

Assim, a nossa capacidade de controlo sobre os nossos desejos, vai além dos desejos proximais, incluindo os desejos proximais de *A* que competem com os nossos melhores juízos e são mais fortes do que os nossos desejos proximais de não fazer *A*.

Em termos de tentações, quanto maior a ocupação de motivação incontrolável, mais fraca será a nossa perspectiva de conduta autónoma. Porque a capacidade de autocontrolo, permite vislumbrar a ideia de autonomia¹⁰⁶.

3.1.10. A pessoa idealmente autocontrolada

Para Alfred Mele, o agente perfeitamente auto controlado manifesta um auto controlo perfeito e essa perfeição tem quatro dimensões: alcance, objeto, frequência e efetividade.

Alcance – *Range* – considerando a distinção entre auto controlo local e global, o auto controlo perfeito, é global. Manifesta-se em ações evidentes, ações mentais, intenções, crenças e emoções, raciocínio prático e melhores juízos decisivos, na avaliação, aceitação e revisão de valores e princípios, etc. Este agente tem a categoria máxima de alcance.

Objeto – *Object* – os exercícios de auto controlo combatem algo em defesa de outra coisa. A pessoa idealmente auto controlada, nunca exerce o auto controlo de forma errada, mas apenas em favor de itens puros – melhores juízos decisivos adquiridos não acriticamente, valores, princípios, etc. Os objetos dos exercícios de auto controlo, na pessoa idealmente auto controlada são, portanto, perfeitos.

Frequência – *frequency* – a pessoa idealmente auto controlada, exerce o auto controlo sempre que refletidamente entende apropriado fazê-lo.

¹⁰⁶ MELE, *Autonomous Agents*, op. cit., pp. 54-55.

Efetividade – *effectiveness* – os exercícios de auto controlo da pessoa idealmente auto controlada sucedem sempre, ao apoiarem aquilo que é suposto apoiar. Esta situação não é uma questão de sorte ou desvio causal. Antes, a pessoa idealmente autocontrolada consistentemente e intencionalmente consegue o sucesso dos seus exercícios de auto controlo.

Obviamente, esta pessoa idealmente auto controlada é um ser imaginário. O indivíduo auto controlado não precisa de ser rígido no seu comportamento, nem precisa de ser excessivamente calculista. A forma como se conduzem a si próprios depende dos seus valores, princípios, melhores juízos decisivos, etc, que constituem itens psicológicos que não precisam de regimes rígidos ou rigorosos¹⁰⁷.

Ou seja, o alcance, objeto, frequência e efetividade do auto controlo são apenas dimensões a considerar no exercício do auto controlo, em agentes como nós. Estas dimensões serão sujeitas a um ajuste a cada cenário em particular, bem como à intensidade do desejo existente. Fica plasmada mais uma vez a noção de que o auto controlo funciona num *continuum* e em função de grau.

3.1.11. Desejo e motivação

Alfred Mele refere alguns pontos relativamente ao uso Filosófico dos termos desejo, vontade e motivação.

O desejo de *A*, sendo que *A* é um curso de ação potencial, constitui motivação para *A*. Por exemplo, o meu desejo de fazer exercício hoje, constitui uma motivação para fazer exercício hoje. Os desejos-ação, são vistos por Alfred Mele, como partes importantes da classe de atitudes que englobam motivação. E uma importante característica destas atitudes é o facto de terem a função de contribuir para a sua própria satisfação, ao induzir o agente a agir segundo as mesmas. O autor distingue seis teses acerca da motivação, presentes na literatura Filosófica e Psicológica:

- a motivação está presente no reino animal, mas não se estende completamente;
- os agentes motivados têm a capacidade de representar objetivos e meios para os alcançar;

¹⁰⁷ MELE, *Autonomous Agents*, op. cit., pp. 121-123.

- uma atitude que engloba motivação deve ter um objetivo ou um meio como seu objeto;
- a motivação varia em força;
- quanto maior for a motivação de um agente para *A*, em comparação com a motivação para cursos de ação alternativos, mais provável será que o agente faça *A*, permanecendo iguais todas as outras coisas;
- Sempre que os agentes agem intencionalmente, existe algo para o qual estão efetivamente motivados para fazer.

Alfred Mele procura dar uma noção de motivação consistente com as seis características enunciadas¹⁰⁸.

3.1.11.1. Atitudes que essencialmente englobam motivação para a ação

Alfred Mele faz uma distinção entre atitudes que abrangem essencialmente motivação para *A* e atitudes que o fazem contingentemente (sendo *A* uma ação variável). Portanto, uma atitude de um agente *S*, essencialmente abrange motivação para *A*, se e só se, abrange motivação para *A* não apenas na situação atual de *S*, mas em todos os cenários possíveis nos quais *S* tem essa atitude. E uma atitude abrange contingentemente motivação para *A*, se e só se, apesar da situação atual de *S*, abranger motivação para *A*, este não é o caso em alguns cenários possíveis nos quais *S* tem essa atitude. O termo abranger motivação – *encompass motivation* – é usado pelo autor como um termo que abrange o ser, constituir ou incluir motivação para *A*.

Alfred Mele pretende desenvolver uma formulação de atitudes focadas em *A*, que essencialmente abranjam motivação para *A*. Atitudes focadas para *A* são atitudes diretamente relacionadas com *A*.

Apenas alguns dos nossos desejos correspondem a desejos ação – *action-desires* – que são desejos do agente de que *A*, onde *A*, é uma variável para potenciais cursos de ação que se forem satisfeitos, são satisfeitos apenas através de ações (e verdades, estados, ou factos para os quais a satisfação das ações são conceptualmente suficientes).

Para Alfred Mele, «*All intentions to A and all desires to A essentially encompass motivation to A*¹⁰⁹». O autor sugere que intencionar *A* é, pelo menos em parte, estabelecido

¹⁰⁸ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 69-71.

¹⁰⁹ MELE, *Motivation and Agency*, op. cit., p. 135.

como fazer *A*, e fazer *A*, é ter a atitude que abrange motivação – *motivation-encompassing attitude* em relação a fazer *A*: um agente que não possui essa atitude não possui o compromisso psicológico para fazer *A*, que é essencial para ficar estabelecido o facto de fazer *A*.

Assim, parece ser incontroverso que as intenções para *A* abrangem motivação para *A*.

Relativamente aos desejos ação – *action-desires* – se desejar que *A* não exclui a presença de certos estados cognitivos, então um sujeito pode desejar que *A*, apesar de não ter motivação para *A*. Por exemplo, se um agente desejasse que *A*, apesar de estar convencido que não pode fazer *A*, o agente poderia desejar que *A*, apesar de não ter qualquer motivação para *A*. No entanto, desejar que *A* exclui a presença de estados cognitivos deste tipo. Se eu estou convencida que não posso viajar mais rapidamente do que a velocidade da luz e alterar o passado, então, apesar de poder querer fazê-lo, não o desejo fazer. Assim, existindo uma inclinação deste tipo, um desejo ação constitui motivação para agir. Desejar que *A* exclui estar convencido de que não podemos fazer *A* ¹¹⁰.

Por outro lado Alfred Mele procura também perceber se uma esperança de que *A*, é constituída essencialmente por um desejo de que *A* e a crença, para o efeito, de que um sujeito tem a oportunidade de fazer *A*, então, sendo verdade que as intenções para *A* e os desejos para *A* abrangem essencialmente motivação para *A*, logo, a esperança de que *A*, essencialmente abrange motivação para *A*. Mas não em virtude de terem como componente, uma atitude que essencialmente abranja motivação para *A*. Atitudes compostas por duas ou mais atitudes são atitudes combinadas.

Alfred Mele pretende apurar a natureza de atitudes não combinadas que focam *A* – atitudes EMA – que abrangem essencialmente motivação para *A*. As atitudes não combinadas, são atitudes que não possuem atitudes distintas como partes.

Para isso começa por referir que a efetividade de uma ação desejo – *action-desire* – define-se da seguinte forma: para que um desejo de que *A* seja efetivo, esse desejo deve ser a causa para o agente fazer *A*. O fato de fazer *A* satisfaz o desejo e assim uma ação desejo efetiva, é a causa da sua própria satisfação (o mesmo é verdade para as intenções). Além disso, vemos os desejos para *A* como tendo, de forma mais completa, a função de

¹¹⁰ MELE, *Motivation and Agency*, op. cit., pp. 134-136.

contribuir para a sua própria satisfação, através da sua contribuição para conduta intencional do agente, que é direcionada diretamente para fazer *A* (o mesmo é verdade para as intenções). Tentar, é essencialmente intencional e qualquer tentativa para *A* ou ter presente fazer *A* é diretamente relacionado com fazer *A*.

Tentar, não requer esforço especial em tentar fazer *A*, e é diferente de tentar ter presente fazer *A*. Por exemplo, Ed pretende fazer exercício diariamente, mas teme deixar-se vencer pela preguiça, então tenta ter presente que se exercita diariamente, dizendo aos seus amigos que pretende fazê-lo. Ed tenta ter presente que se exercita todos os dias, ao aumentar a sua motivação para o fazer, porque ao anunciar que se exercita e caso depois não faça, então isso terá o custo adicional de ficar mal visto perante os amigos. Ao fazer o anúncio, Ed não está a tentar exercitar-se, mas está a tentar ter presente que o faz. Assim, o desejo de tentar *A* e o desejo de tentar ter presente fazer *A*, são satisfeitos respetivamente, se o sujeito tentar *A* e se tentar ter presente fazer *A* ¹¹¹.

As atitudes EMA ou unidades constituídas pelas atitudes e respetivas realizações físicas, podem ser relacionadas com ações, de maneira análoga à forma pela qual as câmaras se relacionam com fotografias. Porque, as câmaras estão envolvidas na produção de fotografias e as fotografias são coisas, cuja produção requer câmaras.

As atitudes EMA (ou unidades constituídas pelas atitudes e suas realizações físicas) são coisas envolvidas na produção das ações e as ações são coisas, cuja produção requer atitudes EMA. As atitudes EMA têm, essencialmente, uma conexão funcional com a ação intencional, uma conexão do tipo eminentemente motivacional. E diferem funcionalmente das atitudes de procura de verdade.

Pelo menos algumas atitudes EMA, nomeadamente ações desejos – *action-desires* e intenções são atitudes paradigmáticas nos agentes humanos ¹¹².

3.1.11.2. Força Motivacional

Alfred Mele defende uma posição acerca da força motivacional de determinadas formas de atitudes EMA.

¹¹¹ MELE, *Motivation and Agency*, op. cit., pp. 138-140.

¹¹² MELE, *Motivation and Agency*, op. cit., pp. 155-156.

Defende um princípio que relaciona a força motivacional dessas atitudes com a ação intencional e mostra que a combinação de uma visão causal da explicação da ação, com um princípio plausível que liga força motivacional com ação intencional, não nos faz vítima dos nossos desejos mais fortes. Clarifica a visão causal do papel da motivação na agência humana, através de dois elementos: força motivacional e controlo¹¹³.

Um desejo de que *A* é a inclinação do sujeito para *A*. E a força dessa ação-desejo, é a força dessa inclinação.

Os princípios que ligam a força motivacional à ação intencional têm importância na explicação da ação intencional, na medida em que podem articular pressupostos de fundo, acerca da ação intencional, no contexto no qual as pessoas oferecem explicações de ações intencionais particulares em termos dessas atitudes, tais como desejos, crenças e intenções.

O pressuposto da existência de uma conexão estreita entre a força motivacional e ação intencional, irá levar-nos a reunir informação acerca do porquê do agente estar mais fortemente motivado para fazer o que fez, de forma a conseguir perceber porque o sujeito fez aquilo.

Em muito casos, um agente que faz *A* a fim de *B*, também faz *B* a fim de *C*, etc. Por exemplo, Sam liga um interruptor para iluminar a sala, de forma a chegar mais facilmente às chaves do carro e para que possa ter maior probabilidade de chegar a tempo ao emprego. Suponhamos que Sam tem um desejo proximal de aumentar a probabilidade de chegar a tempo ao emprego e, para isso, executou o plano descrito. Suponhamos também que desejando não acordar os filhos, tem o desejo proximal de não iluminar a sala e, portanto, um desejo proximal concordante de não ligar o interruptor.

Apenas e só quando a força de um desejo proximal do agente para *A*, é diretamente menos relevante, se ele intencionalmente fizer *A* ou, pelo menos, tentar fazer *A*, do que é a força de alguma ação desejo proximal competitiva ou algum desejo proximal competitivo do sujeito para o não fazer, o seu desejo proximal para *A*, é vencido por esse desejo competitivo. O desejo proximal de Sam de aumentar a probabilidade de chegar a tempo

¹¹³ MELE, *Motivation and Agency*, op. cit., p. 161.

ao emprego, mas não o seu desejo proximal de ligar o interruptor, é vencido pelo seu desejo proximal de não ligar o interruptor.

Alfred Mele dá relevância a três questões: 1. o enviesamento de mecanismos que ligam ações desejo proximais, a tentativas associadas, 2. desejos de não fazer – *not-doings* - e 3. a questão acerca da conexão conceptual entre desejos e intenções.

Relativamente ao enviesamento de mecanismos que ligam ações desejo proximais a tentativas associadas (1), o facto da capacidade de quem vai realizar desejos ter um papel na produção de tentativas apropriadas a esse desejo, depende de propriedades intrínsecas de quem as realiza e outras propriedades do agente. Estas outras propriedades relacionam-se com mecanismos que ligam realizações de ações desejo proximais não vencidas, com tentativas. Quando estes mecanismos não favorecem as tentativas apropriadas, para nenhum dos desejos relevantes, mais do que tentativas apropriadas para qualquer dos outros, os desejos podem designar-se - *bias free*.

Sobre o desejo de não fazer (2), esses desejos podem estar numa base motivacional positiva ou negativa ou ações desejo proximais não vencidos *bias-free*. Por exemplo, um desejo de não *A* pode ajudar a racionalizar e contribuir para a força do desejo de fazer algo que assegure isso, para fazer nesse caso, ou para aumentar a probabilidade de que o sujeito não faça *A*, incluindo desejos proximais não vencidos deste tipo. O desejo de não fumar hoje pode estar na base motivacional positiva de um desejo proximal não vencido de atirar os cigarros ao lixo. Além disso, desejos para não fazer, podem ajudar a motivar decisões e intenções, que ajudam a produzir ações desejo proximais não vencidas¹¹⁴.

Relativamente à conexão conceptual entre desejos e intenções (3), refere que o facto das intenções para *A* serem irredutíveis a complexos de outras atitudes, é consistente com o facto de terem desejos de *A*, como constituintes.

Intenções para *A*, podem abranger desejos para *A*, sendo o resto irredutível.

Consideremos o seguinte cenário: o Al adora um determinado tipo de cerveja e, ao voltar para casa, vê por acaso, uma loja com essa cerveja em promoção. Então decide estacionar o carro, entrar na loja e efetuar a compra. Podemos dizer, neste caso, que

¹¹⁴ Por exemplo, para não ofender os colegas vegetarianos, Bob procura no menu algo sem carne, mas apetitoso. O desejo proximal não vencido, mais forte de Bob, pode ser pedir uma pizza vegetariana e esse facto pode ser explicado, em parte, pelo desejo negativo mencionado.

psicologicamente o facto de A1 ter visto o anúncio, suscitou primeiro a sua intenção de comprar a cerveja, antes de suscitar o desejo.

Intencionar *A* pode ser desejar *A*, tal como insistir que *p* é interiorizar que *p*, (de forma mais robusto sem englobar a sua contraparte menos robusta). Para Alfred Mele, a robustez comparativa de uma intenção deve-se, não ao facto da sua força motivacional ser maior do que a de um mero desejo, mas sim, ao facto de ter uma conexão funcional mais estreita com a ação, do que um mero desejo.

Assim, o autor não pretende afirmar que intenções para *A* não englobam, necessariamente, desejos para *A*, mas pretende sim, deixar em aberto essa possibilidade. Assim, perante esta ideia e tendo presente os desejos não vencidos, formula o seguinte princípio:

Mu. Assuming bias-free desires and the absence of relevant intentions thus far, if, at a time, an agent has an untrumped proximal desire to A that is stronger than any untrumped proximal desire that he has to do anything else and than any untrumped proximal desire that he has for a relevante not-doing, he proceeds to A or at least to try to A then, provided that he proceeds to do something intentionally then^{115, 116, 117}.

Relativamente às ações-desejo, Alfred Mele defende que agentes como nós têm ações desejo, incluindo os não enviesados e ações desejo proximais não vencidos. E estes

¹¹⁵ Desejar que *A*, quando *A* não é tentar fazer alguma coisa, não precisa de ser sempre acompanhado de um desejo de tentar *A*, de forma a resultar em fazer *A*, intencionalmente. Geralmente, se não estamos em dúvida acerca de conseguirmos fazer *A* (sendo *A* algo que queremos fazer), não temos um desejo específico de tentar fazer *A*. No entanto, deve notar-se que, mesmo que um desejo para fazer uma ação *A* que o sujeito pode desempenhar, poderá ser vencido. Ilustrando esta questão: Sam foi levado a acreditar falsamente que, pelo menos um dos seus braços está temporariamente paralisado, havendo uma boa hipótese de que ambos possam estar: Sam acredita que existem 90% de hipóteses de que o seu braço direito esteja paralisado e 70% de hipóteses de que o seu braço esquerdo esteja paralisado. Ofereceram a Sam \$100 para que levantasse o seu braço direito, quando lhe fosse retirada a camisa-de-forças e \$50 para que levantasse o esquerdo, na condição de que não levantar ambos os braços em simultâneo. Ao querer mais \$100 do que \$50, Sam quer levantar primeiro o braço direito, mas levanta antes o esquerdo. Sam quer tentar fazer isso, mais do que quer tentar levantar o braço direito. Em circunstâncias especiais deste género, o desejo para uma ação *A*, que um sujeito pode executar, é vencido por um desejo de tentar *B*, e se o primeiro desejo do agente leva a que intencionalmente faça *A*, ele vai fazer, de uma forma que depende do agente ter um desejo proximal distinto de tentar *A*, que é instrumental. (p. 176).

¹¹⁶ MELE, *Motivation and Agency*, op. cit., p. 172.

¹¹⁷ MELE, *Motivation and Agency*, op. cit., pp. 164-172.

desejos são concretizados em estados físicos. A concretização dos estados têm propriedades físicas adequadas ao desempenho de um papel de causalidade da ação.

Nos agentes, a força de uma ação desejo é sustentada em características físicas dos estados físicos, que concretizam o desejo. As diferenças na força das ações desejo dos sujeitos, são sustentadas em diferenças físicas na concretização de estados.

Esta sugestão relaciona o conceito de força de desejo com o conceito de ação intencional. A força relativa de uma ação desejo proximal minha, não enviesada e não vencida, é concebida, em parte, nos termos do que eu, provavelmente, vou tentar fazer numa determinada altura, se eu fizer uma tentativa que seja apropriada a algum dos meus desejos, devido às propriedades relevantes das bases desses desejos.

A interpretação que Alfred Mele deu acerca da ideia de que ações desejo proximais não vencidos, varia em força motivacional, contraria, portanto, a ideia de que a força motivacional é uma noção vazia. Os desejos não enviesados deste género, estão mais intimamente ligados à ação intencional do que outros desejos.

Em conclusão, alguns dos nossos desejos são motivacionalmente mais fortes do que outros, noção esta, que é relativamente importante considerando a sua aplicação nas ações-desejo proximais, não enviesadas e não vencidas.

Por outro lado, a força motivacional das atitudes constituídas por motivação, reside na sua capacidade de combinação com probabilidades subjetivas que influenciem a força das ações-desejo proximais, não enviesadas e não vencidas¹¹⁸.

3.1.12. *Core weak-willed action*

Mele apresenta duas teses acerca da *core weak-willed action*:

T1 – os nossos melhores juízos são normalmente baseados pelo menos em parte, nas nossas avaliações dos nossos objetos de desejo (aquilo que desejamos).

T2 – A força motivacional dos nossos desejos, nem sempre coincide com as nossas avaliações dos objetos de desejo.

¹¹⁸ MELE, *Motivation and Agency*, op. cit., pp. 172-175.

A partir destas teses, não será surpreendente imaginar que algumas vezes, apesar de julgarmos melhor *a*, e julgarmos melhor fazer *a* do que *b*, somos mais motivados para *b*, do que para *a*.

A tese 1 encaixa na concepção de racionalidade prática, na medida em que, ao raciocinar acerca do que fazer, tentamos descobrir aquilo que é o melhor ou suficientemente bom a fazer e não nos focamos naquilo que estamos mais motivados a fazer. Quando fazemos este raciocínio, ao mesmo tempo que temos desejos conflitantes, os nossos juízos são tipicamente formados parcialmente na nossa avaliação dos objetos de desejo, a qual pode estar desalinhada com a força motivacional desses desejos, considerando que T2 é verdade.

Somente se fossemos agentes ideais, as nossas avaliações dos objetos dos desejos, determinariam e coincidiriam com a força desses desejos e, neste caso, estaríamos sempre mais fortemente motivados para fazer aquilo que julgamos melhor.

Alfred Mele recorda a tese de George Ainslie, que defende que a força motivacional dos desejos tende a incrementar-se hiperbolicamente, à medida que se aproxima a altura da sua satisfação. Se não existe uma tendência de concordância nas nossas avaliações dos objetos dos nossos desejos, pode ocorrer frequentemente que a força de um desejo está seriamente desalinhada com a avaliação do agente acerca dos seus objetos¹¹⁹.

Estudos empíricos acerca do papel das representações dos objetos desejados no comportamento impulsivo e adiamento das recompensas, sugerem forte evidência de que as nossas representações dos objetos desejados têm duas dimensões, a motivacional e a informacional.

Os nossos melhores juízos poderão ser mais sensíveis à dimensão informacional das nossas representações, resultando daqui que esses juízos, por vezes, recomendam ações, que estão desalinhadas com aquilo que estamos mais fortemente motivados para fazer, na altura. É deste modo que se tornam possíveis as *weak-willed actions*, porque, pelo menos

¹¹⁹ Por exemplo, um sujeito dá pouco valor a uma sobremesa que poderia comer mais logo, julgando que é preferível suprimir a sobremesa e tem um desejo de força moderada por essa sobremesa. Mas quando a sobremesa é servida, após o jantar de baixas calorias, a força do seu desejo pode-se aumentar dramaticamente, conforme previsto no modelo de Ainslie. Assim, teremos um desalinhamento entre motivação e avaliação. O sujeito julga melhor não comer a sobremesa, enquanto está muito mais motivado para a comer do que para a suprimir.

algumas ações intencionais, que estão em conflito com o nosso raciocínio de melhor juízo, são efetuadas de forma livre, sã e na ausência de depressão.

De facto, o agente normal é capaz de influenciar a força do desejo das mais variadas formas, exercendo *self-control*. Por exemplo, pode recusar focar a sua atenção nos aspetos atrativos de um curso de ação sedutor e concentrá-los naquilo que deve ser realizado, agindo conforme julga melhor. Pode procurar aumentar a sua motivação para seguir a ação julgada melhor, prometendo a si próprio recompensas por fazê-lo. Pode imaginar um item desejado de forma inatrativa. Os desejos, normalmente, não têm forças imutáveis e a plasticidade da força motivacional é pressuposta por concepções *standard* de auto controle.

A chave para entender as *weak-willed actions*, passa por perceber que a força motivacional de uma atitude motivacional, não tem que estar alinhada com a avaliação que o agente faz do objeto daquela atitude. Os nossos melhores juízos racionais, são baseados, numa parte significativa, na nossa avaliação dos objetos dos nossos desejos e quando a avaliação e a força motivacional não estão alinhados, podemos acreditar que *a* é melhor, e portanto é melhor fazer *a* do que *b*, mas estar mais motivados para *b* do que para *a*. Enquanto esta crença, baseada na racionalidade prática existir, podemos executar *b* de forma livre, sã e na ausência de depressão, sendo que *b* é uma ação *core weak-willed*¹²⁰.

Aquilo que o agente está mais motivado a fazer e aquilo que o agente faz, depende da sua capacidade de *self-control*. Em todos os casos de ação incontinente, um agente faz aquilo para que está mais motivado a fazer, pelas suas próprias razões. O agente que acriticamente faz *a*, faz *a* por uma razão.

As razões, no entanto, não bastam para compreender a ação acrática, e é necessária que a explicação psicológica seja completada com outros fatores acerca do agente: fatores tais como a sua capacidade de *self control*. Se o agente fosse perfeitamente *self-controlled*, teria balanceado as suas motivações de modo a não ter agido acriticamente. Neste sentido, Alfred Mele rejeita a P2 proposta por Donald Davidson, na medida em que entende que

¹²⁰ MELE, *Backsliding. Understanding Weakness of Will*, op. cit., pp. 71-82.

a conexão entre o melhor juízo e o balanço das motivações do agente, são mais complexas do que Donald Davidson faz supor¹²¹.

Uma explicação adequada da ação acrática deve preservar os elementos de racionalidade envolvidos na ação acrática sem interpretar as ações, elas próprias, racionais.

A ação acrática é subjetivamente irracional, dado que é irracional do ponto de vista do agente, por ser contrário ao seu melhor juízo. No entanto, os elementos da explicação racional aplicam-se ao comportamento acrático, pois os agentes acratas agem por razões. A questão é que as razões pelas quais agem, não são as razões pelas quais eles julgam melhor agir. Quando um agente faz *a*, enquanto conscientemente detém o melhor juízo decisivo a favor de fazer a ação competitiva *b*, o agente está claramente mais motivado a fazer *a* do que a fazer *b*. Explicar como o agente chegou a este balanço de motivação é conseguir explicar a ação acrática¹²².

3.1.13. Hipótese modificada da proximidade¹²³

Em relação à proximidade da recompensa, Alfred Mele tinha defendido já em *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, que a percepção da proximidade de recompensa pode afetar a atenção, de uma forma tal, que a motivação para alcançar a recompensa aumenta, bem como atenua a motivação para alcançar recompensas mais distantes.

O agente pode ter falhado aqui no exercício do *self-control*, por várias razões: não lhe ocorreu o exercício do *self-control*; exerceu o *self-control*, mas de forma inadequada; ou, contra o seu melhor juízo, fez a vontade a si próprio e refreou o exercício do *self-control*. A *modified proximity hypothesis* vê incorporado o elemento atencional na hipótese de proximidade, noção a partir da qual muitas *strict akratic actions*, segundo Alfred Mele,

¹²¹ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 48.

¹²² MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 75-84.

¹²³ Este conceito surge da modificação que Alfred Mele faz da noção de George Ainslie e Amelie Rorty. Segundo G. Ainslie, existe uma força motivacional da proximidade de recompensa que gera uma explicação a que A. Mele chamou de *proximity explanation*, na qual intervêm quatro elementos: 1. O nível de motivação do agente para fazer *c* (ação continente possível), 2. Nível de motivação inicial para fazer *a* (ação acrática), 3. Falha do agente no uso efetivo de *self-control*, 4. Proximidade. Segundo Amelie Rorty, uma das estratégias atrativas da acrasia é a estratégia atencional, que consiste no facto de que, o balanço de motivação do agente, no momento da ação, é por vezes influenciado de forma importante, pelo foco da sua atenção. Este conceito, segundo Alfred Mele, ajudará a explicar a influência motivacional da proximidade percebida.

são adequadamente explicadas nos seguintes termos: 1. a proximidade percebida de recompensa, derivada da alternativa incontinente; 2. o nível de motivação do agente para a alternativa continente e o seu nível inicial de motivação para fazer a alternativa acrática; 3. a falha do agente de fazer uma tentativa eficaz de *self-control*; 4. a condição atencional do agente.

Uma explicação baseada nestas condições, poderá explicar a ocorrência de uma *strict akratic action*, explicando o balanço de motivação do agente no momento da ação. Alfred Mele chama a atenção para o facto de que, os itens envolvidos na explicação modificada de proximidade são, eles próprios, sujeitos a explicação em casos particulares (tal como o são as crenças, desejos, intenções etc., em casos de explicação racional). Por outro lado, os quatro elementos mencionados podem ter diferentes níveis de influência.

A proximidade percebida pode afetar a condição atencional do agente, que por sua vez pode contribuir para a falha no *self-control* do agente e para a atenuação da motivação do agente para fazer a alternativa continente¹²⁴.

3.1.14. O papel das intenções

Se as intenções são pensadas como mediadoras entre as razões para a ação e a ação intencional, então Alfred Mele pretende examinar as implicações da ação acrática contra as intenções de aqui e agora.

Segundo Alfred Mele, as intenções aqui e agora são derrotadas, ou não, pela motivação adversária, dependendo do balanço da motivação do agente na altura da ação.

Com efeito, não é possível explicar adequadamente uma ação de um agente, seja continente ou incontinente, na presença de motivação competitiva, explicando a sua intenção para fazer determinada ação, uma vez que as intenções podem ser derrotadas por motivação contrária. Por outro lado, explicar porque o agente agiu segundo a sua intenção de *a* no caso de conflito motivacional, passa por explicar porque o agente estava mais motivado para fazer *a* na altura. A intenção faz a mediação entre a motivação global e ação intencional, facto que é compatível com a ideia de que, intencional, tem uma componente motivacional.

¹²⁴ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 84-92.

Com efeito, uma intenção pode mediar entre a motivação que contribui para a sua formação e a ação intencional, que a intenção vai gerar. E esta motivação pode ser a fonte da componente motivacional da intenção.

Alfred Mele acrescenta ainda que a intenção tem um papel cognitivo crucial na iniciação e guia da ação intencional, por exemplo, ao incorporar a representação daquilo que é intencionado e que ajuda a guiar os movimentos do agente.

Relativamente à *strict akratic action*, se as intenções mediam entre razões e ação intencional, então as intenções aqui e agora, derrotadas por uma ação acrática, são substituídas por intenções competitivas. E são precisamente estas intenções competitivas que geram as ações acráticas.

Para Alfred Mele, o auto controlo envolve a motivação para a ação, conforme o agente julga melhor, bem como um conjunto de competências e poder de resistência. A motivação e o exercício das capacidades de resistência do agente, podem modificar a condição motivacional na altura da ação e, portanto, o comportamento intencional. Através do exercício de *self-control*, o agente pode intervir no seu próprio processo motivacional gerando o balanço de motivação favorável à ação continente¹²⁵.

Esta intervenção que Alfred Mele refere, poderá ser considerada em analogia com a aplicação de mecanismos, noção na qual se centra Jon Elster.

Jon Elster efetua análises interdisciplinares e procura de mecanismos e entende que as ciências sociais não podem encontrar leis¹²⁶. Aliás, refere que o antónimo de mecanismo é a Lei científica (A lei estabelece que perante certas condições iniciais, um determinado evento – a causa – vai sempre produzir um evento de outro tipo – o efeito)¹²⁷

Em alternativa, devem focar-se nos mecanismos da racionalidade e irracionalidade, mecanismos estes que têm um poder explicativo.

Estes mecanismos dizem respeito ao eu e à sua relação consigo próprio, através, por exemplo, de auto imposição de restrições à nossa liberdade de escolha. Estas relações

¹²⁵ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 102-108.

¹²⁶ CADILHA, S, MIGUENS, (coord). *Ação e Ética. Conversas sobre racionalidade prática*. Edições Colibri, 2011, pp. 61-64.

¹²⁷ ELSTER, Jon. *Alchemies of the Mind. Rationality and the Emotions*. Cambridge University Press, United Kingdom 1999, p. 5.

referem-se à posição de cada um de nós ao longo do tempo e à nossa relação com o nosso próprio eu, no futuro.

Segundo Jon Elster, o raciocínio por mecanismos¹²⁸, permite explicar, mas não prever¹²⁹. Com efeito, os mecanismos são padrões causais, frequentes e facilmente reconhecíveis, desencadeados, geralmente, sob condições desconhecidas ou com consequências indeterminadas¹³⁰.

Jon Elster contrasta mecanismos do tipo B (os que ocorrem quando conseguimos prever o desencadeamento de duas cadeias causais, que afetam uma variável independente em direções opostas¹³¹) com mecanismos do tipo A (os que ocorrem quando a indeterminação se relaciona a qual (se existir alguma) das várias cadeias causais poderão ser desencadeadas¹³²).

Quando se refere à explicação por mecanismos, Jon Elster explica que uma Lei estabelece que, dadas determinadas condições iniciais, um evento de certo tipo (a causa), vai sempre produzir um evento de outro tipo (o efeito). Assim, sugere que a teoria para os mecanismos, move-se da expressão: se *A* então sempre *B*, para a expressão: Se *A*, então *B*. E sugere ainda a expressão; Se *A*, então por vezes *C*, *D* e *B*, para assegurar explicações mais detalhadas¹³³.

Jon Elster não defende a explicação por mecanismos como um ideal ou norma, admitindo que a explicação por Lei é melhor, mas muito mais difícil. Por outro lado também não sugere que os mecanismos sejam identificados por condições formais, análogas às que existem na formulação das Leis: Se *p*, então por vezes *q*, que considera ser um *insight* inútil.

¹²⁸ Por exemplo, mecanismo exemplificado na fábula das uvas amargas de La Fontaine: a raposa esfomeada, vê ao longe umas uvas apetitosas, mas que sabe nunca lhes vai chegar. Então diz a si própria que estão demasiado verdes para comer.

¹²⁹ CADILHA, S, MIGUENS, S., (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op.cit., pp. 61-64.

¹³⁰ ELSTER, *Alchemies of the Mind. Rationality and the Emotions*, op. cit., p. 1.

¹³¹ Por exemplo, plantei camélias em solo favorável a estas flores, mas plantei quando o solo anda estava quente, facto que sabia ser desfavorável. Quando as camélias morrerem eu sei que foi devido ao solo quente.

¹³² Por exemplo, um estímulo ambiental pode desencadear num animal a fuga, a luta ou paralisia. Cada um destas reações vai desencadear um comportamento diferente. Sabemos o que desencadeia cada reação, mas não sabemos o que vai desencadear a fuga em vez da luta, por exemplo.

¹³³ ELSTER, *Alchemies of the Mind. Rationality and the Emotions*, op. cit., pp. 3-6.

A explicação por mecanismos funciona quando e porque podemos identificar um padrão causal determinado, que reconhecemos ao longo de situações e que constitui uma resposta inteligente à questão: porque motivo ele fez aquilo¹³⁴?

Jon Elster refere dois tipos de restrições benéficas ao agente. As restrições incidentais – *incidental constraints* – constrangimentos que beneficiam o agente constrangido, mas que não são escolhidos pelo agente em função desses benefícios. As restrições podem ser escolhidas por outras razões, ou por outro agente, ou mesmo não ser escolhido por ninguém, por resultar simplesmente de ser um facto da vida que o agente tem que respeitar (por exemplo, a necessidade de gravar filmes a preto e branco antes da invenção da gravação a cores).

Por outro lado existem as restrições essenciais – *essential constraints* – que são as restrições que o agente impõe a si próprio, em função de algum benefício que espera para si próprio. A este fenómeno o autor chama de *precommitment* ou *self-binding*.

As restrições essenciais correspondem ao desejo do agente de criar obstáculos para a sua escolha futura de alguma opção, ou opções, específicas.

Nesta perspectiva, o *precommitment* envolve uma certa racionalidade ao longo do tempo. Num tempo 1 o sujeito quer fazer *A* no momento 2, mas antecipa que quando chegar o momento 2, pode querer ou fazer *B*, a menos que esteja prevenido de o fazer. Nesses casos, o comportamento racional no momento 1 pode envolver medidas de precaução para prevenir a escolha *B* no momento 2, ou pelo menos, tornar essa opção menos provável¹³⁵.

Jon Elster faz uma análise das razões e das formas pelas quais as pessoas se podem pré comprometer a si próprios e dos mecanismos disponíveis para esse fim.

¹³⁴ ELSTER, *Alchemies of the Mind. Rationality and the Emotions*, op. cit., p. 10.

¹³⁵ O compromisso prévio é ilustrado pelo exemplo de Ulisses e das Sereias. Neste exemplo, Ulisses pensa que, em princípio, não conseguirá resistir ao canto das sereias num futuro próximo e por esse motivo pede para o amarrarem ao mastro da embarcação. Assim, o seu eu futuro estará preso quando ouvir as sereias, permitindo-lhe resistir. Este caso ilustra o mecanismo do compromisso prévio, estratégia que permite ao indivíduo lidar com situações que se prevê uma eventual falta de controlo. (CADILHA, S, MIGUENS, (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op. cit., p. 63).

Assim o conjunto de razões seria: superar a paixão, superar o auto interesse, superar o desconto hiperbólico, superar a inconsistência temporal¹³⁶ estratégica e neutralizar ou prevenir a alteração das preferências. E os mecanismos disponíveis são os seguintes: eliminação de opções, imposição de custos, estabelecimento de recompensas, criação de atrasos, alteração de preferências, investimento do poder de negociação, induzir ignorância e induzir paixão.

Estes tópicos indicam apenas algumas das principais conexões entre razões e mecanismos de pré compromisso, mas nem todos os mecanismos de pré compromisso servem todas as razões para o pré compromisso¹³⁷.

As competências a que Alfred Mele se refere, quando fala de capacidade de resistência podem concretizar-se em mecanismos específicos, tais como os presentes no pensamento de Jon Elster, mecanismos esses que dependerão, tanto da capacidade de auto controlo do agente, bem como da sua motivação para o exercício desses mecanismos de resistência e, consequentemente, a motivação para a ação congruente com o comportamento intencional favorável à ação continente. Ao exercer o autocontrolo, o agente envolver-se-á num *precommitment*, que o liga à ação racional.

3.1.15. Sentimentos Acráticos

Sem dúvida alguma que as emoções têm um importante papel nas nossas vidas e nesta matéria, Alfred Mele centra-se nos *feeling-tokens*, que incluem *emotion-token*, ou seja, instâncias particulares de medo, alegria, tristeza, ansiedade e outras instâncias análogas. Foca-se portanto em estados de curta duração.

Para exemplificar a que se refere, em concreto, Alfred Mele relata a história de um jovem que se despreza a si próprio por amar uma mulher superficial e distante. Este jovem está convencido de que o seu amor por aquela pessoa, prejudica o seu bem-estar e tenta erradicar essa emoção de si mesmo. Mas os seus esforços falham e não consegue livrar-

¹³⁶ A inconsistência temporal ocorre quando a melhor política planeada no momento, para o futuro, já não se revela ser o melhor quando chega esse momento. A reversão de preferência envolvida na inconsistência temporal não se deve a mudanças ambientais exógenas ou imprevistas, nem a uma mudança subjetiva no sujeito. A reversão é causada apenas pela mera passagem do tempo.

¹³⁷ ELSTER, Jon. *Ulysses Unbound*. Cambridge University Press, United Kingdom 2000, pp. 4-7.

se daquela paixão que o obceca e se torna degradante. Ou seja, o jovem falha em alinhar os seus sentimentos, de acordo com o seu melhor juízo.

Coloca então a questão: como pode um agente evitar adquirir ou continuar a sentir algo que vai contra aquilo que julga melhor, ou evitar não sentir algo que o agente entende que seria melhor sentir?

Uma resposta natural seria que os nossos melhores juízos acerca dos sentimentos não têm que estar necessariamente alinhados com os determinantes desses estados de sentimentos (incluindo também estados de não ter certos sentimentos).

Mas, tal como na ação acrática, temos que perceber de que forma é possível esta disparidade e a resposta não pode ser a de que não temos qualquer controlo sobre o que sentimos. Tal como a ação acrática se distingue de ações compelidas, efetuadas contra o melhor juízo do agente, os sentimentos acrátricos, concebidos da mesma forma, também não são compelidos¹³⁸.

Os desejos que motivam as ações acrátricas e as ações daí resultantes estão sujeitos ao controlo do agente: apesar do agente poder controlar os desejos que o levam à ação acrática, o sujeito falha em fazê-lo. De forma similar, as emoções acrátricas e outros sentimentos, que sejam genuinamente acrátricos, devem ser sujeitos a um controlo apropriado pelo agente, que deve ter à sua disposição meios apropriados de prevenir ou erradicar tais sentimentos ou alinhar a intensidade desses sentimentos com os seus melhores juízos.

Se os melhores juízos não têm qualquer influência relevante nos estados de sentimento – *feeling states* – o agente pode estar completamente desprovido de controlo racional por esses sentimentos. Ou seja, pode ser que, quanto mais falha for a conexão entre melhor juízo e sentimento, menor controlo racional temos sobre aquilo que sentimos ou não¹³⁹.

¹³⁸ Distinguímos os viciados em heroína de sujeitos que, na procura de adrenalina, utilizam acriticamente, drogas pela primeira vez. Apesar de ambos os grupos agirem contra o seu melhor juízo, o primeiro possui um desejo compelido irresistível por heroína.

¹³⁹ MELE, *Autonomous Agents*, op. cit., pp. 102-104.

3.1.15.1. *Strict Akratic Feelings*

Os *Strict Akratic Feelings*, tal como as *Strict Akratic Actions*, estão em desacordo com o melhor juízo do agente.

Devemos atender as várias perspectivas pelas quais as emoções e outros sentimentos, podem ser avaliados, de forma a perceber que tipo de melhores juízos poderão ser vencidos por um sentimento e para isso, Alfred Mele retoma a noção de crença.

Quando considerarmos uma crença como garantida, tipicamente temos em mente o que pode ser denominado de garantias evidenciais, pois as crenças visam a verdade e a evidência será a melhor guia.

Mas, mesmo crenças falsas e não garantidas por evidência podem trazer vantagens práticas, conforme aquilo que, em parte, ocorre no auto engano¹⁴⁰. Parece ser possível que uma crença considerada como sendo melhor manter, consideradas todas as coisas, por vezes carece de evidência suficiente.

Quando falamos de um sentimento ou emoção garantidos, temos em mente uma emoção ou sentimento que se constituem como uma resposta apropriada ao estímulo¹⁴¹. Além disso atribuímos a muitos sentimentos um valor moral¹⁴². Assim, muitas vezes, o facto de ser apropriado, o valor de prudência e a moralidade apontam na mesma direção.

Mas nem sempre as coisas se passam desta forma, por exemplo, é possível que um agente considere que o sentimento de raiva e ressentimento, podem ser respostas apropriadas ao facto de uma esposa ter um caso extra conjugal, mas, caso o marido pretenda salvar o seu casamento, então estes sentimentos podem ser imprudentes.

O melhor juízo decisivo na esfera dos sentimentos, tal como o melhor juízo decisivo na esfera da ação, pode ser formado com base numa variedade de considerações.

Se as emoções fossem ações, a conceção de ação akrática poderia ser diretamente aplicável às emoções, mas os sentimentos e as crenças akráticas não são ações akráticas.

Portanto, uma conceção apropriada de *strict akratic feeling* tem que acomodar esta diferença: estando ou permanecendo o sujeito *S* num estado de sentimento *x*, durante um

¹⁴⁰ Por exemplo, se um sujeito acredita na fidelidade da sua esposa, isso pode permitir o seu normal funcionamento, mesmo sem suporte evidencial.

¹⁴¹ Por exemplo, o medo pode ser uma resposta apropriada perante determinada ameaça, na medida em que leva o agente a tomar medidas de prevenção. Existe portanto um valor de prudência.

¹⁴² É moralmente apropriado sentirmos simpatia em algumas circunstâncias e indignados em outras.

tempo t , constitui uma instância de *strict akratic feeling*, se e só se, o facto do sujeito S estar ou permanecer em x durante o tempo t , é não forçado e durante o tempo t , o sujeito S detém, conscientemente, o juízo de que existe uma razão boa e suficiente para não estar ou permanecer em x . (de novo, a noção de estados de sentimento inclui estados de não ter determinado sentimento)¹⁴³.

Então coloca-se a seguinte questão: todos os estados de sentimento, em desacordo com o melhor juízo detido conscientemente pelo agente, estão além do seu controlo?

Parece claro que temos algum controlo sobre o que sentimos e sobre a intensidade das nossas emoções e outros sentimentos. O funcionário tímido que acredita que apenas consegue ter coragem de pedir um aumento, caso se zangue com o seu superior, pode deliberadamente zangar-se, ao representar vividamente a injustiça que vem a sentir ao longo do tempo. Este é um exemplo de controlo interno, mas muitas emoções são sujeitas a controlo externo, controlo efetuado através de uma ação manifesta. Por exemplo, o Jack sabe que consegue moderar a sua depressão telefonando para a sua irmã¹⁴⁴.

De forma a ser possível defender e entender a possibilidade da *strict akratic feeling*, devemos ter presente a relação entre melhores juízos e os determinantes da ação e crença. Por vezes, os nossos melhores juízos acerca da ação residem significativamente nas nossas avaliações dos objetos dos nossos desejos ou vontades, mas conforme já vimos, a avaliação de um agente acerca de um item desejado não tem que estar alinhada com a força motivacional do desejo¹⁴⁵. Similarmente, a avaliação que sustém o melhor juízo acerca de crenças, não tem que determinar nem medir precisamente o poder causal dos itens que influenciam as crenças¹⁴⁶.

A etiologia de muitas emoções e outros sentimentos é bastante complexa. As nossas emoções são, em muitos casos, produto de desejos, hábitos de interpretação, padrões de resposta emocional adquiridos, outras emoções, a nossa condição fisiológica, etc. A possibilidade psicológica das *Strict Akratic Feelings*, residem na possibilidade

¹⁴³ MELE, *Autonomous Agents*, op. cit., pp. 104-105.

¹⁴⁴ MELE, *Autonomous Agents*, op. cit., p. 106.

¹⁴⁵ Posso ter o melhor juízo de que é melhor deixar de fumar, mas estar motivado para continuar.

¹⁴⁶ Uma classificação relativamente baixa em relação ao mérito, pode, devido à vividez da representação do sujeito acerca dessa classificação, exercer uma influência desproporcional naquilo em que o sujeito acredita.

psicológica de que, avaliações que sustentam alguns melhores juízos, acerca de questões de sentimento não resolvem completamente, nem são indicadores do poder causal dos itens influenciadores das crenças. Se existir uma falta de correspondência entre os determinantes do sentimento de um agente e o seu melhor juízo, e se, apesar desta disparidade, o que o agente sente ou continua a sentir está sujeito ao seu controlo, então está aberta a possibilidade para o sentimento acrático.

A possibilidade da *Strict Akratic Feeling*, não depende da derrota da motivação hedónica de um melhor juízo. Os determinantes de um sentimento podem estender-se além dos itens considerados na formação de melhor juízo decisivo pertinente do agente e esses determinantes podem produzir sentimentos desagradáveis e não desejados¹⁴⁷.

Alfred Mele ilumina estas conclusões através do exemplo de Jane, que descobriu que o marido tem um caso extra conjugal. A Jane acredita que a fúria e o ressentimento são respostas apropriadas e moralmente adequadas e acredita ainda que estas respostas podem favorecer a hipótese de salvar o seu casamento, conforme ela deseja. A Jane julga existirem razões boas e suficientes para manter estas emoções. Mas esta situação apenas aumenta a confusão do seu estado de sentimento, uma vez que, de facto, não são estas as emoções experienciadas, mas sim um sentimento de inadequação e este sentimento persiste, apesar de Jane acreditar que é inapropriado e contra produtivo. O melhor juízo de Jane, é que seria melhor não se sentir inadequada. Mas, se olharmos para a sua história, ficamos a saber que os seus pais se divorciaram quando ainda era criança, tendo-se sentido responsável por isso. Jane acha que se tivesse sido melhor filha, os seus pais teriam ficado juntos, ou seja, sentiu-se inadequada. E, apesar de vir a saber que foi o pai que deixou a mãe por outra pessoa, Jane culpou a mãe.

Podemos depreender daqui que a situação do marido de Jane, veio abrir feridas antigas, reativando um padrão de auto recriminação e talvez até tenha desenvolvido uma associação entre Jane e o caso da sua mãe.

Assim, a avaliação das considerações que Jane teve em conta ao formar o seu melhor juízo, não resolveram nem foram indicadores do poder causal dos determinantes do seu

¹⁴⁷ Por exemplo, hábitos controlados de auto recriminação podem produzir sentimentos de culpa, medo, ou inadequação, que estão em desacordo com o melhor juízo decisivo.

estado de sentimento. Aqui, tal como no caso da ação acrática, a avaliação pode tomar um sentido oposto ao da força causal preponderante dos itens pertinentes.

Mas, apesar desta disparidade, esta situação não transforma o agente numa vítima indefesa. A Jane tinha ao seu dispor, uma variedade de técnicas que lhe permitiam alinhar o seu estado de sentimento com o seu melhor juízo. Poderia ter adotado estratégias de focar a atenção no seu valor e focar atenção nos defeitos do marido, por exemplo, ou procurar compreender a fonte profunda dos seus sentimentos de inadequação, procurado suporte junto de amigos, etc. Supondo que Jane não deseja, de facto, sentir-se inadequada, ela falhou em corrigir o seu estado de sentimento, ou seja, apesar de julgar melhor tentar livrar-se do sentimento, Jane acriticamente falhou em levar esse plano avante.

Mas esta situação não leva a concluir que o *strict akratic feeling* deriva da *strict akratic action* (*strict akratic action* associada com o facto do agente não alinhar os seus sentimentos com o melhor juízo), porque a falha em erradicar o sentimento não é uma *strict akratic action*. As *strict akratic actions* violam os melhores juízos práticos, juízos acerca do que é melhor fazer (ou abster-se de fazer). Talvez simplesmente não ocorra à pessoa, que deve tomar medidas para corrigir o seu estado de sentimento.

Alfred Mele faz ainda três considerações, em relação às *strict akratic actions*: o grau segundo o qual as nossas emoções e outros sentimentos estão sujeitos ao nosso controlo; a frequência com que fazemos melhores juízos decisivos acerca dos nossos sentimentos e a frequência com que as nossas emoções e outros sentimentos falham, em adaptar-se a esses julgamentos. A investigação empírica indica que temos controlo considerável sobre uma boa parte das nossas emoções e sentimentos.

Para Alfred Mele, quanto mais influência nos obrigarmos a ter, no que diz respeito ao que sentimos e não sentimos, mais provável será tomarmos melhores juízos acerca dos nossos sentimentos. E dada a profunda influência das emoções e outros sentimentos no comportamento, uma intervenção de sucesso na vida afetiva de cada um de nós, tenderá a aumentar a qualidade de vida de um modo geral.

Se os agentes podem acriticamente entrar ou permanecer em estados de sentimento que vão contra os seus melhores juízos decisivos, então os agentes podem exercer auto

controle em favor de estados de sentimento, recomendados pelos melhores julgamentos decisivos¹⁴⁸.

Existem na literatura, outros autores que defendem e explicam a influência das emoções nas nossas decisões, intenções e ações.

Jon Elster, por exemplo, refere-se aos efeitos da paixão, noção que engloba as emoções e estados tais como a embriaguez ou a dor. Quando um sujeito age sob influência de uma paixão, poderá desviar-se dos planos delineados num momento calmo e ter esta consciência, cria um incentivo a comprometermo-nos a nos próprios de forma a garantir a execução dos planos criados¹⁴⁹. Ou seja, esta ideia alinha-se com a de Alfred Mele, quando refere que o agente pode exercer auto controle, a favor de estados de sentimento recomendados pelos melhores julgamentos decisivos.

Segundo Jon Elster, existem quatro formas segundo as quais as paixões poderão criar discrepâncias entre planos: distorcendo a cognição (criação de falsas crenças acerca das consequências); turvando a cognição (deturpação da consciência das consequências); induzindo fraqueza de vontade (opções com piores consequências percebidas, são escolhidas em vez das opções com melhores consequências percebidas); induzindo miopia (alteração dos pesos de decisão associados às consequências). Os dois primeiros mecanismos envolvem irracionalidade cognitiva e o terceiro envolve irracionalidade motivada.

A paixão pode distorcer o nosso pensamento acerca das consequências do comportamento e essa paixão pode ser de tal forma forte, que nos leva a ignorar todas as outras considerações¹⁵⁰.

Esta posição de Jon Elster estabelece desde já uma ponte com o fenómeno do auto engano, que veremos mais adiante.

Segundo Jon Elster, estas discrepâncias, sendo diferentes, têm naturalmente implicações nas estratégias de auto compromisso – *self binding*. Por exemplo, quando um sujeito consegue ter os incentivos em conta, mesmo que se encontre no calor da paixão, então o pré-compromisso pode tomar a forma de custo a aplicar a uma escolha que o sujeito

¹⁴⁸ MELE, *Autonomous Agents*, op. cit., pp. 108-111.

¹⁴⁹ ELSTER, *Ulysses Unbound*, op. cit., p. 7.

¹⁵⁰ ELSTER, *Ulysses Unbound*, op. cit., pp. 7-10.

pretende evitar. Mas existem paixões de tal forma fortes, que a única forma prática de as neutralizar será evitar as ocasiões que as despoletam¹⁵¹.

Sobre esta matéria e numa outra perspetiva, também António Damásio distingue três fases que fazem parte de um processo contínuo: o estado de emoção, que pode ser executado e iniciado de forma inconsciente; o estado de sentimento, que pode ser representado de forma inconsciente e o estado de sentimento tornado consciente, ou seja o estado que é conhecido pelo organismo onde se passa a emoção e o sentimento.¹⁵²

Defende que nos últimos anos a neurociência e a ciência cognitiva passaram a ocupar-se das emoções, tendo deixado de ser aceite a suposta oposição entre emoção e razão. Aliás António Damásio refere que os seus trabalhos laboratoriais demonstraram que a emoção integra os processos de raciocínio e tomada de decisão¹⁵³.

Assim, uma redução seletiva das emoções é considerada prejudicial, tanto quanto a emoção excessiva, para a racionalidade do sujeito. Fica assim sugerido que a emoção, bem dirigida, suporta a razão eficaz, não a substituindo como é óbvio, pois um estado de perturbação emocional pode levar a decisões irracionais¹⁵⁴.

As emoções derivam de uma base biológica, na medida em que constituem um conjunto de respostas químicas e neurais, que desempenham um papel regulador, cuja finalidade é ajudar o organismo a manter a vida. As emoções são processos determinados biologicamente, apesar de a sua expressão ser passível de alteração pela aprendizagem e pela cultura. Os dispositivos responsáveis pela produção das emoções, localizam-se num conjunto restrito de regiões cerebrais e estes dispositivos podem ser ativados automaticamente, sem deliberação consciente. As emoções modificam o estado do corpo e também o estado cerebral¹⁵⁵.

Por outro lado, as emoções não se dissociam da nossa ideia de recompensa ou castigo, prazer ou dor, aproximação ou afastamento, vantagem ou desvantagem pessoa, e do bem

¹⁵¹ ELSTER, *Ulysses Unbound*, op. cit., p. 11.

¹⁵² DAMÁSIO, António, *O Sentimento de Si*, Publicações Europa-América, Lda., Mira-Sintra 2004, p.57.

¹⁵³ Esta é a hipótese dos marcadores somáticos, segundo a qual pacientes com lesões em áreas da região pré-frontal, levava a uma alteração consciente da capacidade de decidir vantajosamente em situações de risco e conflito, bem como uma redução seletiva da ressonância emocional.

¹⁵⁴ DAMÁSIO, *O Sentimento de Si*, op. cit. pp.61-62.

¹⁵⁵ DAMÁSIO, *O Sentimento de Si*, op. cit. pp 72-73.

e do mal¹⁵⁶. Esta afirmação parece corroborar o pensamento de Alfred Mele, quando defende que, quanto mais influência o agente se obriga a ter, em relação aos seus sentimentos, maior probabilidade existirá para assumirmos o melhor juízo.

António Damásio explica os diversos mecanismos das emoções. Alguns são evidentes, tais como as mãos húmidas traduzirem ansiedade, ou na quase paragem do coração perante um momento de terror. Outras não são visíveis e referem-se às modificações que ocorrem no corpo, por exemplo, a libertação de neurotransmissores e neuromodeladores. Durante as emoções, determinados neurónios transmitem estas substâncias químicas para neurónios de outras regiões cerebrais, facto que transforma, temporariamente, a forma de funcionamento de muitos circuitos neurais. Neste processo, pode existir como consequência, a sensação de aumento ou diminuição da velocidade dos processos mentais ou sensações de prazer ou desprazer¹⁵⁷.

António Damásio identifica as alterações cerebrais mais importantes que derivam da libertação de neuromodeladores: indução de determinados comportamentos (estabelecimento de laços familiares e sociais, por exemplo); alteração de estados corporais em curso (os estados corporais podem ser seletivamente inibidos ou realçados, por exemplo); alteração do modo de processamento cognitivo (modificação de imagens visuais de um modo lento para rápido, ou modificação de imagens nítidas para desfocadas).

Os estados emocionais são responsáveis por muitas alterações químicas corporais, mas também por modificações no conjunto de estruturas neurais¹⁵⁸. Esta descrição parece pertinente, porque permite corroborar em termos neurológicos, a tese de Alfred Mele, relativamente ao auto controlo das emoções do agente.

Conforme Alfred Mele defende, também António Damásio refere que, apesar das emoções serem úteis, o processo de sentir, fornece ao organismo o incentivo para instanciar o resultado das emoções. O sentimento de conhecer o que sentimos permite planear as respostas específicas, feitas à medida de cada situação, que complementam

¹⁵⁶ DAMÁSIO, *O Sentimento de Si*, op. cit. p. 77.

¹⁵⁷ DAMÁSIO, *O Sentimento de Si*, op. cit., pp. 81-83.

¹⁵⁸ DAMÁSIO, *O Sentimento de Si*, op. cit., p. 322.

essa emoção e garantem a preservação ao longo do tempo, das vantagens imediatas obtidas pela emoção¹⁵⁹.

Para António Damásio é indiscutível a importância das emoções, nos fenómenos de racionalidade prática, nomeadamente no planeamento das ações próprias e decisões.

Foi a partir dos mecanismos biológicos de regulação corporal, aos quais se associam as emoções, que a natureza construiu o dispositivo cerebral para a racionalidade, sendo que o mecanismo de racionalidade prática fica perturbado, na ausência do normal funcionamento das emoções¹⁶⁰.

As emoções são desencadeadas através de imagens de objetos ou acontecimentos, que estão a ocorrer nesse momento, ou que são invocadas, ou mesmo imaginadas, pelo sujeito. Portanto, a situação contextual é central para o funcionamento do aparelho emocional¹⁶¹. Mais uma vez, e em sintonia com o pensamento de Alfred Mele, António Damásio defende que, apesar das emoções serem geneticamente determinadas, ou seja, constituírem um programa de ação inato e automático, nada obsta ao seu molde pelo sujeito, facto que alude ao auto controlo defendido por Alfred Mele.

Por outro lado, importa referir que as emoções variam de sujeito para sujeito: por exemplo, há coisas que eu receio e outras pessoas não. Além de individuais, as emoções também são reguladas pelo contexto social e educação individual. Portanto, as emoções podem ser reguladas voluntariamente, sendo que esta regulação se passa apenas ao nível das manifestações externas.

António Damásio diferencia as emoções de fundo, das quais são exemplos o entusiasmo e o desencorajamento das emoções sociais, tais como, compaixão, embaraço, vergonha, culpa, desprezo, ciúme, inveja orgulho, admiração. As emoções de fundo podem ser ativadas por circunstâncias contextuais, mas também por estados internos, como uma doença ou fadiga. As emoções sociais funcionam na vida social do sujeito e revestem-se de princípios morais, formando uma base natural para os sistemas éticos¹⁶².

¹⁵⁹ DAMÁSIO, *O Sentimento de Si*, op. cit., p. 325.

¹⁶⁰ MIGUENS, Sofia, *Compreender a Mente e o Conhecimento*. Mlag Discussion Papers., Porto 2009, p. 207.

¹⁶¹ DAMÁSIO, *O Livro da Consciência*, op. cit., pp. 140-145.

¹⁶² DAMÁSIO, *O Livro da Consciência*, op. cit., pp. 159-162.

3.1.16. A decisão

Alfred Mele designa por decisão prática, o facto de o agente decidir agir e por decisão cognitiva o facto de o agente decidir, que algo é o caso. Tem ainda em conta decisões de não fazer *A*, por exemplo, não votar numa eleição, como uma decisão prática¹⁶³.

Alfred Mele assume que as decisões práticas são ações mentais momentâneas, de formação de intenções¹⁶⁴.

Se um agente adquire a intenção para decidir *A*, ou o agente a adquire de forma não ativa ou o agente decide, decidir *A*. A última disjunção implica uma regressão impossível, que pode ser eliminada supondo que, em determinado ponto, o agente adquire não ativamente a intenção decisional pertinente, ou seja, a intenção de decidir, que vai decidir *A*. No entanto, considerando que acabaremos, em qualquer dos casos, com uma intenção decisional não ativa, parece melhor considerarmos a primeira disjunção.

Mas, porque razão pode acontecer que um agente adquira a intenção não ativa de decidir *A*, em vez de adquirir a intenção não ativa de *A*?

Para o autor, um agente adquire a intenção não ativa de decidir *A*, apenas se existirem alguns obstáculos à sua aquisição não ativa da intenção de *A*, que também não é um obstáculo para a sua aquisição não ativa de intenção para decidir *A*.

Em casos normais, não existe razão para adquirir não ativamente uma intenção para decidir *A*, que não seja servida, mais eficientemente, por uma aquisição não ativa de uma intenção de *A*. Na ausência de uma razão, que suponha que a nossa vida mental não ativa é, normalmente, ineficientemente complicada nesta matéria, a assunção de maior eficiência e simplicidade é bastante plausível.

Mas, se as intenções para decidir *A* não desempenham, normalmente, um papel na produção das nossas decisões de *A*, devemos então abandonar a ideia de que desempenhamos ações na decisão prática?

Existem muitas coisas que decidimos intencionalmente, sem ter intenções específicas de o fazer. Quando caminho para o trabalho, por exemplo, os meus passos individuais são

¹⁶³ MELE, Alfred *Aspects of Agency. Decisions, Abilities, Explanations, and Free Will*. Oxford University Press, New York 2017, p. 7.

¹⁶⁴ MELE, *Motivation and Agency*, op. cit., p. 202.

ações intencionais, mas não existe necessidade de supor que cada passo requer a sua própria intenção distinta. Neste caso típico, uma intenção única mais geral, a intenção de caminhar para o trabalho, pelo caminho habitual, é intenção suficiente.

Se um agente pode fazer intencionalmente *A* sem ter a intenção específica de fazer *A*, talvez esse agente possa decidir *B* sem ter a intenção de decidir *B*. Se determinado tipo de intenção está normalmente a funcionar na produção das decisões, então pode ser uma intenção de decidir o que fazer. Num agente normal que decide *B*, na ausência de uma intenção de decidir *B*, uma intenção de decidir o que fazer, pode ter uma contribuição importante na produção dessa decisão. A ideia de que em todas as instâncias de decisão prática, algumas intenções, ou outras, desempenham um papel produtivo, não têm a implicação infeliz de que decisões para *B* são sempre ou normalmente produzidas (em parte), por intenções para decidir *B*.

O desafio relativamente à afirmação de que existem ações mentais momentâneas na decisão prática consiste nas seguintes suposições: 1. decisões práticas são ações intencionais e 2. as intenções desempenham um papel na produção de todas as ações intencionais.

A preocupação é que não existem candidatos plausíveis para intenções que desempenham um papel produtivo na grande variedade de decisões práticas comuns. Apesar de inclinado a aceitar a segunda proposição, em parte por questões de simplicidade, Alfred Mele defende a primeira proposição. Não desenvolvendo a segunda proposição, sugere que se entenda como verdadeira e, nesse caso, existe um tipo de intenção que está plausivelmente envolvida na produção das decisões práticas comuns – uma intenção para decidir o que fazer¹⁶⁵.

Mas, se as decisões práticas são ações mentais momentâneas de formação de intenções, é necessário defender que existem de facto, decisões práticas.

Por vezes decidimos fazer o que parece ser melhor, mas apenas o fazemos, às vezes. Outras vezes, tomamos decisões para as quais temos motivação mais forte na altura, mas alguém que tenha forte motivação para decidir *A* do que para decidir outra coisa qualquer, pode simplesmente continuar a deliberar acerca do que fazer.

¹⁶⁵ MELE, *Motivation and Agency*, op. cit., pp. 202-205.

Excetuando casos muito excepcionais, um agente que intencionalmente age, está a agir por uma ou outra razão. Se considerarmos que parte daquilo que é, o agir por uma razão, e que essa razão desempenha um papel na explicação da ação do agente, então poderíamos também acreditar, justificadamente, que as razões desempenham um papel na explicação das nossas ações intencionais comuns, ainda que num contexto determinista. A mesma analogia aplica-se ao facto de decidir por razões.

Alfred Mele não pretende afirmar que os deterministas deviam aceitar a visão de que as ações intencionais têm razões explicativas. Antes, afirma que 1. uma defesa apropriada da ação intencional deve revelar que, deixando de lado casos realmente excepcionais, as ações intencionais são efetuadas por razões; 2. uma análise apropriada de agir por uma razão revelará que essa ação requer que a razão relevante desempenhe um papel na explicação da ação; 3. ambos os pontos são consistentes com o determinismo causal.

Mesmo que descobríssemos que o determinismo é verdadeiro, iríamos continuar a acreditar que desempenhamos ações intencionais. Continuaríamos a acreditar que razões desempenham um papel na explicação das nossas ações intencionais comuns. E esta mesma analogia aplica-se às ações intencionais que são decisões práticas.

A alusão ao determinismo permite mostrar que a complexidade da tarefa de construir um esquema explanatório para a decisão prática, é consistente com a explicabilidade das decisões práticas. Num mundo determinista todos os eventos são explicáveis, pelo menos em termos de condições e leis antecedentes. O ponto em questão é que alguns destes eventos explicáveis podem ser ações intencionais, incluindo decisões práticas, e considerando a natureza das ações intencionais, as razões desempenham um papel na explicação também das ações não excepcionais¹⁶⁶.

3.1.16.1. A Decisão prática

Na tentativa de resolver uma incerteza prática particular, acerca de fazer *A* ou *B*, podemos fazer uma variedade de coisas. Podemos reunir informação relevante ou contratar alguém para o fazer, podemos avaliar ou analisar a informação ou pedir a um profissional que o faça. No entanto, mesmo um agente convencido de que detém toda a

¹⁶⁶ MELE, *Motivation and Agency*, op. cit., pp. 206-209.

informação relevante e acredite que seria inútil fazer análises ou deliberações adicionais, pode ficar inseguro acerca do que fazer.

A questão que se coloca é: o que é decidir fazer algo?

Para o autor, intenções para agir são atitudes executivas direcionadas para planos. Portanto, decidir agir – estabelecer ativamente um curso de ação - pode ser entendido como um ato mental do assentimento executivo de um plano de ação da primeira pessoa. Ao decidir agir, o sujeito forma uma intenção para agir e, ao fazê-lo, o sujeito tem presente que possui uma intenção que incorpora um plano. E o sujeito aprova esse plano ao executar uma decisão prática. A decisão prática é uma forma direta de resolver uma incerteza prática¹⁶⁷.

Tal como as crenças e desejos, as intenções têm conteúdo representacional e esse conteúdo representacional é um plano de ação. Delimitando a situação, um componente do plano de uma intenção tem um nó individual. É o caso, por exemplo, da representação futura de alguém que vai viajar até Lisboa no próximo inverno, que nada inclui acerca de meios para esse fim ou sobre atividades específicas de férias.

Mas por vezes, os planos com intenções incorporadas são mais complexos. Por exemplo, a intenção de verificar o seu *email* que a Jan executou esta noite, incorporou um plano que incluiu, clicar no ícone de email, inserir a palavra passe, clicar em prosseguir, etc. Um agente que executa com sucesso uma intenção, é guiado por um plano com intenções incorporadas.

Para Alfred Mele, a dimensão executiva de uma intenção, é intrínseca à orientação atitudinal da intenção. Podemos ter uma variedade de atitudes relativamente a planos: por exemplo, admirar o plano *x*, odiar o plano *y*, e desejar executar o plano *z*. Ter a atitude de intencionar um plano é estabelecer que o vamos executar.

A atitude de intencionar e desejar planos difere na medida em que o primeiro, por si só, inclui o facto de o agente estabelecer executar o plano, conforme referido. Alguém que deseja *A* ou deseja executar um determinado plano para fazer *A*, mesmo alguém que o deseje mais fortemente do que deseje não fazer *A*, ou não executar esse plano, pode estar

¹⁶⁷ MELE, *Motivation and Agency*, op. cit., pp. 209-210.

ainda a deliberar acerca de *A*, ou acerca de executar o plano, caso em que não estabeleceu o facto de fazer *A*, ou não estabeleceu executar o plano¹⁶⁸.

Na contribuição das intenções de *A*, para intencionalmente fazer *A*, é crucial considerar a característica de formar o estabelecimento das intenções, que não é capturável em termos de força de desejos nem em termos de crenças. As intenções para *A*, englobam essencialmente motivação para *A*, mas sem se tornarem numa redução a uma combinação de desejo e crença.

Parte do que deve ser estabelecido em fazer *A*, é ter uma atitude que engloba motivação relativamente a fazer *A*. Sem essa atitude o sujeito não possui o compromisso psicológico para fazer *A* que é intrínseco ao estabelecer fazer *A* e, em consequência, ter uma intenção para *A*¹⁶⁹.

Recordemos que, para Alfred Mele, as intenções para agir, são atitudes executivas direccionadas para planos.

Para Alfred Mele, podemos distinguir entre adquirir meramente tal atitude e estabelecer ativamente um curso de ação e, nesta linha, decidir *A* será esta última proposição. Decidir *A*, ou ativamente estabelecer um curso de ação, pode ser entendido como um ato mental de assentimento (aprovação) executivo, de um plano de ação da primeira pessoa. Se me for dito que o Mike é um excelente jogador de futebol e eu expressar o meu completo acordo, eu concordo com essa afirmação. Isto é uma concordância cognitiva.

Se me é proposto que devia ir ver o jogo do Mike hoje à noite e eu expresso a minha completa aceitação, então eu concordo com a proposta. Isto é uma concordância executiva: eu aceitei executar uma proposta para essa ação.

Agora, talvez o meu ato de aceitação a essa proposta tenha sido uma questão de dar voz a uma intenção adquirida não ativamente, de ir ver o jogo do Mike. Por exemplo, após ouvir a proposta eu podia não estar, de todo, incerta do que fazer e portanto imediatamente, eu adquiri não ativamente a intenção de ir ver o jogo e exprimi essa intenção num ato de aceitação da proposta. Ou, posso ter pesado os prós e contras e decidi

¹⁶⁸ Por exemplo, a Pat deseja mais fortemente responder a um recente insulto do que não o fazer, mas devido a questões morais está a deliberar se o deve fazer.

¹⁶⁹ MELE, *Aspects of Agency*, op. cit., pp. 18-20.

que seria de facto melhor, ir ver o jogo e, com base nessa decisão cognitiva, adquirir a intenção não ativa de ir ver o jogo. Ou talvez eu tivesse outros planos, mas tendo entendido que a proposta era atrativa, revi esses planos. E talvez tenha efetuado uma ação mental de assentimento à proposta e expressado o assentimento interno, com a pessoa que me fez a proposta. Com certeza, ao efetuar essa ação mental, eu decidi ir ver o jogo: o meu assentimento mental à proposta, foi um ato de formação de intenção, ou seja, um ato de resolver ir ver o jogo do Mike esta noite, conforme me propuseram.

A proposta teórica de Alfred Mele acerca de decidir agir, é que se trata de um ato mental de concordância executiva, para um plano de ação da primeira pessoa. Ao decidir agir, o sujeito forma uma intenção para agir, e ao fazê-lo, o sujeito tem presente, que tem uma intenção que incorpora o plano, com o qual concorda.

No caso de uma decisão de não *A*, na qual não fazer *A* é representado pelo agente como um não fazer genuíno, o plano associado com a intenção formada é simplesmente não *A* e essa intenção, não é uma intenção para agir. Intenções nas quais o agente se assegura que não faz *A*, no entanto, são intenções para agir. E uma intenção de não *A* pode ajudar a racionalizar uma intenção do agente de assegurar-se que não faz *A* quando, por exemplo, acredita que estará em perigo de sucumbir à tentação de *A*.

A proposta de Alfred Mele requer que existam planos com os quais o sujeito concorda, ao efetuar uma decisão prática.

Apesar da decisão prática ser uma forma direta de resolver incerteza prática, o facto de um sujeito decidir *A*, pode gerar novas incertezas práticas. Por exemplo, apesar de ter decidido passar férias no Havai daqui a dois anos, estou incerta de como irei chegar lá, onde ficarei, etc. Estes assuntos podem ser estabelecidos por futuras decisões práticas.

Da mesma forma, em algumas circunstâncias, agentes que decidem não fazer *A*, decidem subsequentemente o que fazer em vez disso.

Portanto, se existem ações básicas, as decisões práticas construídas no presente, estão entre elas.

As decisões práticas, portanto, são estruturadas de forma relativamente simples. A sua complexidade reside no conteúdo da intenção formada e não na estrutura do ato de tomar a decisão em si. Decidir que *A*, será mais próximo do facto de alguém levantar o braço,

do que o sujeito indicar, ao levantar o braço, que pretende levantar uma questão. A proposta de que a decisão prática é uma ação mental momentânea de uma concordância executiva de um plano de ação da primeira pessoa, é coerente com esta relativa simplicidade.

Alfred Mele defende que a decisão prática é um ato mental do tipo atrás descrito e que difere da aquisição de uma intenção de forma não ativa, porque esta última não é um ato de nenhum tipo.

Nós tomamos decisões práticas somente quando estamos incertos acerca do que fazer. Este facto é também verdade, normalmente, quando raciocinamos acerca daquilo que seria melhor ou suficientemente bom ou permissível fazer e fazemo-lo em situações que envolvem incerteza prática.

As nossas decisões práticas podem estar de acordo com as decisões cognitivas que alcançamos ou pelo contrário, conflitar com elas, tal como nos casos de ação acrática. Se um agente que julgue melhor *A*, decide *A* ou decide não *A*, a sua decisão prática, resolve a incerteza prática que a sua conclusão cognitiva deixou por resolver.

Alfred Mele sugere que algumas decisões cognitivas resolvem incertezas práticas neste sentido: se um agente está incerto acerca de *A*, a aquisição da crença de que seria melhor *A* pode, de imediato, resultar na aquisição de uma intenção de *A* e ao adquirir essa intenção, o sujeito fica comprometido a fazer *A*, ficando portanto removida a incerteza prática relevante.

Mas o autor sugere também que, algumas incertezas práticas, incluindo algumas incertezas que qualquer decisão cognitiva a que alguém chegue mas falhe em resolver, são resolvidas através das decisões práticas. A decisão prática, é uma forma importante de resolução de incertezas e um modo de ação direto, modo esse, que é uma ação mental momentânea de formação de intenções.

O poder da formação de decisões práticas está no panorama de muita literatura acerca de liberdade de vontade e liberdade de ação, quer literatura compatibilista quer incompatibilista. Alfred Mele sublinha que queremos pensar que, de facto, formamos intenções, que decidimos fazer certas coisas e não outras.

Colocando de lado decisões práticas de não fazer, é razoável entender as decisões práticas como ações mentais de estabelecimento executivo de planos de ação da primeira pessoa¹⁷⁰.

3.1.16.2 Decisões conscientes

Alfred Mele deixa em aberto se todas as decisões são decisões conscientes, ou seja, que algumas ações momentâneas ou estabelecimento executivo de um plano de ação da primeira pessoa, não são formadas conscientemente¹⁷¹.

De novo, decidir que *A* é desempenhar uma ação momentânea de formar uma intenção de *A*. Decidir que *A*, não pode ser confundido com nenhum processo que leva a uma decisão para *A*, incluindo, por exemplo, deliberação acerca do que fazer. E decidir que *A* não precede o início da intenção de *A*, formado no ato de decidir. Antes, aquilo que é decidir que *A* consiste em formar, ativamente, uma intenção para *A*. A intenção emerge na ação formadora da intenção momentânea e não após a mesma.

Alfred Mele sugere que podemos decidir conscientemente que *A*, ou seja, desempenhar conscientemente a ação momentânea de formar uma intenção para *A*, no mesmo sentido de desempenho consciente e, se esta sugestão está correta, o momento do início de uma intenção proximal para *A* pode, pelo menos algumas vezes, ser idêntico ao momento de início da consciência do agente de tal intenção. Estes momentos são idênticos sempre que um agente conscientemente decide que *A*, no sentido da expressão, decisão consciente, ou seja, no sentido que requer que o agente esteja consciente da ação de decidir fazer *A* durante todo o tempo em que desempenha *A*¹⁷².

3.2 O auto engano

Considerando a abordagem das instâncias de agência irracionais, não pode ser ignorado o facto de que a nossa interpretação da realidade, pode ser deturpada pelos nossos afetos, ainda mais se a realidade em questão, afeta os nossos receios ou desejos. Refiro-me aqui a um viés cognitivo, pois, se nossas crenças afetam os nossos juízos

¹⁷⁰ MELE, *Motivation and Agency*, op. cit., pp. 209-213.

¹⁷¹ MELE, *Effective Intentions*, op. cit., p. 16.

¹⁷² MELE, *Effective Intentions*, op. cit., pp. 43-44.

avaliativos, em função dos quais deliberamos, logo, se essas crenças são formadas irracionalmente, ou são deturpadas pelos nossos afetos, muito facilmente poderemos cair em ações consideradas acráticas. E neste caso, o agente acrático não estaria a agir contra o seu melhor juízo, conforme defende Donald Davidson, mas estaria a agir em função de um melhor juízo enviesado.

Pode tornar-se útil a comparação do auto engano com a ação acrática. Quando um agente age acriticamente, ele age intencionalmente, mas o facto de agir incontinentemente, tipicamente, não faz parte da sua intenção. Ou seja, o agente não pretende agir incontinentemente e isso não faz parte do seu plano de ação. Apesar de, contra o seu melhor juízo, o agente intencionalmente comer a sobremesa que está diante dele, o agente não tenciona, nesse momento, agir incontinentemente. Se, similarmente, o agente auto enganado iniciar um comportamento intencional, que o coloque num estado de auto engano, talvez ele o faça sem pretender iludir-se a si próprio.

Em *Irrationality* Alfred Mele defendeu que o auto engano parece ser motivado por desejos ou receios do agente¹⁷³.

Para Alfred Mele, o auto engano ocorre, aparentemente, de duas formas diferentes: *straight* ou *twisted*, a que chamarei de forma direta ou forma indireta.

Nas formas *straight*, de auto engano, os sujeitos iludem-se a si próprios ao acreditar em algo, que querem que seja verdade. Por exemplo, acreditar que não estão seriamente doentes ou que o seu filho não se droga.

Nos casos *twisted*, de auto engano, o agente ilude-se a si próprio ao acreditar em algo que deseja que seja falso e que não quer que seja verdade. Por exemplo, um marido inseguro e ciumento, pode acreditar que a esposa tem um caso, apesar de não ter evidências concretas do facto e apesar de não querer que seja o caso. Então, se o auto engano pode existir nesta forma *twisted*, pelo menos um argumento comum acerca do auto engano é falso – o argumento de que o auto engano de *S* de que *p*, requer o desejo de *S* de que *p*¹⁷⁴.

¹⁷³ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 123-124.

¹⁷⁴ MELE, Alfred, *Self-Deception Unmasked*, Princeton University Press, New Jersey 2001, pp. 4-5.

3.2.1 Crenças e agência motivacionalmente enviesadas

Alfred Mele refere que podemos ter a tendência para acreditar em proposições que queremos que sejam verdadeiras, mesmo quando uma investigação imparcial indica que essa proposição é provavelmente falsa. Uma hipótese plausível para esta tendência, será que o nosso desejo de que algo seja verdadeiro, por vezes, exerce uma influência enviesadora naquilo em que acreditamos. E existe evidência de que as nossas crenças acerca dos nossos próprios traços se tornam mais enviesadas, quando esses traços são altamente desejáveis ou indesejáveis¹⁷⁵.

Alfred Mele apresenta as duas grandes teses acerca das crenças motivacionalmente enviesadas:

T1 – A visão da agência: todas as crenças motivacionalmente enviesadas são produzidas ou protegidas intencionalmente. Em todos os casos de crenças motivacionalmente enviesadas que *p*, tentamos ter presente que adquirimos ou retemos a crença que *p*, ou pelo menos, tentamos facilitar a nós próprios o facto de adquirir ou reter a crença.

T2 – A visão anti agência: nenhuma crença motivacionalmente enviesada é produzida ou protegida intencionalmente. Em nenhum caso de crença motivacionalmente enviesada que *p*, o sujeito procura ter presente que adquire ou retém a crença, ou tenta facilitar a si próprio o facto de adquirir ou reter a crença.

As tentativas a que apela a visão da agência parecem enfraquecê-la. Se estou a tentar ter presente que sou uma boa condutora (não pelo aperfeiçoamento das competências de condução, mas ignorando alguma evidência de que sou uma condutora inferior) não chegarei à conclusão de que os fundamentos para essa crença são ilegítimos? E não continuarei a não ter crença de que sou boa condutora?

Na visão anti agência, encontramos dificuldades em tentar compreender de que forma, crenças motivacionalmente enviesadas podem emergir, sem ser através de esforços do género postulado pela visão da agência. Por exemplo, como pode a minha vontade de querer ser boa condutora motivar-me a acreditar que sou uma boa condutora, exceto se me motivar a tentar ter presente que acredito nisso ou se me motivar a tentar tornar mais fácil, para mim própria, acreditar nisso? Esta perspetiva tem o desafio de obter uma visão

¹⁷⁵ MELE, *Self-Deception Unmasked*, op. cit., p. 11.

alternativa do mecanismo pelo qual os desejos conduzem a crenças motivacionalmente enviesadas¹⁷⁶.

Alfred Mele distingue três tipos de atividades cognitivas, que contribuem para as crenças motivacionalmente enviesadas: 1. atividades não intencionais (ou seja, foco não intencional em informação de certo tipo); 2. atividades intencionais (ou seja, foco intencional em informação de certo tipo) e 3. atividades intencionais envolvidas, como fazendo parte de uma tentativa de iludir-se a si próprio ou causar a si próprio uma crença ou tornar mais fácil a si próprio acreditar em algo (foco intencional em informação de certo tipo, como fazendo parte de uma tentativa de se iludir a si próprio ao acreditar que *p*)¹⁷⁷. As atividades cognitivas intencionais que contribuem, mesmo que de uma forma relativamente direta, para uma crença motivacionalmente enviesada, falsa e injustificada, não têm que ser guiadas por uma intenção, nem têm que envolver tentativas associadas para manipular aquilo em que se acredita^{178, 179}.

Para Alfred Mele, existe larga evidência de que desejar que *p* fornece uma contribuição causal importante na aquisição e retenção de crenças injustificadas de que *p*. E estes desejos tornam-se explicações causais de crenças enviesadas pertinentes. É um erro assumir que o papel característico dos desejos na explicação das ações intencionais é o único papel explanatório que os desejos podem ter. Os desejos contribuem para a produção de crenças motivacionalmente enviesadas (incluindo crenças que o sujeito está iludido, ao tê-las) numa variedade de formas, relativamente bem entendida, que encaixa no modelo anti agência¹⁸⁰.

¹⁷⁶ MELE, *Self-Deception Unmasked*, op. cit., p. 13.

¹⁷⁷ Alfred Mele apresenta o exemplo de Beth, que perdeu o Pai há pouco tempo e encontra algum conforto emocional ao rever memórias nas quais brincavam os dois. Ao mesmo tempo, sente desconforto em visualizar memórias em que o seu pai brincava mais com os seus irmãos e deixava Beth para trás. Se Beth intencionalmente prestar atenção às memórias agradáveis, pode adquirir uma falsa crença de que o seu pai gostava mais dela do que qualquer pessoa. Mas, apesar destas atividades cognitivas intencionais poderem ser explicadas pela motivação atrativa de que o seu Pai gostava mais dela, essas atividades não têm que ser também explicadas por um desejo ou uma tentativa de se iludir a si própria, ao acreditar nisso, ou causar a si própria o facto de acreditar nisso, ou tornar mais fácil para si própria acreditar nisso.

¹⁷⁸ O caso de Beth pode ser explicado pela hipótese de que procura experiências agradáveis e tenta evitar as desagradáveis, sem que em nenhum momento tente influenciar aquilo em que acredita.

¹⁷⁹ MELE, *Self-Deception Unmasked*, op. cit., pp. 17-19.

¹⁸⁰ MELE, *Self-Deception Unmasked*, op. cit., p. 24.

3.2.2 Processos Psicológicos envolvidos na *Straight Self Deception*

3.2.2.1 O paradoxo

O paradoxo estratégico de Alfred Mele pretende resolver a ideia Freudiana de que os seres humanos são parcialmente compostos por subagentes, capazes de comportamento intencional, ideia com a qual o autor não concorda. Segundo esta visão, existiria um subsistema mental (o enganador) que opera num outro subsistema mental (o enganado) sem que este saiba, ou sequer suspeite, daquilo que o enganador está a fazer.

Ora, segundo Alfred Mele esta situação não configura o auto engano, mas sim o engano de uma entidade por outra. O paradoxo estratégico não envolve a existência de entidades desta natureza.

Com efeito, para Alfred Mele quando o sujeito de auto engana, assume estratégias, tais como interpretação errónea positiva ou negativa, foco seletivo e recolha de evidência seletiva. E este tipo de comportamentos são estratégias de auto engano, no sentido lato. Não é possível compreender o auto engano, sem prestar atenção à influência da motivação na formação e retenção de crenças¹⁸¹.

As estratégias de auto engano dividem-se em dois tipos: um que se constitui pela manipulação de informação que o sujeito já tem - *internal-biasing*, e outro que consiste no controlo do agente (em determinado grau) na informação que adquire – *input-control*. A interpretação errónea positiva ou negativa e o foco seletivo, são estratégias do primeiro tipo, quando o agente foca seletivamente ou presta atenção a informação que já tem.

Quando o agente procura informação que corrobore determinada hipótese, estamos a falar de estratégias do segundo tipo, considerando que a sua atividade atencional é uma forma de *input-control*¹⁸².

As estratégias de *input-control* dividem-se em dois grupos: o que opera em informação que já existe no mundo e outro que gera informação.

O primeiro grupo é similar à estratégia de *internal-biasing*.

¹⁸¹ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 138-139.

¹⁸² MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 143.

Uma das formas de exercer controlo sobre *data-input* é gerar informação para nós próprios e uma estratégia de criação dessa informação consiste, no que Alfred Mele chama de *acting as if*, que veremos mais adiante¹⁸³

3.2.2.2 Desejos e enviesamentos

Alfred Mele descreve a variedade de processos, que contribuem para as crenças motivacionalmente enviesadas.

Nos casos de *straight self deception*, o nosso desejo que *p*, contribui, através de uma variedade de formas, para a nossa crença de que *p*, conforme comprovam quatro exemplos que o autor descreve.

1 – Interpretação equivocada negativa – *Negative misinterpretation* – o nosso desejo de *p*, pode levar-nos a interpretar erroneamente, ou não valorizar o suficiente informação contra *p*, informação essa que facilmente reconheceríamos como válida ou suficientemente válida, na ausência do desejo¹⁸⁴.

2 – Interpretação equivocada positiva – *Positive misinterpretation* – o nosso desejo de *p* pode levar-nos a interpretar informação como sendo a favor de *p*, informação essa que facilmente reconheceríamos como sendo contra *p*, na ausência do desejo¹⁸⁵.

3 – Atenção Seletiva – *Selective focusig/attending* – o nosso desejo de *p* pode levar-nos a não prestar atenção em evidência contra *p* e colocar o nosso foco em evidência sugestiva de *p*¹⁸⁶.

4 – Recolha seletiva de evidência – *Selective evidence-gathering* – o nosso desejo de *p* pode levar-nos a ignorar evidência obtida facilmente para $\sim p$ e procurar evidência de *p*,

¹⁸³ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 143-144.

¹⁸⁴ Por exemplo, Don entende que a rejeição de um artigo que havia submetido foi erroneamente rejeitada, mas alguns dias mais tarde, quando relê o artigo e as críticas, num estado de espírito mais imparcial, fica claro para ele que a rejeição foi, de facto, justificada.

¹⁸⁵ Por exemplo, Sid nutre um sentimento pela colega Roz, com quem estuda frequentemente e, desejando que a colega também sinta o mesmo, interpreta a sua recusa em sair com ele, bem como o facto de o ter lembrado que tem namorado, como estando a fazer-se de difícil, para que Sid se sinta encorajado a continuar a aproximar-se dela.

¹⁸⁶ Alfred Mele convida-nos a recordar o exemplo de Beth, que desejava ter sido a favorita do seu Pai e portanto, procurava conforto em memórias que a punham em destaque com ele ao mesmo tempo que negligenciava as memórias que a colocavam em segundo plano.

mesmo que seja muito menos acessível¹⁸⁷. A recolha de evidência seletiva pode ser analisada como sendo uma hipersensibilidade a evidência e fontes de evidência favoráveis ao que se deseja, combinada com uma cegueira por evidência e fontes de evidência contrárias¹⁸⁸.

Em cada um dos casos enunciados, existe um tratamento inapropriado de informação, influenciado pelo desejo. O desejo do sujeito de que p , leva-o a falhar a apreciação da relevância de evidência de que $\sim p$, de forma a interpretar erroneamente tal evidência, focar-se seletivamente em evidência que suporta p , ou a falhar a localização de evidência disponível de que $\sim p$, enquanto procura evidência de que p . Se esta manipulação da informação leva à aquisição de uma falsa crença de que p , então podemos estar perante um exemplo de auto engano.

Para Alfred Mele, o que gera o auto engano do sujeito na crença de que p , é uma manipulação de informação influenciada pelo desejo, que é ou parece ser relevante para a verdade de p . Parte daquilo que o sujeito auto enganado faz, em muitos casos, é prevenir-se de deter determinada crença verdadeira e, é por essa razão, que o sujeito não acredita que $\sim p$ enquanto acredita que p . A interpretação errónea – *misinterpretation* – pode ser, por vezes, explicada pela suposição simples de que o agente reconhece ou acredita que a informação apenas pode ser tomada em conta como sendo contra p , ou, considerando a sua condição emocional, o agente pode simplesmente ver a informação de determinada forma¹⁸⁹.

Em nenhum destes casos, o sujeito detém a crença verdadeira de que $\sim p$ e, intencionalmente, sustém a consciência de que acredita em p .

Através destes casos, ficamos com uma intuição acerca da forma como o desejo de que p , pode iniciar ou suster cada um dos quatro processos.

Por outro lado, reconhecemos que as pessoas tendem a procurar prazer e evitar a dor e algumas atividades atencionais são prazerosas ou dolorosas. Além disso, ficamos também

¹⁸⁷ Por exemplo, Betty que faz parte do *staff* de uma campanha política, ouve rumores de que o líder é sexista, facto que ela tem esperança de que não seja verdade. Essa esperança leva-a a procurar informação acerca do comportamento do líder, mas junto de pessoas que trabalham na própria campanha e vasculha o seu passado político. Betty pode ter ignorado informação óbvia de que o líder é de facto sexista e centrou-se em informação menos óbvia que favoreceu a sua visão.

¹⁸⁸ MELE, *Self-Deception Unmasked*, op. cit., pp. 25-27.

¹⁸⁹ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 126-129.

com a intuição de como a atenção seletiva para evidência de p , pode aumentar a probabilidade de adquirirmos a crença de que p .

O facto de sermos auto enganados ao adquirir ou manter uma crença, constitui uma crença enviesada. E, no auto engano, o enviesamento é motivado, mas Alfred Mele considera importante relevar três potenciais fontes de crenças enviesadas não motivadas (ou frias – *cold*), que decorrem da literatura psicológica:

1 – Vividez da informação - *Vividness of information* – Existe uma maior probabilidade de reconhecer, prestar atenção e recordar informação vívida do que informação mais fraca ou opaca e, conseqüentemente, a informação vívida tende a possuir uma influência desproporcional na formação e retenção de crenças.

2 – Heurística disponível – *The availability heuristic* – Quando formamos crenças acerca da frequência, probabilidade, ou causas de um evento, podemos ser influenciados pela relativa disponibilidade dos objetos ou eventos, ou seja, a sua acessibilidade nos processos de percepção, memória ou construção imaginativa. Da mesma forma, as tentativas de localizar a causa de um evento, são significativamente influenciadas por manipulações que focam a nossa atenção numa potencial causa específica.

3 – A confirmação enviesada – *The confirmation bias* – Quando as pessoas testam uma hipótese, tendem a procurar, na maior parte das vezes, na sua memória e no mundo, a sua confirmação em vez da não confirmação, ao mesmo tempo que reconhecem mais rapidamente a confirmação do que a não confirmação¹⁹⁰. As implicações do viés da confirmação na formação e retenção de crenças são óbvias.

Obviamente, a informação mais vívida e disponível tem, por vezes, um importante valor probatório, ou de evidência, e a influência de tal informação nem sempre é uma influência enviesadora.

Portanto, é importante destacar que, apesar das fontes de crenças enviesadas poderem funcionar independentemente da motivação, podem também ser desencadeadas e sustentadas pela motivação, na produção de crenças motivacionalmente enviesadas particulares. Por exemplo, a motivação pode aumentar a vividez ou saliência de

¹⁹⁰ Por exemplo, sujeitos que testaram a hipótese de que uma pessoa estava zangada, interpretaram a expressão facial, como transmitindo fúria, enquanto que, sujeitos que testaram a hipótese de que de que a pessoa estava feliz, interpretaram a mesma expressão facial como transmitindo felicidade.

determinada informação. Informação que conta a favor da verdade de uma hipótese, que o sujeito gostaria que fosse verdadeira, pode tornar-se mais vívida ou saliente, e informação vívida e saliente, mais propensa ao reconhecimento e evocação, tende a ser mais disponível do que outras contrapartidas mais ténues. Similarmente, a motivação pode influenciar as hipóteses que vão ocorrer ao sujeito, e afetar a saliência de hipóteses disponíveis, configurando assim o contexto para a confirmação enviesada. Porque as hipóteses favoráveis, são mais agradáveis de contemplar do que as desfavoráveis e tendem a estar mais rapidamente disponíveis na nossa mente. Desejar que p , aumenta a probabilidade de que o sujeito se foque mais em p do que em $\sim p$, ao testar a sua hipótese. A motivação que desencadeia e sustém a confirmação enviesada – *confirmation bias* – promove o comportamento cognitivo que os epistémicos rejeitam. Pelo menos, em alguns exemplos de auto engano, as falsas crenças adquiridas face a evidência em contrário, podem ser produzidas e sustentadas pelo fenómeno da motivação tal como descrito. E, nestes casos, o auto engano não requer que o agente pretenda ou tente produzir ou sustentar uma crença em si próprio, ou comece a acreditar em algo que vai acabar por desacreditar.

Obviamente, o enviesamento frio não é intencional e, em casos particulares, o funcionamento dos mecanismos descritos podem ser iniciados ou sustentados pela motivação, independentemente de qualquer intenção ou tentativa de ilusão.

Alfred Mele expõe uma breve explicação da sustentação motivacional do funcionamento destes mecanismos.

Um desejo passageiro que, momentaneamente, aumente a vividez ou disponibilidade de uma informação, ou sugira uma hipótese momentânea, não terá, permanecendo iguais todas as outras coisas, um efeito forte na aquisição de crenças, como terá um desejo de duração significativamente maior. Retomando o exemplo de Sid, que deseja que Roz esteja apaixonada por ele, pode interpretar evidência desta atração como sendo particularmente vívida e disponível, enquanto o desejo persistir. Ao longo do tempo e devido a estas características da evidência, ele pode acreditar que ela gosta dele. Mas, quando perder o desejo relevante (talvez devido a apaixonar-se por outra pessoa), é expectável um decréscimo na vividez e disponibilidade da evidência em questão acerca

da Roz, e, permanecendo iguais todas as outras coisas, um decréscimo correspondente na probabilidade de que adquire a crença em questão.

Portanto, não é misteriosa a forma como a confirmação enviesada, a disponibilidade heurística e a vividez de informação, podem contribuir para os dois tipos de interpretação equivocada e os dois tipos de seletividade identificados por Alfred Mele. Os efeitos previsíveis da motivação na confirmação enviesada – *confirmation bias*- aumentam a probabilidade de interpretação enviesada, foco seletivo e recolha de evidência seletiva. O mesmo será verdade para os efeitos da motivação na disponibilidade e vividez de informação¹⁹¹.

3.2.2.3 *Acting as if*

Conforme referido, uma estratégia de criação de informação para nós próprios (controlo sobre *data-input*), é o *acting as if*.

O *acting as if* permite ao agente gerar informação que suporta *p*, de forma direta ou intrapessoal ou social. Por vezes, fazemos inferências acerca de nós próprios, tendo por base a observação do nosso próprio comportamento e, agindo como se *p* fosse o caso, o sujeito pode adicionar informação à evidência de base.

Além disso, *acting as if p*, o agente pode influenciar a perceção dos outros ao gerar para si próprio, informação social que suporta *p*^{192, 193}.

Alfred Mele sugere que esta tendência fria, pode ser explorada pela motivação, em casos de auto engano. Apesar de um agente, conscientemente, acreditar que está *acting as if p*, de forma a proporcionar, a si próprio, evidência de que *p*, tenderia a minar esse projeto, pois o agente pode não saber o que o motiva para *act as if*. E esta é a maior dificuldade em perceber de que forma o *acting as if* pode levar ao auto engano.

Todos sabemos que existe a tendência de formarmos crenças, incluindo falsas crenças, acerca de nós próprios, ao observar o nosso próprio comportamento. E também não é difícil ver porque alguém pode estar motivado para *act as if p* por um desejo de que *p*

¹⁹¹ MELE, *Self-Deception Unmasked*, op. cit., pp. 27-31.

¹⁹² Por exemplo, o agente que, consistentemente, se comporta de forma confiante em público, é percebido e tratado como um indivíduo confiante pelos outros sujeitos que observam o seu comportamento, mesmo que esse indivíduo não seja de todo confiante.

¹⁹³ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 151.

(por exemplo, uma pessoa insatisfeita com uma situação que entende ser difícil corrigir, tal como uma fase de infelicidade no casamento, pode *act as if* estivesse feliz, de forma a tornar tolerável a situação ou tornar mais fácil o adiamento de uma decisão definitiva¹⁹⁴.

Parte da atração pela noção de subagentes deve-se ao facto das atividades que levam ao auto engano e à ação acrática, serem explicáveis num modelo tradicional de crença-desejo da explicação da ação. Porque, mesmo que a pessoa acredite ou aja irracionalmente, a etiologia da crença ou ação, envolve centralmente o trabalho de um subagente que se comporta racionalmente à luz das suas crenças e desejos.

Mas, Alfred Mele entende que o modelo psicológico, que presta um papel explicativo importante às crenças e desejos, pode levar a explicações adequadas das ações irracionais, sem assentar na hipótese de uma subagência intencional.

As estratégias apresentadas por Alfred Mele pretenderam mostrar que o auto engano não depende de uma execução intencional ou de uma monitorização de um subagente¹⁹⁵

3.2.3 Condições suficientes para o auto engano, na aquisição de crenças

Alfred Mele identifica algumas condições que, em conjunto, considera suficientes para o auto engano, na aquisição de uma crença de que p :

- 1 – A crença de que p que o sujeito S adquire é falsa;
- 2 – O sujeito S trata informação relevante, ou pelo menos, aparentemente relevante, para a verdade de p de uma forma motivacionalmente enviesada;
- 3 – O tratamento enviesado é uma causa não desviante da aquisição da crença de que p , pelo sujeito S ;
- 4 – O conjunto da informação que o sujeito S possui no momento, proporciona maior garantia de que $\sim p$ do que de p .

A condição 4 pretende evidenciar que a crença auto enganadora do sujeito existe, perante o peso da evidência que possui e Alfred Mele não considera esta condição como necessária do auto engano. Em alguns casos, de recolha de evidência motivacionalmente enviesada, por exemplo, o sujeito pode tomar consciência da sua crença na falsidade de

¹⁹⁴ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 157.

¹⁹⁵ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., p. 159.

p , quando $\sim p$ é melhor suportado pela evidência disponível e, mesmo assim, devido ao processo de seletividade de recolha de evidência, a evidência que realmente possui, na altura, favorece p em vez de $\sim p$.

Alfred Mele refere que alguns Filósofos discordam destas quatro condições, que consideram insuficientes, pelo facto de não acomodarem a tensão necessária ao auto engano, tensão essa, entendida como sendo representada com o reconhecimento de que p , coexistente com o conhecimento ou, pelo menos, a crença verdadeira de que $\sim p$. Esta tensão consistiria, portanto, num conflito cognitivo.

Mas Alfred Mele não nega a existência de conflito nos casos de auto engano. Nega sim, que a alegada tensão seja conceptualmente necessária para o auto engano, na aquisição de uma crença de que p . Com efeito, a evidência pode favorecer a condição indesejada pelo sujeito, mas não é necessário supor que o agente está consciente disso, para explicar a sua cognição enviesada¹⁹⁶.

A conceção de auto engano de Alfred Mele pode ser contrastada com o modelo intencional de auto engano. Defende que não existe razão para supor que este fenómeno pressupõem que o agente esteja na condição doxástica particular de acreditar que p e acreditar também que $\sim p$ ¹⁹⁷.

3.2.4 Auto engano na retenção de crenças

As condições mencionadas referem-se especificamente ao auto engano na aquisição de crenças, mas as conceções deste fenómeno envolvem o auto engano de um sujeito, ao reter uma crença.

Alfred Mele ilustra este facto com o exemplo de Sam, que acreditou por muitos anos, que a sua mulher nunca teria um caso com outra pessoa. No passado, esta proposição poderia ser considerada como boa, considerando as evidências: o comportamento de Sally era revelador de ser feliz e adorar o marido e a sua vida. Mas, recentemente, este comportamento alterou-se e Sally chega frequentemente tarde a casa, dá desculpas para sair sempre sozinha e um amigo próximo do casal afirmou a Sam, ter visto a sua esposa

¹⁹⁶ MELE, *Self-Deception Unmasked*, op. cit., pp. 50-53.

¹⁹⁷ MELE, *Irrationality. An essay on akrasia, self-deception and self-control*, op. cit., pp. 132-133.

na companhia de Mr. Jones. Mesmo assim, Sam continua a acreditar que Sally não tem nenhum caso e, no entanto, Sam está errado.

Neste caso, Sam incorre no auto engano, mas ao reter uma crença.

Em geral, quanto mais forte for a evidência percebida contra a proposição (ou crença) em que se acredita, mais difícil será reter a crença. Mas, no caso de Sam, ele pode interpretar a informação erroneamente, de forma positiva, pensando que, se Sally realmente tivesse um caso, teria o cuidado de se esconder em público com Mr. Jones e, consequentemente, isso indica que eles não estão, de facto, envolvidos. Por outro lado, pode interpretar erroneamente a informação de forma negativa, ao pedir-lhe uma explicação ou sugerindo, para aprovação da sua mulher, algumas hipóteses aceitáveis acerca da sua conduta. O foco seletivo pode ter um papel óbvio e mesmo a recolha de informação seletiva pode ter uma influência no auto engano de Sam. Poderá até procurar investigar por conta própria, mas devido ao seu desejo de Sally não ter um caso, localizará informação menos acessível favorável ao estado desejado, e ignorará informação importante e disponível que suporte o juízo contrário.

De novo, o auto engano comum é explicável, independentemente de assumirmos que o sujeito manipula a informação para se iludir a si próprio, ou como um esforço para proteger ou produzir a crença favorável.

Por outro lado, não existe necessidade de supor que Sam acredita em p , ao mesmo tempo que acredita em $\sim p$ ¹⁹⁸.

3.2.5. *Wishful Thinking*

A literatura acerca do auto engano, inclui a tentativa de distinção entre este fenómeno e um outro fenómeno relacionado: o pensamento ilusório – *wishful thinking*. Alfred Mele recorre à sugestão de Bela Szabados relativamente à distinção entre os dois termos: tanto o *self deceiver* como o *wishful thinker* detêm a falsa crença e ambos estão influenciados por um desejo ou motivo de algum tipo, mas enquanto o *self deceiver* cumpre uma das condições para o conhecimento (esse agente tem boas razões para pensar que a proposição em que acredita, é falsa) o *wishful thinker* não possui essas razões.

¹⁹⁸ MELE, *Self-Deception Unmasked*, op. cit., pp. 56-58.

Considerando que Alfred Mele aceita a proposição de que aquilo que o *wishful thinker* faz, é acreditar falsamente que *p*, então, se houver alguma distinção entre os dois fenómenos, pode simplesmente dever-se a que o *wishful thinking* é uma espécie, do género abrangido pelo termo *self deception*. Se Bela Szabados está correta acerca da evidência contrária, pode apenas considerar-se que o *wishful thinking* seja uma forma de *self deception*, na qual, devido a algum comportamento influenciado pelo desejo, o *self deceiver* não possui boas razões para rejeitar a proposição na qual, auto enganadamente, acredita.

É verdade que o termo *wishful thinking* evoca a imagem do idealismo da juventude. Mas suponhamos que em vez de *wishful thinking* nos referíamos a *wishful false believing*, termo melhor apropriado, se considerarmos correta a explicação de Bela Szabados acerca do pensamento envolvido neste fenómeno.

Alguns casos de auto engano são também casos de *strict akratic believing*. Outros envolvem acrasia. Muitos casos que se situam na fronteira do auto engano e *strict akratic belief* são casos fortes de auto engano. A pessoa culpada de uma *strict incontinent belief*, ao contrário da maioria dos agentes auto enganados, acredita que a evidência constitui razões boas e suficientes para que não acredite que *p*, a proposição em que acredita auto enganadamente¹⁹⁹.

Os defensores da teoria de que o auto engano é essencialmente, ou normalmente, intencional poderão suportar a distinção entre auto engano e pensamento ilusório – *wishful thinking*, defendendo que apesar do pensamento ilusório não precisar de uma intenção para iludir o sujeito, o auto engano pode distinguir-se, precisamente, devido ao facto de ser intencional.

Se o pensamento ilusório não é uma crença ilusória, então é óbvia uma diferença entre o pensamento ilusório de que *p* e estar auto enganado ao acreditar que *p*. Se, no entanto, o pensamento ilusório é a crença ilusória, em particular, uma falsa crença motivacionalmente enviesada, então se não se sobrepõe com o auto engano, a diferença pode residir na força relativa da evidência relevante contra a proposição em que se acredita: o sujeito que pensa ilusoriamente - *wishful thinker*, pode deparar-se contra

¹⁹⁹ MELE, *Irrationality. An Essay on Akrasia, Self-Deception and Self-Control*, op. cit., pp. 134-136.

provas mais fracas do que o sujeito auto enganado. Esta diferença requer uma diferença na intenção, apenas se a força relativa da evidência contra a proposição na qual o sujeito auto enganado acredita for tal, que requer que a aquisição ou retenção dessa crença, dependa da intervenção de uma intenção pertinente e, para Alfred Mele, a tese relativa à força das evidências, é falsa²⁰⁰.

Jon Elster discorda desta interpretação de Alfred Mele, defendendo uma perspectiva oposta. Para Elster, o *Wishful Thinking* é a tendência para a formação de crenças quando e porque o sujeito prefere um estado do mundo, no qual são verdadeiras do que outros estados, nos quais são falsas.

A decisão de acreditar é moldada por um desejo consciente e o *wishful thinking* é moldado por uma ânsia não consciente. Por outro lado o *wishful thinking* é distinguível do auto engano, uma vez que o auto engano envolve paradoxos que não estão presentes no *wishful thinking*. E a raiz deste paradoxo é o fato particular do agente que se auto engana, esconder intencionalmente uma das suas crenças, dele próprio, enquanto defende o outro, como sendo a sua posição oficial. O seu argumento nesta distinção é que, no *wishful thinking*, o sujeito pode chegar às suas crenças que são verdade, mas são também bem fundamentadas, na medida em que correspondem a evidência disponível. O *wishful thinking* não pode ser causalmente fundamentado em evidência, mas pode ser sustentado pela mesma. Por contraste, o auto engano envolve necessariamente uma dualidade entre crença que o sujeito detém, oficialmente, e a que o sujeito acredita estar sustentado pela evidência²⁰¹.

3.2.6 *Twisted Self Deception*

O trabalho Filosófico e Psicológico, tem vindo a focar-se naquilo a que Alfred Mele chama de *Straight Self-Deception*, noção segundo a qual, conforme vimos, o agente auto engana-se, ao acreditar em algo que deseja que seja verdade.

Mas existe uma outra noção, menos comum, e teoricamente mais complexa que é a forma *twisted* de auto engano.

²⁰⁰ MELE, *Self-Deception Unmasked*, op. cit., pp. 73-74.

²⁰¹ ELSTER, Jon. *Sour Grapes. Studies in the Subversion of Rationality*. Cambridge University Press, United Kingdom 1983, pp. 148-151.

Esta forma de auto engano, foi ilustrada com o exemplo do marido inseguro e ciumento, que acredita que a esposa tem um caso, apesar de não ter evidências concretas desse facto e apesar de não querer que isso seja verdade.

Alfred Mele desenvolve duas abordagens para explicar a *twisted self-deception*: uma abordagem centrada na motivação e uma abordagem híbrida, que inclui motivação e emoção.

Estas abordagens permitem explorar e explicar a *twisted self-deception*, sendo consistentes com a posição de Alfred Mele em relação à *straight self-deception*²⁰².

3.2.6.1 A abordagem da *twisted self-deception* centrada na Motivação

Existem alguns modelos de testes de hipóteses quotidianas, desenhados para acomodar evidência de crenças motivacionalmente enviesadas. Alfred Mele recorre ao modelo *Friedrich-Trope-Liberman* (FTL), visão segundo a qual, os testadores de hipóteses procedem da forma que procedem, em parte, porque pretendem evitar ou minimizar erros graves²⁰³. Alfred Mele encontra implícito neste modelo, um diagnóstico motivacional daquilo que acontece na *twisted self-deception*, sem se tornar necessário aceitar o modelo na sua generalidade.

No modelo FTL, o custo de um erro para um indivíduo, depende dos seus interesses e desejos. Enquanto, para algumas pessoas, poderá ser mais importante evitar adquirir a falsa crença de que a sua esposa está a ter um caso, do que evitar adquirir a falsa crença de que os dois não estão tão envolvidos, o inverso pode muito bem ser verdade para pessoas inseguras e ciumentas. A crença de que a esposa é infiel, causa um significativo desconforto psicológico. Mesmo assim, evitar acreditar falsamente que a esposa é fiel pode ser tão importante para algumas pessoas, que essas pessoas testam as hipóteses relevantes, de formas menos propensas a levá-los a uma falsa crença na fidelidade da esposa, do que uma crença falsa na infidelidade da esposa. Além disso, informação sugestiva de infidelidade, pode ser particularmente saliente para essas pessoas, enquanto que, em comparação, informação contrária pode ser bastante fraca.

²⁰² MELE, *Self-Deception Unmasked*, op. cit., p. 94.

²⁰³ MELE, *Self-Deception Unmasked*, op. cit., p. 31.

Mas, porque razão pode alguém estar nestas condições psicológicas?

É certamente aceitável que, perante determinado perfil psicológico, um forte desejo de manter uma relação, influencia o facto de cometer um potencial erro grave, ou falsamente acreditar que o cônjuge é inocente de infidelidade, erro grave esse, no sentido FTL, e mais grave do que o erro de acreditar falsamente que o cônjuge é culpado de infidelidade. Porque afinal, o primeiro erro reduz a possibilidade do marido fazer alguma coisa, para proteger a sua relação em relação a intrusos.

A análise FTL sugere uma visão unificada do auto engano, especificamente, a visão de que, em todos os casos de auto engano (*straight* ou *twisted*), a tendência para minimizar erros graves, dado o perfil motivacional do agente, tem um papel explicativo central. O modelo FTL suporta a visão de Alfred Mele relativamente aos casos de *Straight self-deception*.

Se o fenómeno central ao modelo FTL se aplica a todos os casos de auto engano, irá depender de várias coisas, entre as quais, a força da evidência para a verdade da hipótese FTL e ainda das vantagens de abordagens alternativas na explicação do auto engano.

É uma dessas abordagens alternativa que veremos a seguir²⁰⁴.

3.2.6.2. Abordagem híbrida da *twisted self-deception*, que inclui motivação e emoção

Nos casos comuns de *straight self-deception*, pelo menos parte daquilo que acontece é que, um desejo de que *p* torna mais fácil para o agente acreditar que *p* ao levar o agente a um incremento da vividez de evidência para *p*, a valorizar a saliência de várias memórias que suportam *p*, ao prontificar o agente para que considere hipóteses que suportam *p*, etc. Por vezes, estes efeitos de um desejo de que *p* contribuem para a aquisição ou retenção de uma falsa crença injustificada de que *p*.

Se os desejos fazem isto, então talvez as emoções também o façam.

Retomando o exemplo do marido ciumento e na ausência de desejo, relativamente à infidelidade da sua esposa, bem como ausência de qualquer desejo de acreditar que ela é infiel, pode este homem inseguro, tornar a evidência aparente de que a esposa está a ter um caso, significativamente mais vívida, ao mesmo tempo que torna menos clara

²⁰⁴ MELE, *Self-Deception Unmasked*, op. cit., pp. 96-98.

evidência competitiva? Pode este homem focar a sua atenção em memórias raras de situações de sedução da sua esposa e ignorar memórias que pareceriam mais reveladoras a um agente imparcial? Poderá encetar a situação favorável ao *confirmation bias* que confirma a sua crença de que ela é infiel?

Existem evidências empíricas de que as emoções podem operar destas formas.

Talvez as emoções desempenhem um papel na *twisted self deception* análogo a alguns papéis desempenhados pelos desejos na *straight self deception*²⁰⁵.

Para Alfred Mele, mesmo que alguns casos de *twisted self deception* possam ser explicados, em parte, pela motivação, tal como no modelo FTL, outros casos podem ser explicados pela emoção²⁰⁶.

E é a partir daqui que Alfred Mele desenvolve uma versão híbrida (motivacional/emocional) da *twisted self deception*.

No modelo FTL, os sujeitos tendem a testar as hipóteses, de forma a minimizar erros graves e quais os erros que são considerados graves, pode resultar de desejos e emoções relevantes. O desejo de que p , considerando determinado perfil psicológico, pode contribuir para que o agente tenha motivação mais forte para evitar acreditar, falsamente, que $\sim p$, do que para evitar acreditar, falsamente, de que p , facto que pode contribuir para a *straight self deception* acerca de p . E não existe razão para negar que uma emoção, tal como ansiedade relativamente a p , pode influenciar este cenário.

Nos casos de *twisted self deception*, ao acreditar que p , no modelo FTL, a identificação de quais os erros considerados mais graves para o agente, dependerá de desejos, mas não do desejo de que p .

Para ilustrar esta situação, Alfred Mele refere o exemplo de Dolores, uma mãe que se culpa a ela própria pela morte da sua filha, devido a uma leucemia, sendo pessoalmente inaceitável para Dolores, que a filha morreu de forma aleatória. Dolores insiste em acreditar que a sua filha foi vítima de leucemia, devido ao contacto com um animal de estimação doente, e que Dolores acredita poderia ter retirado de casa e, desta forma, ter evitado a doença da filha.

²⁰⁵ MELE, *Self-Deception Unmasked*, op. cit., pp. 98-99.

²⁰⁶ MELE, *Self-Deception Unmasked*, op. cit., p. 104.

Este caso, envolve um desejo para explicar a morte da sua filha, que enfraqueça a hipótese pessoal, inaceitável, de que a criança ficou doente aleatoriamente. A motivação de Dolores para evitar acreditar falsamente que a morte da sua filha foi aleatória, pode ter sido mais forte, do que a sua motivação para evitar acreditar falsamente que ela própria (Dolores) foi responsável, em parte, pela morte da filha.

Esta situação é inteiramente consistente com a ideia de que as emoções influenciam o enviesamento. Talvez não fosse apenas o desejo, por si só, que influenciou o pensamento de Dolores, mas o desejo em conjunto com uma aversão à ansiedade que Dolores experiencia, perante a hipótese da doença aleatória.

Segundo a hipótese *E*, que Alfred Mele defende, as emoções desempenham um papel na *twisted self deception*, em analogia a alguns dos papéis desempenhados pelos desejos na *straight self deception*; além disso, a *twisted self deception* não tem desejos como causas significantes. Ou seja, de acordo com esta hipótese, as emoções têm um efeito enviesador na cognição, que está em paralelo com alguns efeitos enviesadores da motivação. Além disso a hipótese *E*, nega que a motivação tem um papel significativo na *twisted self deception*.

Nesta linha, Alfred Mele, menciona dois psicólogos que sugeriram que o complexo do ciúme (pensamentos, sentimentos e comportamentos, tipicamente associados a episódios de ciúme), é a manifestação de motivos que refletem preocupações sexuais e de apego. O ciúme está intimamente ligado aos desejos (motivação) das pessoas ciumentas, nas suas relações com a pessoa da qual têm ciúme.

Com efeito, é plausível considerar que o ciúme romântico, é parcialmente constituído por um desejo de um vínculo romântico estreito com a pessoa relevante, ou, se isso falhar, que esse desejo é uma parte significativa da causa do ciúme.

Sendo verdade o que se expôs, então, se o ciúme do marido em relação à sua esposa, afeta o foco da sua atenção, a vividez da evidência e a hipótese que vai formar em relação à sua esposa, poderemos concluir que a motivação influenciou este cenário.

Ora, isto sugere que os efeitos do ciúme são parcialmente explicados por um desejo de manter uma relação próxima e esse desejo, poderá ter estado presente na cognição enviesada do marido, que suspeita que a esposa tem um caso. O desejo, considerando o

contexto psicológico, incluindo a ciúme associado, pode ajudar a aumentar a saliência de evidência de ameaças à relação entre um marido ciumento e a sua esposa, pode ajudar a priorizar a confirmação enviesada, que favorece a crença de que a esposa está a ter um caso, etc.

Retomando o modelo FTL, e considerando o contexto mencionado, o desejo do marido ciumento, pode ter contribuído para a motivação mais forte, de evitar acreditar, falsamente, que a sua mulher é fiel, do que para evitar acreditar, falsamente, que a esposa é infiel e, ao mesmo tempo, contribuiu para que o marido possa aceitar, mais facilmente, a hipótese que a sua mulher está a ter um caso, do que aceitar a hipótese contrária.

Devido à conexão estreita entre emoções e desejos associados, o teste empírico de casos de auto engano nos quais a emoção, e não a motivação, desempenha um papel enviesador, será difícil (o ciúme é visto como tendo um desejo constituinte ou pelo menos que esse desejo é uma causa parcial. O mesmo acontece com o amor, ódio, ansiedade, terror, inveja, orgulho, etc.). Seria difícil demonstrar que existem casos de auto engano, nos quais uma destas emoções desempenha um papel enviesador, mas nenhum desejo associado o faz.

Mas seria um erro ignorar a potencial contribuição das emoções no auto engano, pois em muitos dos casos deste fenómeno, incluindo a forma *twisted*, as influências enviesadoreas provêm tanto das emoções como dos desejos, que estão intimamente relacionados com emoções enviesadoras, tanto como sendo uma parte do todo, ou sendo uma causa parcial ou um efeito, ou como resposta às emoções (como o caso de um desejo de ficar livre de uma ansiedade).

Em alguns casos, a influência enviesadora das emoções pode ser direta. Talvez uma emoção possa permitir o *confirmation bias* ou reforce a saliência de informação congruente com a emoção, sem que um desejo faça parte deste cenário e sem que os efeitos sejam causalmente mediados por um desejo²⁰⁷.

²⁰⁷ Por exemplo, Carl ficou zangado com o amigo Dan devido a uma ofensa. A sua raiva pode sugerir uma hipótese, congruente com essa emoção, acerca do comportamento de Dan (Dan comporta-se ofensivamente), e portanto permitindo o *confirmation bias*, podendo aumentar a saliência de informação que suporte essa hipótese.

Mas voltando ao caso do marido ciumento, se apenas soubéssemos a sua evidência relevante e do seu desejo de manter uma relação próxima com a sua esposa, sem mais informação adicional, não estaríamos na posição ideal para compreender a sua crença de que a esposa estaria a ter um caso. Por outro lado, pessoas com maior evidência de infidelidade (evidência que garantiria uma crença acerca dessa infidelidade), por vezes acreditam na fidelidade das esposas, ainda que desejem manter o relacionamento com a esposa. No caso descrito, a informação de que o marido tinha um perfil ciumento²⁰⁸, ajudou-nos a compreender a sua crença acerca da infidelidade da esposa. Ou seja, o facto de ser ciumento, constitui uma parte importante do contexto psicológico, no qual a crença é adquirida. É pois plausível, pensar que o ciúme desempenha um papel na produção da sua crença enviesada, que não é influenciada apenas por um desejo pertinente, por si só. Assim, se uma emoção pode desempenhar um papel enviesador direto no auto engano, está aberto o caminho para a possibilidade de que uma emoção pode contribuir para uma instância de auto engano, que não tem nenhum desejo como causa significativa²⁰⁹.

Alfred Mele defende pois, uma abordagem motivacional/emocional híbrida, acerca da *twisted self-deception* e acerca da *self-deception* em geral, que afirma a possibilidade de um viés motivacional direto, na ausência de causas motivacionais das crenças enviesadas em questão.

As visões motivacional e híbrida, relativamente à *twisted self-deception* são consistentes com a abordagem do autor em relação à *straight self-deception*²¹⁰.

Para Alfred Mele, a nossa preocupação relativamente aos casos do auto engano, deve ser similar à dos casos de ação acrática, na medida em que ambos os casos apresentam, tipicamente, uma irracionalidade influenciada pelos desejos.

O sujeito que age incontinentemente ao fazer *A*, age contra o seu melhor juízo, devido a um desejo contrário e o agente auto enganado, também devido ao desejo, tipicamente,

²⁰⁸ Parece possível e relevante destacar um paralelo entre o que Alfred Mele apelida de perfil e a noção de traço de personalidade, que constitui uma tendência para o indivíduo se comportar, sentir, perceber e pensar de formas relativamente constantes através do tempo e das situações nas quais o traço se pode manifestar. (American Psychiatric Association, *DSM-5*, op.cit, p. 988).

²⁰⁹ Por exemplo, o Carl instanciou o auto engano, ao adquirir a crença que o Dan está a agir ofensivamente e não existe um desejo do Carl a desempenhar o papel enviesador.

²¹⁰ MELE, *Self-Deception Unmasked*, op. cit., pp. 112-118.

acredita contra a sua melhor evidência, ou contra a melhor evidência, que poderia ter obtido ou facilmente adquirido, se não existisse o desejo em questão.

É óbvia a ocorrência destes dois tipos de irracionalidade e uma consideração importante a fazer é que, em ambos os casos, a influência causal dos nossos desejos não está totalmente sob o nosso controlo.

Por outro lado, tal como o agente incontinente não age tipicamente com a intenção de se comportar de forma incontinente, o agente, tipicamente, não cai no auto engano como resultado de uma intenção de se auto enganar. Mas ainda que não tenha existido intenção, em ambos os casos, o agente permanece culpado. Ou seja, pensamos ser correto culpar uma pessoa pelas suas ações incontinentes e, por outro lado, as pessoas são culpadas pelos autoenganos não intencionais, porque deviam ter sido mais cuidadosos em obter factos claros e diretos²¹¹.

3.2.7 Considerações

Se a interpretação da realidade, for deturpada pelos afetos, e o sujeito pode iludir-se de tal forma que as crenças adquiridas, irão afetar os juízos avaliativos, em função dos quais deliberamos, então podemos dizer que o sujeito de auto engana, ao formar as suas crenças de forma irracional.

Por exemplo, na sociedade atual é normal ouvirmos dizer que devemos manter o pensamento positivo (talvez aqui se possa estabelecer um paralelo com o *wishful thinking*), no sentido de sustentar a hipótese psicologicamente mais confortável e também como um mecanismo psicológico de resiliência, perante determinadas situação da vida. Mas, se considerarmos que, neste caso, um agente adota uma crença irracional, então o seu curso de ação poderá estar seriamente comprometido, não apenas ao optar pelo curso de ação/informação errado, mas ao ignorar outros cursos de ação/informação que, racionalmente, seriam mais benéficos.

Por exemplo, o positivismo exagerado, ou o extremo da expressão – querer é poder – pode levar o sujeito a entrar numa situação de esforço físico e emocional tão exagerado, que chega a provocar a exaustão física ou o tão mencionado *burn out*. Neste caso o sujeito

²¹¹ MELE, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, op. cit., pp. 136-137.

acreditou, falsamente, em algo que à luz de uma opinião imparcial, seria impossível efetuar, exerceu o auto controlo ao adotar estratégias de compromisso para o alcance desse objetivo, facto que o levou a um comportamento irracional e prejudicial a si próprio. Este caminho de adoção de uma crença irracional resultaria também da aplicação das estratégias de que Alfred Mele nos fala, tais como interpretação errónea positiva ou negativa, foco seletivo e recolha de evidência seletiva.

Daqui se compreende que algumas ações acráticas possam sê-lo, porque não coincidem com as razões racionais do agente. Com efeito, o agente está a representar a realidade através de um viés cognitivo e portanto, vai agir em função de um melhor juízo enviesado. Parece-me que o próprio autor partilha desta ideia ao referir que, o agente pode exibir auto controlo, não apenas em ações evidentes que estão de acordo com os seus julgamentos avaliativos, mas também na aquisição ou retenção de crenças. Tal como os agentes podem agir acriticamente, também podem acreditar acriticamente, tal como vimos nos casos de auto engano²¹².

Ainda relativamente à adoção e manutenção de crenças enviesadas, parece de referir um importante aspeto de cariz fisicalista, referido por António Damásio. Na relação entre a mente e o corpo, entra em funcionamento o sistema músculo-esquelético, órgãos internos, ao ambiente interno, mas também aos dispositivos de perceção, através dos quais entram no corpo os sinais vindos do exterior. E estes dispositivos possuem dois planos, um a que Damásio chama de portadas sensoriais (por exemplo, os músculos ao redor dos olhos, as componentes do globo ocular, a membrana do tímpano) e um outro plano, a sonda neural (por exemplo a retina do globo ocular onde são projetadas as imagens óticas).

A combinação das portadas sensoriais e da sonda neural, constituem a fronteira transposta pelos sinais externos. Ou seja, a representação do mundo apenas penetra no cérebro através do próprio corpo. E é através desta mapeação do corpo que o cérebro cria aquilo que se torna no eu²¹³.

²¹² MELE, *Autonomous Agents*, op. cit, p. 7.

²¹³ DAMÁSIO, *O Livro da Consciência*, op. cit., pp. 121-122.

Ora, o funcionamento complexo das portadas sensoriais, poderá contribuir para erros relativamente à fonte de uma percepção²¹⁴. As portadas sensoriais desempenham, portanto, um papel essencial na nossa perspetiva em relação ao mundo, ou seja afetam o nosso ponto de vista em relação ao que acontece²¹⁵.

Alfred Mele pretende justificar o paradoxo do auto engano, ultrapassando a tese da divisão da mente, segundo a qual seria o inconsciente do sujeito que estaria a enganar a sua parte consciente. Assim, segundo a conceção motivacional de Alfred Mele, e através das estratégias já mencionadas, seriam os desejos e as emoções do sujeito, que se tornam causas irracionais das crenças ilusórias do sujeito.

Estas estratégias podem levar o sujeito, de forma involuntária, à aquisição de crenças que vão estar em desacordo com aquilo que a informação disponível sugere. Ou seja, a interpretação da realidade pode, de facto, ser deturpada pelos desejos e emoções e esta deturpação será tanto maior, quanto mais essa realidade representar ou reforçar para o sujeito, um forte desejo ou emoção ou receio.

Outro aspeto pertinente é que, apesar de Alfred Mele defender uma irracionalidade motivada, que não depende da intencionalidade do sujeito, a aquisição de crenças auto enganadoras não inibe o sujeito de culpa. Assim, parece razoável supor que a noção de autocontrolo possa aqui ser introduzida, pois cada um de nós tem a responsabilidade de ponderar o próprio processo de formação de crenças. O grau de deliberação das decisões, conforme já referido em momento anterior, parece ser um tópico fundamental para a decisão do curso de ação a tomar, sendo esse o curso de ação desejavelmente racional.

A noção de grau de ponderação das decisões poderá ser colocado em paralelo com os métodos de raciocínio que Vasco Correia refere.

Se supusermos que as crenças irracionais são adquiridas de forma involuntária e inconsciente, então a responsabilidade epistémica pertence ao processo de formação de crenças e cada sujeito pode exercer um certo controlo sobre as condições nas quais forma

²¹⁴ Por exemplo, recordarmos que primeiro vimos determinado objeto e só depois o ouvimos num determinado acontecimento, quando se passou exatamente o contrário.

²¹⁵ DAMÁSIO, *O Livro da Consciência*, op. cit., p. 248.

essas crenças. Assim, a racionalidade das nossas crenças, que por sua vez culminaria em ações racionais, beneficia com a elaboração de métodos de raciocínio²¹⁶

Esses métodos de raciocínio são capazes de aumentar a fiabilidade na formação de crenças e Vasco Correia defende que somente com o estudo dos mecanismos de enviesamento do juízo, podemos refletir acerca dos métodos e estratégias capazes de evitar a irracionalidade das nossas crenças.

Essas estratégias passam por obrigar o sujeito a impor determinados constrangimentos às condições nas quais forma o seu juízo. Esta estratégia passa por adotar um controlo indireto. Por exemplo, pedir a um colega que reveja um texto que vamos publicar. Uma outra estratégia é a aquisição de certas competências de raciocínio, através do estudo das regras de inferência lógica ou regras de argumentação (por exemplo, existem muitas correlações entre as heurísticas e os enviesamentos). Outra estratégia reside nas virtudes epistémicas (hábitos propícios à salvaguarda da validade dos nossos raciocínios e das nossas crenças, ou seja, subjaz aqui a tendência de que temos raciocínios falaciosos involuntários, devido às ilusões cognitivas sobre os processos cognitivos). Por fim, as estratégias de compromisso prévio (uma pessoa que reconheça a sua tendência de tomar decisões irracionais impulsivamente, sob o efeito de forte emoção, pode adotar o hábito de nunca tomar decisões de cabeça quente)²¹⁷.

Apesar das influências internas (ao agente) e externas (contextuais), serem peças integrantes na intenção e ação, o ponto de vista do agente, relativamente a esse conjunto de fatores internos e externos, é que permitirá a real compreensão do processo de formação de crenças, bem como permitirá a interpretação da ação do sujeito, quer seja acrática ou não.

Em momento anterior, defendi que os fatores contextuais devem ser colocados a par com os fatores internos ao sujeito, facto que parece estar corroborado pela tese de Alfred Mele quando se refere ao desalinhamento entre a força motivacional e a avaliação dos objetos dos nossos desejos.

²¹⁶ CORREIA, Vasco, «Os limites da racionalidade: auto-engano e acrasia» *Disputatio*, Vol. III, No. 28 (2010) p. 290.

²¹⁷ CADILHA, S, MIGUENS, (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op.cit, pp. 184-185.

Nesta linha, podemos analisar a abordagem contextualista que Vasco Correia desenvolveu, segundo a qual as estratégias cognitivas para prevenir o enviesamento têm que ser complementadas por mecanismos extrafísicos, que dependem dos constrangimentos sociais e ambientais, de forma a promoverem o raciocínio racional.

Aceita a ideia de que os enviesamentos podem ser adaptativos em determinados contextos, ou seja, são racionais do ponto de vista instrumental²¹⁸, e que esses enviesamentos podem ser toleráveis sem precisar de correção.

Mas por outro lado, defende que esses enviesamentos podem ser desadaptados em muitos outros contextos²¹⁹, precisando assim de ser contrariados, através de estratégias de desenviesamento – *debiasing strategies* (estratégia ou estratégias que suprimem ou atenuam os enviesamentos ou, pelo menos, os seus efeitos).

Estudos sugerem que certas estratégias de desenviesamento parecem ser eficazes na redução de alguns tipos de viés, mas não em outros e existem estratégias ineficazes.

De forma a superar este impasse, Vasco Correia apresenta uma visão segundo a qual o viés cognitivo, pode ser complementado por uma variedade de mecanismos contextuais. Ou seja, tendo em mente a abordagem contextualista, os mecanismos extrafísicos (ou ambientais), são mais eficazes contra os enviesamentos do que o pensamento crítico, por si só, na medida em que assentam em constrangimentos externos e estruturas sociais, que contrariam as causas principais do enviesamento persistente, nomeadamente: a não consciência do sujeito acerca do seu próprio viés; limitações cognitivas; falta de motivação para evitar o enviesamento; e correção inadequada.

O que acaba de ser dito não significa que as estratégias de desenviesamento intrafísicas (ou cognitivas) são ineficazes na prevenção do viés, mas apenas que deve existir uma complementaridade entre ambas as abordagens.

²¹⁸ Alguns autores sugerem, por exemplo, que as avaliações em função do auto interesse do sujeito, ilusões de controlo ou otimismo irrealista contribuem para a motivação, humor, criatividade, produtividade, persistência, performance a até a habilidade de cuidado com os outros. Ou seja, ao invés de procurarem a otimização da racionalidade, servem a maximização dos interesses do agente, em contextos da vida real.

²¹⁹ Por exemplo, o otimismo irrealista e excesso de confiança, podem levar o sujeito a correr riscos desnecessários.

Defende que algumas formas de desenviesamento contextual fortalecem o pensamento crítico e a alteração cognitiva, tendo desenvolvido uma distinção entre dois tipos de estratégias contextuais:

Correção situacional – aquela que apenas reduz os efeitos de um determinado viés; Mas se o propósito das estratégias é simplesmente atenuar as repercussões indesejadas do viés, sem modificar o processo cognitivo do sujeito, então o viés irá permanecer em simultâneo com as respetivas características disposicionais²²⁰.

Correção disposicional – aquela que implica a redução do próprio viés.

Estes últimos mecanismos parecem ser mais capazes de promover as competências do pensamento crítico, bem como as disposições ao nível do indivíduo, ainda que indiretamente. Estes tipos de estratégias requerem, tipicamente, constrangimentos externos que forçam o sujeito a detetar e corrigir potenciais enviesamentos no seu raciocínio.

Vasco correia identifica os seguintes mecanismos de correção disposicional:

- responsabilização – *accountability*. A eficácia deste mecanismo reside no facto de motivar o sujeito a acionar esforços de desenviesamento, acreditando que outros irão analisar os seus argumentos. Assim, o sujeito irá querer assegurar-se de que os seus argumentos são fortes e racionais. Esta estratégia tem, no entanto, as suas limitações na medida em que pode amplificar outros tipos de viés: um sujeito perante um grupo com perspetivas conhecidas, pode acabar por dizer aquilo que a audiência quer ouvir.
- interação de grupo – *group interaction*. Os grupos funcionam como um sistema de verificação de erros durante a interação e é mais provável que surjam perspetivas diversas, o que promove o pensamento crítico. Este mecanismo funciona em grupos com diferentes visões, senão pode ocorrer o fenómeno de polarização de grupo, no qual os sujeitos reforçam a sua opinião partilhada.
- Incentivos - *incentive*. Este mecanismo pode contribuir para reduzir certos enviesamentos, na medida em que aumenta a motivação do sujeito para a redução do

²²⁰ Um exemplo destas estratégias é o controlo da fonte de informação – *blinding* – uma vez que, se o estímulo nunca entrar na mente do sujeito então não poderá enviesar o seu julgamento ou emoções. O problema deste método é a dificuldade na sua implementação, pois não sabemos de antemão quais os estímulos que provocarão o viés. E por outro lado envolve manipulação de informação.

enviesamento. Parece razoável assumir que um sujeito com mau desempenho cognitivo tem maior probabilidade de se tornar epistemicamente vigilante, se puder ganhar algo com esse processo. No entanto este mecanismo pode atenuar o problema de falta de motivação para o desenviesamento, mas pode por outro lado, exacerbar o problema da correção inadequada.

Assim, o cerne da questão não reside no facto de saber se, no geral, o desenviesamento funciona ou não, mas sim perceber que técnica de desenviesamento específica irá funcionar melhor, em cada tipo de tarefa e quais os efeitos colaterais que poderá ter.

Deste modo, ao submeter o seu pensamento às estruturas sociais e ambientais certas, os indivíduos têm mais probabilidade de estar conforme as normas de racionalidade. Mas, garantidamente, cada técnica de desenviesamento terá as suas falhas.

A tese que Vasco Correia defende é que cada mecanismo pode levar a melhorias significativas, como um todo²²¹.

3.3 *Free Will*

A noção de auto controlo do agente abre caminho para suportar a ideia de autonomia ou de livre arbítrio. Somos, de facto, agentes que agem livremente? E somos responsáveis por essas instâncias de agência?

Em *Autonomous Agents*, Alfred Mele refere a relação entre o auto controlo (considerando esta noção como oposta à noção de acrasia) e autonomia. Através do autocontrolo, o agente autorregula-se e autogoverna-se²²². Este mecanismo permite a projeção do agente tanto no presente, como no futuro, através das decisões mais imediatas ou decisões que fazem parte de planos de médio ou longo prazo.

Quando estão em causa, entidades que desempenham determinados movimentos corporais baseados nessas crenças, desejos e intenções, ou seja, agentes, forçosamente levantam-se questões que se relacionam com a responsabilidade desse agente. Ora, apenas os agentes humanos são responsáveis por algumas das suas ações, porque para que

²²¹ CORREIA, Vasco. «Contextual Debiasing and Critical Thinking: Reasons for Optimism» *Springer Science + Business Media Dordrecht*, (2016).

²²² MELE, *Autonomous Agents*, op. cit., prefácio.

um agente seja responsável, não basta que a sua ação seja intencional, é também necessário o auto controlo.

O passo seguinte será a autonomia desse agente e ainda deve ser considerado o facto de o agente ter tido a oportunidade de agir de outra forma.

Assim, a racionalidade ou irracionalidade da ação dos agentes, pode ser analisada em relação às suas razões para agir.

A deliberação é uma das formas pela qual o agente age, por razões e, para essa deliberação, o agente adquire a intenção da ação. Essa intenção é a atitude executiva relativamente aos planos.

Considerando as relações entre crenças desejo, deliberação e motivação, a questão do livre arbítrio coloca-se de forma global, considerando a nossa natureza como agentes. Somos, de facto, livres²²³?

Conforme vimos, Alfred Mele considera que a dimensão do auto controlo, se relaciona com a autonomia individual. A vertente do auto controlo como sendo o contrário da acrasia e a vertente que o autor pretende explorar, no sentido de explicar a relação de uma compreensão apropriada do auto controlo e manifestação comportamental, com a autonomia pessoal e comportamento autónomo²²⁴. A autonomia está associada com outros conceitos de liberdade, tais como, ação livre, livre escolha, ação livre, etc²²⁵.

«Is it true that at least some human beings sometimes act freely and are morally responsible for some of what they do?»²²⁶. Esta é a questão com a qual Alfred Mele inicia *Free Will and Luck*.

As defesas filosóficas tradicionais acerca destas questões incluem a defesa de posições sobre se a ação livre e a responsabilidade moral são compatíveis ou incompatíveis com o determinismo.

Alfred Mele define-se agnóstico acerca deste tema e avança alguns conceitos básicos que facilitam a compreensão daquilo que vai defender:

²²³ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 5-7.

²²⁴ MELE, *Autonomous Agents*, op. cit., Prefácio.

²²⁵ MELE, *Autonomous Agents*, op. cit., p. 4.

²²⁶ MELE, Alfred, *Free Will and Luck*, Oxford University Press, New York 2006, p. 3.

Determinismo – a tese segundo a qual existe, a cada instante exato, um futuro fisicamente possível, ou seja, em cada instante exato, só um futuro é compatível com o estado do universo nesse momento e com as leis da natureza.

Compatibilismo – a tese segundo a qual a ação livre e a responsabilidade moral são compatíveis com a verdade do determinismo (um compatibilista pode defender ou não que existem seres humanos livres e moralmente responsáveis e pode defender ou não que o determinismo é verdadeiro²²⁷).

Incompatibilismo – a tese segundo a qual nem a ação livre nem a responsabilidade moral são compatíveis com a verdade do determinismo (um incompatibilista pode defender ou não que existem seres humanos livres e moralmente responsáveis e pode defender ou não que o determinismo é verdadeiro).

Libertarianismo – conjunção do incompatibilismo com a tese de que alguns seres humanos agem livremente e são moralmente responsáveis por algumas coisas que fazem. Obviamente que o libertarianismo e o compatibilismo são mutuamente exclusivos.

Alfred Mele considera-se agnóstico relativamente à verdade do compatibilismo e defende a posição a que chama de autonomismo agnóstico, que consiste na conjunção deste agnosticismo com a afirmação de que, existem seres humanos livres e moralmente responsáveis.

Para Alfred Mele, o autonomismo agnóstico é mais credível do que a visão segundo a qual não existem seres humanos livres e moralmente responsáveis (NFM) e apresenta as seguintes proposições:

- a. Alguns seres humanos, por vezes, agem de forma livre e moralmente responsável em parte daquilo que fazem, e o determinismo é compatível com isso (crença compatibilista em seres humanos livres e moralmente responsáveis).

²²⁷ Em princípio um teórico pode ser um compatibilista acerca do determinismo e ação livre, ao mesmo tempo que é incompatibilista acerca do determinismo e responsabilidade moral, e vice-versa. Na discussão destas visões, poderá haver quem seja mais refinado no uso da expressão "compatibilismo". Por razões de estilo, Alfred Mele entende que a definição mais abrangente de compatibilismo é preferível, considerando o propósito da obra em referência (MELE, *Free Will and Luck*, op. cit., p. 25).

- b. Alguns seres humanos, por vezes, agem de forma livre e moralmente responsável em parte daquilo que fazem e o determinismo é incompatível com isso (libertarianismo).
- c. Ou a) ou b) (autonomismo agnóstico)
- d. Nenhum ser humano age alguma vez de forma livre e nenhum ser humano é moralmente responsável por nada (NFM).

Alfred Mele utiliza os recursos do libertarianismo e do compatibilismo para defender a sua posição agnóstica, procurando desenvolver as posições compatibilistas e libertárias que considera melhores. Constrói a ideia de um agente perfeitamente autocontrolado e defende que, mesmo esse agente, pode exibir falta de autonomia sendo necessário perceber o que faltaria a esse agente para ser entendido como um agente autónomo.

E para responder a esta dúvida, Alfred Mele oferece duas respostas, uma para os compatibilistas e outra para os libertários. Propõe então que o autonomismo agnóstico é mais credível que a NFM²²⁸.

A posição do autonomismo agnóstico consegue recorrer a recursos tanto do compatibilismo como do libertarianismo, oferecendo um conjunto de condições suficientes e robustas para a autonomia compatibilista e um conjunto de condições coerentes para a autonomia incompatibilista. Além disso, o autonomismo agnóstico não se compromete com a ideia de que o autonomismo é compatível com o determinismo, nem insiste que somos internamente indeterministas, na forma a que se refere o libertarianismo²²⁹.

O autor apresenta, portanto, uma crença disjuntiva da crença do libertarianismo e do compatibilismo acerca do agente livre e moralmente responsável²³⁰.

Alfred Mele desenvolve vários modelos acerca do livre arbítrio: o libertarianismo modesto – *Modest Libertarianism*, Compatibilismo suave – *Soft compatibilism*, Libertarianismo suave – *Soft libertarianism* e o Libertarianismo suave ousado – DSL – *Daring soft Libertarianism*.

²²⁸ MELE, *Free Will and Luck*, op. cit., pp. 3-5.

²²⁹ MELE, *Autonomous Agents*, op. cit. p. 253.

²³⁰ MELE, *Free Will and Luck*, op. cit. p. 6.

Considerando que Alfred Mele é recetivo tanto ao compatibilismo, como ao libertarianismo, desenvolve, portanto, duas respostas, uma para os compatibilistas e outra para os libertários²³¹. E estas repostas sustentam os melhores argumentos de ambas as posições.

3.3.1. A resposta libertária

Os libertários alegam que, por exemplo, um agente fez livremente *A* num momento *t*, se e só se, pudesse ter feito de outra forma além de *A* ou, podem alegar que, um agente que não pudesse ter agido, num momento *t*, além de *A* então pode, no entanto, fazer *A* livremente, num momento *t*, desde que o sujeito tenha desempenhado anteriormente alguma(s) ação/ações livre(s) relevante(s), num determinado momento ou momentos, nos quais pudesse ter feito de outra forma, ao invés de desempenhar essas ações.

Relativamente ao acaso – *luck* – o controlo do agente é o critério pelo qual é medida a influência do acaso na liberdade ou responsabilidade moral. Quando o acaso (bom ou mau), é problemático, é porque parece impedir significativamente o controlo dos agentes sobre si próprios, ou destaca lacunas importantes nesse controlo. Por exemplo, um agente tenciona de acordo com o seu juízo acerca do que é melhor fazer, perde o controlo sobre a sua intenção, pode ser alegado que uma conexão determinista positiva, entre o melhor juízo considerado e a intenção, teria sido mais condutora à liberdade e responsabilidade moral²³².

No mundo atual um agente decide fazer *A* num momento *t* e noutro mundo com as mesmas leis da natureza e o mesmo passado, esse agente decide num momento *t*, não fazer *A*. Se, acerca dos poderes desse agente, capacidades, estados mentais, caráter moral, nada conta para essa diferença, em nenhum dos mundos, então a diferença parece ser uma questão de acaso. E dado que os mundos não divergem entre si, antes do momento *t*, não existe diferença nenhuma no agente, nestes dois mundos, que releve para a diferença nas suas decisões. Ou seja, o que poderá explicar isto é que esse agente é um decisor

²³¹ MELE, *Free Will and Luck*, op. cit. pp. 163-164.

²³² MELE, *Free Will and Luck*, op. cit., pp. 6-7.

indeterminista. Esta explicação é consistente com o facto da diferença nas suas decisões, ser uma questão de acaso - *luck*²³³.

3.3.1.1. Libertarianismo Modesto

De acordo com a visão libertária do típico evento causal, as causas proximais das ações livres causam-nas, indeterministicamente. Ou seja, a consequência da ideia do típico evento causal libertário é que as ações livres têm causas proximais e se um agente desempenhou *A*, livremente, num momento *t*, num mundo *W*, esse agente não vai fazer *A* num momento *t*, num outro mundo possível com as mesmas leis da natureza e o mesmo passado até ao momento *t*.

As causas proximais das ações, incluindo ações que são decisões, são internas ao agente, logo a visão libertária do típico evento causal, engloba o compromisso com aquilo que pode ser chamado de indeterminismo interno do agente.

Aquilo a que Alfred Mele designa de libertarianismo modesto – *modest libertarianism* – engloba esse compromisso, mas rejeita a ideia de que as causas proximais das ações livres causam indeterministicamente as ações.²³⁴

3.3.1.2. Libertarianismo suave

O libertarianismo suave - *Soft libertarianism* dá uma perspetiva da responsabilidade moral e liberdade, através da qual os compatibilistas podem achar mais fácil perceber o motivo pelo qual alguém pode, razoavelmente, valorizar a agência indeterminista como podendo contribuir para a responsabilidade moral e a liberdade.

Os libertários tradicionais são incompatibilistas fortes. Defendem que a ação livre e a responsabilidade moral são incompatíveis com o determinismo. Uma linha mais suave estará aberta, para os teóricos com simpatias libertárias, que deixem em aberto que a ação livre e a responsabilidade moral são compatíveis com o determinismo, mas que mantêm a ideia de que a falsidade do determinismo é necessária para as espécies de ação livre e responsabilidade moral, mais desejáveis. Esta é a linha libertária suave.

²³³ MELE, *Free Will and Luck*, op. cit. p. 9.

²³⁴ MELE, *Free Will and Luck*, op. cit., pp. 9-10.

Os libertários suaves ficariam desapontados ao descobrir que o determinismo é verdadeiro, mas não concluem que nunca ninguém agiu livremente ou nunca ninguém foi moralmente responsável por nada. Um libertário suave, em princípio, pode ou não tratar a desejabilidade atrás referida, de forma a relativizá-la para os indivíduos. Assim, o libertarianismo suave, apresentado por Alfred Mele, é relativista, na medida em que mantém que, pelo menos alguns agentes humanos, possuem tipos de liberdade e de responsabilidade moral, incompatíveis com o determinismo e que são razoavelmente desejados por pelo menos alguns desses agentes, a qualquer tipo de liberdade ou de responsabilidade moral, consistentes com a verdade do determinismo.

Ao contrário dos libertários fortes, os libertários suaves deixam em aberto que o determinismo é compatível com o facto de as nossas ações serem algo que depende de nós – *up to us* – numa forma que conduz à liberdade e responsabilidade moral. No entanto, acreditam que uma liberdade e responsabilidade moral mais desejáveis requerem que as nossas ações não façam parte do desenrolar da cadeia de eventos determinista, que estava em progresso mesmo antes de nascermos. Se os libertários suaves podem ver-se a si próprios a desempenhar algumas ações ou decisões, que não são causadas deterministicamente ou são causadas deterministicamente por, por exemplo, algo que inclui juízos deliberativos que não são, eles próprios, deterministicamente causados, então, podem ver-se a si próprios a iniciar alguns processos causais, que não são ligações intermediárias numa longa cadeia causal determinística, que se aproxime do *big bang* (a cadeia pode ir no sentido inverso: duas ou mais sequências de eventos causalmente independentes, podem combinar de forma a emitir um evento subsequente). E os libertários suaves assumem o poder de serem um iniciador deste tipo, um iniciador indeterminista, requisito para uma liberdade e responsabilidade mais desejáveis.

Os libertários suaves não precisam de afirmar que todos os valores dos agentes (ou de todas as pessoas racionais) são iniciadores indeterministas. Algumas pessoas podem preferir pensar que o seu mundo é determinista ou indeterminista, apenas se isso não tiver uma influência importante na sua agência, e podem não ter o desejo de ser um iniciador indeterminista. Outros, evidentemente, entendem o determinismo como desconfortavelmente limitador.

Os libertários suaves garantem que, da sua perspectiva, certas espécies incompatibilistas de liberdade e responsabilidade, são mais desejáveis do que qualquer espécie compatibilista e a sua própria perspectiva não é irracional. Aquilo em que os libertários suaves insistem é que, a partir da sua própria perspectiva, certas espécies incompatibilistas de liberdade e responsabilidade são mais desejáveis, do que qualquer espécie compatibilista e que a sua própria perspectiva não é irracional. Valorizam estas espécies incompatibilistas de liberdade e responsabilidade, mais do que valorizam qualquer espécie compatibilista e não identificam nisso nada de irracional.

O autonomismo agnóstico, é agnóstico acerca do determinismo ser compatível com a agência livre e acerca da existência da agência livre libertária no nosso mundo, mas compromete-se com a coerência conceptual da livre agência libertária e com a tese de que a crença de que existem agentes livres é mais defensável do que a crença de que não existem.

O libertarianismo suave é consistente com o autonomismo agnóstico, mas é um pouco mais arrojado. Os libertários suaves recusam a ideia de que as espécies incompatibilistas de agência livre, são realizadas no nosso mundo²³⁵.

3.3.1.3 *Daring Soft Libertarianism*

Num agente com a arquitetura psicológica do libertarianismo modesto, certas decisões derivam do agente, ou seja, o conjunto dos estados internos do agente e os eventos que explicam essa decisão, não são deterministicamente causados por nada externo ao agente. O agente fica com a responsabilidade última sobre si, pois a natureza indeterminista da deliberação que leva às decisões, garante a ausência de uma conexão determinista ao passado do agente, relevante para a componente final da responsabilidade última. Mas, esta ausência de conexão determinista pode ser garantida durante a deliberação, de modo a não envolver uma operação cerebral do agente indeterminística. O libertarianismo modesto apresentado por Alfred Mele, é consistente com a visão compatibilista de que, parte de um processo que pode dar origem a ações livres, parte essa que começa com uma crença racionalmente adquirida de que é melhor, no todo, fazer *A*,

²³⁵ MELE, *Free Will and Luck*, op. cit., pp. 95-102.

inclui a aquisição imediata de uma intenção de *A*, e termina pouco depois no facto de fazer *A*, o que executa essa intenção²³⁶.

Alfred Mele desenvolve uma visão que, pretende, possa enfrentar o acaso presente – *Daring soft libertarians* – DSL, em vez de seguir o libertarianismo modesto.

Os apoiantes da DSL defendem que o acaso presente está envolvido pelo facto do agente ter um certo poder de iniciação que a DSL valoriza, e cuja presença, em caso de ação, não exclui o facto de a ação ser desempenhada livremente, nem o facto de o agente ser moralmente responsável por essa ação. Os libertários suaves, não defendem que a ação livre e a responsabilidade moral, requerem a falsidade do determinismo.

O DSL que Alfred Mele defende, foca-se na decisão e valoriza o poder de efetuar decisões, que não são deterministicamente causadas – um certo poder de iniciação – *initiatory power*. Este poder de iniciação está patente, apenas quando as causas proximais de uma decisão, não a causam deterministicamente.

Alfred Mele recorda os conceitos de ação basicamente livre – *basically free action* – e ação pela qual um agente é basicamente moralmente responsável.

As ações basicamente livres, são fazer *A* livremente, que ocorrem em momentos nos quais o passado e as leis da natureza são consistentes com o agente não fazer *A*, nesse caso. De forma similar, os desempenhos de *A*, pelos quais o agente é basicamente moralmente responsável, são desempenhos de *A*, pelos quais o agente é moralmente responsável, que ocorrem em momentos nos quais o passado e as leis da natureza são consistentes com o não desempenho de *A*, nesse caso. Estas são noções essencialmente incompatibilistas.

Apenas e só, quando um agente desempenha *A* livremente, num momento no qual tem, pelo menos, uma alternativa possível não robusta, a ação é basicamente^{237*} livre. De forma similar, apenas e só quando um agente desempenha *A*, num momento no qual tem, pelo menos, uma alternativa possível não robusta, e esse agente é moralmente responsável pelo desempenho de *A*, o agente é basicamente* moralmente responsável pelo desempenho de *A*. Desempenhar livremente *A* basicamente*, difere do desempenhar

²³⁶ Esta afirmação é consistente com o exemplo de Beth ter livremente rejeitado uma oferta de emprego. A sua rejeição foi causada deterministicamente pelos antecedentes imediatos e esses antecedentes imediatos foram deterministicamente causados por algo, que incluiu o evento causado indeterministicamente, de Beth ter adquirido a crença de que é melhor, no todo, rejeitar a oferta de emprego.

²³⁷ O (*) diferencia os termos basicamente* e basicamente (sem*).

livremente *A* basicamente, no sentido de que este último (mas o primeiro não), requer que o não desempenho de *A*, esteja em aberto ao agente no momento da ação, dado o passado e as leis da natureza. Paralelamente existe uma diferença entre responsabilidade moral básica* e responsabilidade moral básica para uma ação.

O termo a que se refere *basicness** é mais inclusivo do que o termo *basicness*: o seu requisito para possibilidades alternativas no momento da ação, é mais fraco.

Conforme referido, para os libertários está em aberto que as ações basicamente livres de um agente, adequadas em relação a ações posteriores, que não são basicamente livres, conferem liberdade para essas ações posteriores e a responsabilidade moral pode também ser conferida desta forma.

O modelo DSL adota esta posição e também adotam a mesma posição, relativamente à ação basicamente* livre e moralmente responsável. O modelo vê o facto de o agente ser capaz de desempenhar algumas ações, cujas causas proximais não causem deterministicamente as ações, como uma condição necessária para o agente ter um tipo de poder de iniciação. No modelo DSL, a importância deste poder de iniciação não está limitado aos casos nos quais é exercido. É também importante, na medida em que as ações basicamente* livres e moralmente responsáveis, que dependem do seu exercício, conferem um tipo incompatibilista de liberdade e responsabilidade moral em algumas ações subsequentes.

A visão DSL perspetiva o poder de desempenho de ações basicamente* livres e moralmente responsáveis, como um requisito necessário, para ter um tipo de agência livre e moralmente responsável, que o modelo DSL valoriza mais, do que qualquer espécie de agência compatibilista, mas o modelo valoriza também o trabalho de permitir não ficar sujeito ao acaso, numa esfera prática importante, mesmo que o sucesso deste esforço possa excluir o modelo de outras ações basicamente* livres e moralmente responsáveis nesta esfera.

Alfred Mele compara a visão DSL com o jogo *blackjack*, no qual os jogadores têm opções, por exemplo, de pedir outra carta, recusar cartas, fazer apostas. Ou seja, o *blackjack* envolve um misto de competências e de acaso - *Luck*. Tal como o acaso (apesar de não necessariamente acaso indeterminista) é uma parte do jogo, o facto de estar sujeito

ao acaso presente, de acordo com o DSL, é uma parte essencial do facto de ser um agente capaz de desempenhar ações basicamente* livres e ações pelas quais é basicamente* moralmente responsável.

Considerando de novo o jogo *blackjack*, os jogadores que pretendem maximizar as suas chances de ganhar, devem aprender de que forma podem minimizar as potenciais consequências do azar e maximizar as potenciais consequências da sorte. Assim, podem aprender a contar as cartas memorizar um bom conjunto de mesas de *blackjack* e jogar nessa conformidade. O que deve fazer o modelo DSL relativamente à sorte (ou acaso)?

Considerando o modelo DSL, o agente deve tentar tornar-se tão autocontrolado, que não subsiste a hipótese de agir acráticamente.

Assim, uma vez que o DSL abarca o acaso/sorte, resta então perceber de que forma o DSL responde à preocupação da presença do acaso, numa instância de ação, incompatível com o facto de a ação ser desempenhada livremente e o agente ser moralmente responsável por essa ação.

Portanto, Alfred Mele procura explicar a *strict akratic action* à luz da DSL, considerando que o acaso presente é especialmente importante na esfera da ação acrática e ação continente.

Começa por referir um recurso convencional libertário: na visão libertária de evento causal, um agente tentado a agir contra o melhor juízo, pode tentar decidir de acordo com o mesmo e tentar decidir, em simultâneo, em função do curso de ação tentador. Se o agente decide acráticamente ou não, será uma questão de qual dos dois esforços sucede. Conforme vimos, o caminho comum da aquisição do melhor juízo a favor de *A* para a intenção de *A* é um caminho por *default*. Quando o agente, acráticamente, falha ao não intencionar de acordo com o melhor juízo, então verificou-se alguma oposição que bloqueou o caminho por *default*. De igual forma, o caminho por *default* não necessita de esforços de auto controlo para desempenhar a ação favorecida pelo melhor juízo.

Mas, noutros casos, o agente cujo melhor juízo favoreça fazer *A*, pode tentar desempenhar estratégias, no sentido de agir ou intencionar em conformidade com esse melhor juízo, por exemplo, tentar focar a sua atenção nos resultados desejáveis de fazer *A*, ou nos

aspectos negativos de não fazer *A*. Pode gerar imagens vívidas acerca de ambas as situações ou efetuar auto comandos ou até mesmo procurar um terapeuta comportamental. Consideremos o caso de Drew, que após beber um *whiskey*, entende que seria melhor trocar para o café, de modo a poder conduzir para casa em segurança. Apesar de acreditar que poderia tomar café, decide tomar outro *whiskey* e esta decisão foi livre, ou seja, não fez nenhum exercício de auto controlo para desempenhar a ação que julgou ser melhor, apesar de sentir um desejo modesto de o fazer. Esta descrição corresponde a uma ação acrática.

Segundo a visão DSL, a ocorrência de ações acrálicas e continentes basicamente* livres, implica o acaso, no momento da ação, e esse acaso não exclui a liberdade ou responsabilidade moral. No caso de Drew, várias probabilidades relevantes, baseiam-se em várias características da sua condição psicológica, no momento *t*, no qual adquire o melhor juízo de passar para o café. As probabilidades relevantes incluem a decisão imediata de decidir beber outro *whiskey*, a decisão imediata de trocar para o café ou imediatamente decidir exercer autocontrolo, e trocar para o café, ou permanecer indecisa durante algum tempo, acerca do que fazer. E as características psicológicas que suportam, em parte, estas possibilidades incluem a força de vários desejos que a Drew tem no momento *t*: por exemplo, querer tomar outro *whiskey*, desejar evitar a possibilidade de ter um acidente, desejar não infringir a lei e o desejo de exercer o autocontrolo. Outra característica psicológica pertinente inclui (mas não se limita) ao grau de confiança no momento *t*, de que mudar para o café seria melhor e a força de qualquer desejo genérico, que ela possa ter no momento *t*, acerca do que pensa ser melhor fazer.

Segundo a visão DSL, estas características não são aleatórias. São em parte, produtos do comportamento passado da Drew²³⁸.

A DSL mantém a defesa de que o agente desempenha um papel indeterminístico no facto de moldar as probabilidades de agir continentemente e acriticamente, na medida em que desempenham um papel indeterminístico na modelação de coisas que suportam essas probabilidades. Muitos agentes têm a capacidade de influenciar, no presente,

²³⁸ Por exemplo, o desejo de não querer aumentar a possibilidade de ter um acidente é influenciado pelas suas crenças acerca de ocasiões prévias, nas quais conduziu sob efeito de álcool, crenças essas que não teria se não tivesse já incorrido, no passado, neste comportamento de risco.

probabilidades futuras de ação relevantes e é por vezes verdade que o seu comportamento passado influencia as probabilidades de ação presentes.

Portanto, a influência do passado nas probabilidades de ação presente é considerável.

Em parte, e acerca da influência, a DSL advoga que as probabilidades práticas das ações nem sempre são impostas ao agente, porque pelo seu comportamento passado, o agente molda as probabilidades práticas presentes e, no seu comportamento presente, moldam as probabilidades práticas futuras.

As probabilidades de ação (probabilidades práticas) nem sempre são impostas pelo agente. Através do comportamento passado, o agente molda as suas probabilidades práticas futuras.

Assim, a DSL enfatiza outro tipo de oportunidades que os agentes possuem relativamente ao acaso: a oportunidade de aprendizagem a partir de eventos associados ao azar, especialmente derivados de decisões basicamente* livres, contrárias ao seu melhor juízo racional. Um agente pode aprender através de consequências más, pela reflexão acerca dessas consequências e mesmo que essas más consequências não tenham ocorrido, o agente pode refletir acerca do que poderia ter acontecido de mal.

A visão DSL mantém que, na maioria das ações basicamente* livres e ações pelas quais o agente é basicamente* moralmente responsável, os agentes têm alguma responsabilidade pelas probabilidades práticas relevantes. Por exemplo, têm alguma responsabilidade pela probabilidade de agir acriticamente ou continentemente. E estas probabilidades não são ditadas por forças externas e são influenciadas pelas ações basicamente* livres e moralmente responsáveis, que os agentes desempenharam no passado²³⁹. Infere-se portanto que os agentes podem aprender com os erros acráticos do passado, passando a conseguir gerir com sucesso situações similares que ocorram no futuro.

3.3.2. A resposta compatibilista

A proposta compatibilista adiciona, ao agente idealmente autocontrolado e mentalmente são, que regularmente exerce o seu poder de auto controlo o seguinte: o agente não possui

²³⁹ MELE, *Free Will and Luck*, op. cit., pp. 111-123.

atitudes produzidas de forma compelida ou coerciva. As suas crenças conduzem a uma deliberação informada acerca de todos os assuntos que lhe importam e o agente é um deliberador fiável. Quando todos estes fatores são verdade num agente que, de forma não desviante, desempenha *A*, com base num juízo deliberativo formado racionalmente de que é melhor fazer *A*, então esse agente faz *A* livremente.

Alfred Mele relata a história de Ann e Beth. Ann é uma Filósofa excecional e diligente e Beth, apesar de ser igualmente talentosa, valoriza²⁴⁰ outras coisas, além da Filosofia. O superior hierárquico pretende que Beth seja como Ann e para isso, sem que ninguém saiba, contrata Psicólogos para determinarem a forma de funcionamento de Ann e uma equipa de *brainwashers* que transformem a Beth, para que seja igual a Ann. Os psicólogos concluem que a hierarquia de valores de Ann é o motivo pelo qual é tão produtiva e os *brainwashers* incutem esta hierarquia de valores em Beth (*através de new-wave brainwashing*), ao mesmo tempo que eliminam os valores concorrentes e Beth transforma-se num clone de Ann.

Neste exemplo, percebemos que Ann desempenha o seu trabalho filosófico de forma livre, mas Beth que não teve oportunidade de resistir ao processo de alteração dos valores, viu a sua direção de vida manipulada, ou seja, a sua autonomia foi violada e podemos vê-la como um agente não livre numa esfera importante da sua vida. Então, apesar das similaridades psicológicas entre Ann e Beth, a diferença relativamente à sua liberdade, reside na forma pela qual ambas as agentes adquiriram as suas características psicológicas e também em algo externo à sua constituição psicológica presente. Ou seja, a diferença crucial é histórica, a livre agência está, de certa forma, ligada à história.

O facto de a história ser importante para um compatibilismo plausível, gera aquilo a que Alfred Mele chama de autonomia psicológica, especialmente a posse autónoma de valores particulares e deliberação autónoma. A autonomia psicológica não se estende para lá da cabeça do agente. Para o autor, os compatibilistas conseguem atingir o seu objetivo se demonstrarem de que forma o determinismo é compatível com a autonomia psicológica.

²⁴⁰ Alfred Mele entende o significado de valorizar da seguinte forma: *S* pelo menos valoriza levemente *X*, num momento, se e só se, nesse momento, *S* tem uma atitude motivacional positiva em relação a *X* e acredita que *X* é bom. Estas duas dimensões são tidas em conta, quando os valores são entendidos como estados psicológicos.

Alfred Mele enuncia a seguinte afirmação:

*NF. An agent who performs an overt action A does not freely A, if (1) he expresses unsheddable values in A-ing; (2) owing directly to those values, he could not have done otherwise than A in the circumstances (on a compatibilist reading of “could have done otherwise”); (3) those values were very recently produced in a way that bypassed his capacities for control over his mental life by value engineering to which he did not consent and are seriously at odds with autonomously acquired values of his that were erased in the process; (4) he retains no preexisting value that is promoted by his having the unsheddable values he expresses in A-ing; and (5) A is the first overt action he performs on the basis of his new values*²⁴¹.

O facto de o agente expressar um valor numa ação tem uma dimensão causal essencial. Por definição, uma ação *A* expressa um valor *V* apenas se *V* (ou o seu realizador neural) desempenha um papel causal não desviante na produção de *A*²⁴².

De ressaltar que o autor não considera que as condições suficientes para a ação livre, na proposta compatibilista vencem os libertários, uma vez que Alfred Mele ofereceu aos libertários uma visão que derrota o problema no acaso presente²⁴³.

3.3.3 O conceito de Free Will

Qual é a evidência que suporta a existência do livre arbítrio? Quando é questionado sobre a sua crença na existência do livre arbítrio, Alfred Mele inicia a discussão dizendo que depende do significado atribuído a livre arbítrio – *free will*²⁴⁴.

Alfred Mele, focando-se nos conceitos de *free will*, que não apelam à noção de alma e refere duas concepções da noção de livre arbítrio.

Uma dessas concepções afirma que ter livre arbítrio significa ter a habilidade de formar (e agir nessa base) decisões racionais, informadas, quando não estamos sujeitos a força indevida. Esta é a definição modesta – *modest free will*.

Outra concepção afirma que o livre arbítrio requer aquilo a que Alfred Mele designa de *deep openness* (forte abertura). Nesta concepção, devemos adicionar à concepção modesta algo importante: os agentes livres têm em aberto, decisões alternativas compatíveis com

²⁴¹ MELE, *Free Will and Luck*, op. cit., p.170.

²⁴² MELE, *Free Will and Luck*, op. cit., pp. 164-170.

²⁴³ MELE, *Free Will and Luck*, op. cit., p. 196.

²⁴⁴ MELE, Alfred, *Free. Why science hasn't disproved Free Will*, Oxford University Press, New York 2014, p. 83.

tudo aquilo que já aconteceu e com as leis da natureza. E o resultado desta conceção fica designado por livre arbítrio ambicioso – *ambitious free will*.

De forma a ilustrar a forte abertura, Alfred Mele relata a situação de uma amiga que é frequentemente convidada para ser voluntária num evento de angariação de fundos. Por vezes a sua resposta é sim, outras vezes é não. Nas ocasiões em que acede ao pedido, vários fatores relevantes podem estar em jogo, tais como o seu nível de ocupação, o nível de interesse na causa do evento, etc.

Numa determinada ocasião, decidiu aceitar o convite. Mas, consideremos que o tempo volta atrás, até ao momento imediatamente anterior em que decidiu dizer sim. Apesar de tudo permanecer exatamente igual, desta vez decide dizer não.

Este é o quadro que permite retratar a noção de forte abertura – *deep openness*. Se ela tivesse uma forte abertura, no momento da decisão e se o tempo voltasse atrás uma e outra vez, ela tomaria diferentes decisões, em cada repetição.

Será que existe o livre arbítrio ambicioso? Quando estamos inseguros acerca do que fazer e a refletir sobre isso, sentimos que está em aberto para nós, mais do que uma decisão. Mas como seria sentir que as decisões alternativas não estão abertas, da forma forte referente ao livre arbítrio ambicioso? Como seria sentir, se fosse necessariamente verdade que, dadas as leis da natureza e a condição do nosso universo num tempo passado, nós tomaríamos todas as decisões da forma que fizemos até então, exatamente da mesma maneira e tendo os mesmos sentimentos que tivemos naqueles momentos?

Para Alfred Mele a resposta é: da mesma forma como normalmente sentimos, dessa mesma forma sentimos agora. Ora esta afirmação não significa que a forte abertura não existe. Significa sim, que a diferença entre a forte abertura e a sua ausência, não se trata de algo que pode ser sentido – *can be felt*. Não saber o que vamos decidir é uma coisa. Ter forte abertura relativamente ao que vamos decidir é outra. E descobrir se essa forte abertura existe não pode ser uma questão de ver como sentimos as coisas.

Portanto a resposta de Alfred Mele à questão: o livre arbítrio existe? A resposta é sim, se nos referirmos ao livre arbítrio modesto. Se nos referirmos ao livre arbítrio ambicioso então a questão ainda fica em aberto²⁴⁵.

²⁴⁵ MELE, *Free. Why science hasn't disproved Free Will*, op. cit., pp.78-91.

Considerações finais

«*The occurrence of Strict Akratic Actions seems to be an unfortunate fact of life*²⁴⁶». O mesmo parece ser verdade em relação ao auto engano. Compreender as bases do comportamento irracional permite vislumbrar as suas implicações no comportamento e pensamento racional.

Todos nós somos sujeitos a atitudes racionais e irracionais, dependendo das nossas predisposições psicológicas – tal como nos exemplificou Alfred Mele quando se refere ao sujeito ciumento – e poderemos adotar atitudes acráticas ou de auto engano num determinado contexto, dimensão ou fase da vida, sem que isso nos classifique como sujeito irracional.

A este propósito, Alfred Mele refere que um agente humano é um ser humano que age. E entre as nossas ações intencionais parecem estar ações racionais e irracionais, ações morais próprias e impróprias, ações nas quais o nosso envolvimento, enquanto agentes, é impressionante e ainda ações simples e rotineiras. Existem ações mentais – por exemplo resolver um problema de xadrez – e ações intencionais óbvias, que envolvem essencialmente movimento corporal periférico. Uma teoria da agência humana deverá acomodar ações de todos estes tipos²⁴⁷.

Os fenómenos do auto engano e da acrasia são melhor entendidos perante a explicação das nossas ações intencionais, o poder do raciocínio prático e juízos avaliativos práticos, racionalidade humana e livre arbítrio.

Considerando que a acrasia e o auto engano são ambos fenómenos de irracionalidade motivada, parece não ser possível deixar de relacionar ambos os fenómenos entre si, como aliás espero ter conseguido sugerir ao longo do texto. Com efeito, não pode ser ignorado o facto de que a nossa interpretação da realidade pode ser deturpada pelos nossos afetos, ainda mais se a realidade em questão, afeta os nossos receios ou desejos. Refiro-me aqui a um viés cognitivo, pois, se nossas crenças afetam os nossos juízos avaliativos, em função dos quais deliberamos, logo, se essas crenças são formadas irracionalmente, ou são deturpadas pelos nossos afetos, muito facilmente poderemos cair em ações

²⁴⁶ MELE, *Autonomous Agents*, op. cit., p. 16.

²⁴⁷ MELE, *Motivation and Agency*, op. cit., p. 5.

consideradas acráticas. E neste caso, o agente acrata não estaria a agir contra o seu melhor juízo, consideradas todas as coisas, conforme defende Donald Davidson, mas estaria a agir em função de um melhor juízo enviesado.

Neste raciocínio, também Vasco Correia defende que explicar o auto engano, permite elucidar o fenómeno da acrasia²⁴⁸. Porque ambos os fenómenos emergem da irracionalidade motivada, ou seja, os juízos são formados sob influência de uma emoção. No auto engano, esta ilusão provem de uma forte emoção, que atua nas faculdades cognitivas do agente e no caso da acrasia, a emoção ou desejo fortes, atuam no juízo avaliativo do agente²⁴⁹.

Assim, se a irracionalidade cognitiva pode afetar os juízos avaliativos do agente, que levam o agente a deliberar, fica determinado o problema do efeito desses juízos irracionais sobre a racionalidade prática²⁵⁰. É por este motivo que Vasco Correia também se refere à acrasia como sendo uma fraqueza de juízo e não uma fraqueza de vontade²⁵¹. Defende ainda que é mais plausível referirmo-nos a ações racionais ou irracionais do que sujeitos racionais ou irracionais. Isto porque, os agentes racionais e irracionais somos todos nós, na medida em que todos desempenhamos ações mais ou menos racionais e ações mais ou menos irracionais. O foco da racionalidade deve ser colocado nas ações e não nas pessoas²⁵². Esta afirmação pode ser considerada congruente com o entendimento de Alfred Mele, quando afirma que o desempenho de uma ação acrática não rotula esse sujeito de irracional.

Vasco Correia refere-se à conceção motivacional, segundo a qual os desejos e emoções do sujeito irracional não se constituem como razões para a adoção de crenças intencionais, mas sim, como as causas irrefletidas das crenças ilusórias adquiridas pelo agente. E considera que esta hipótese é fundamentada em estudos empíricos da psicologia social, que demonstram que as emoções enviesam os raciocínios e as crenças do sujeito, ao influenciarem os processos cognitivos, que formam o nosso juízo.

²⁴⁸ CADILHA, S, MIGUENS, S., (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op.cit. p. 165.

²⁴⁹ CADILHA, S, MIGUENS, S., (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op.cit. p. 186.

²⁵⁰ CORREIA, «Os limites da racionalidade: auto-engano e acrasia», art. cit., p. 287.

²⁵¹ CORREIA, «Os limites da racionalidade: auto-engano e acrasia», art. cit. p. 289.

²⁵² CADILHA, S, MIGUENS, S., (coord). *Ação e Ética. Conversas sobre racionalidade prática*, op.cit., pp. 170-171.

Ou seja, as emoções afetam a direção e intensidade da atenção que dispensamos à informação disponível, viciam a heurística na recolha de informação, deturpam a nossa interpretação dos dados e originam o fenómeno de memória seletiva²⁵³.

Cada um de nós possui uma representação do mundo em que vive e no qual age. Esta ação dependerá de vários fatores internos e externos, que além de existirem em simultâneo, relacionam-se complexamente. A explicação e compreensão das ações de cada agente, racionais ou irracionais, terá que ter em conta duas dimensões de igual importância: a dimensão interna e a dimensão externa do agente.

Conforme defende Alfred Mele, as ações só podem ser entendidas com base na consideração da predisposição psicológica de cada indivíduo, os acontecimentos passados que afetam o nosso *self* presente e a representação das condições nas quais podemos ou temos que agir.

No que diz respeito à dimensão interna do agente, parece relevante considerar que o nosso *self* presente é exclusivo de cada agente e apenas conhecido por si próprio. Trata-se de *what is it like to be me*, recorrendo à famosa expressão de Thomas Nagel, ou do eu auto biográfico que António Damásio diz, englobar as histórias e experiências de que fizemos parte, ou não, bem como recordações das experiências emocionais, sendo, portanto, uma autobiografia consciente²⁵⁴. Parece legítimo supor que Alfred Mele partilha desta noção, quando descreve o modelo DSL.

Portanto, considerar este contexto psicológico é fundamental para a compreensão da ação, dado que confere uma predisposição idiossincrática para interpretar, representar, deliberar, intencionar, decidir e agir. Parece ser nesta linha que o modelo DSL refere a probabilidade das ações futuras serem influenciadas pelas ações presentes e passadas e que o passado influencia as probabilidades de ação presentes.

Mas, por outro lado, a representação é orientada e ocorre no contexto em que o sujeito se encontra, contexto esse que afeta de diferente forma cada agente.

Nesta linha parece razoável entender que a perspetiva de Alfred Mele se reveste de elementos externos, de cariz contextual, que pressupõem a conjugação, num momento *t*,

²⁵³ CORREIA, «Os limites da racionalidade: auto-engano e acrasia», art. cit., pp. 284-285.

²⁵⁴ DAMÁSIO, *O Livro da Consciência*, op. cit., p. 263.

de determinado conjunto de circunstâncias que passam a constituir o conhecimento do agente, a partir do qual vai construir a sua intenção para agir, em articulação com as suas características internas.

Esta linha de pensamento parece remeter-nos para o pensamento de Michael Bratman, que defende que as atitudes cognitivas, que guiam o raciocínio prático e a ação, vão além das crenças do agente. O agente planifica a sua ação em coerência e consistência com as crenças internas, as quais são independentes do contexto, mas existe também uma atitude cognitiva que se reveste de uma *context-relativity*. As pressões práticas, inerentes a um contexto, levam-nos à aceitação de uma determinada proposição, no quadro da nossa deliberação, ou seja, o agente exhibe uma aceitação relativa a um contexto, facto que é distinto das suas crenças²⁵⁵.

E este quadro parece-me muito próximo de Alfred Mele, quando se refere à força motivacional dos nossos desejos e as nossas avaliações dos objetos dos nossos desejos, como sendo dois elementos que podem, ou não, estar em concertação.

Ou seja, segundo Michael Bratman, posso acreditar que *p*, mas não aceitar *p* num determinado contexto prático. A aceitação não é uma crença. O agente tem um contexto cognitivo independente do contexto externo, que surge por defeito, tal como o caminho *by default* de que nos fala Alfred Mele, mas esse contexto cognitivo, que não muda, permite ajustamentos em função do contexto prático²⁵⁶. Neste ajustamento, o agente ao experienciar um conflito interno, vai fazer o balanço do peso das opções e agir, sendo que essa ação pode ser, ou não, contrária ao seu melhor julgamento.

Michael Bratman distingue decisão e intenção. As decisões, que resultam da deliberação entre opções, envolvem uma relatividade contextual: estas decisões devem ser avaliadas relativamente a um contexto específico de aceitações. E determinadas aceitações do agente, são peculiares ao contexto específico, no qual a decisão é tomada. As intenções, em contraste, são independentes do contexto, paralelamente à independência de contexto das crenças. Assim, no raciocínio prático, existe um par dependente do contexto: decisão e aceitação e um par independente do contexto: intenção e crença. A aceitação e a crença

²⁵⁵ BRATMAN, Michael «Practical Reasoning and Acceptance in a Context», *Mind*, New Series, Vol. 101, Nº 401 (Jan. 1992) pp. 5-6.

²⁵⁶ BRATMAN, «Practical Reasoning and Acceptance in a Context», *art. cit.*, pp. 9-10.

são elementos cognitivos e a decisão e intenção são elementos práticos. Cada par desempenha um papel diferente mas inter-relacionado na planificação de cada agente²⁵⁷. Portanto, de novo sublinho que conhecer o contexto no qual o agente se encontra no momento em que delibera e age, parece-me ser indispensável à compreensão da ação do agente.

Apesar da importância dos fatores externos na explicação da ação acrática, em conjugação com a capacidade de auto controlo do agente, a representação dos objetos, assumida como sendo interna ao agente, parece influenciar a percepção do quadro contextual de uma determinada situação, do ponto de vista de cada agente. Ainda que a motivação para *a* seja significativa, a forma de interpretação e representação dos objetos e dos resultados da ação para *a*, pode marcar a ação acrática, bem como a formação e retenção de crenças enviesadas.

Considerando que somos agentes sociais, o fator contextual parece revestir-se de igual importância, tanto na planificação da vida do sujeito, numa dimensão individual mas também social. Aliás, a psicologia social defende e apresenta estudos empíricos que demonstram a influência social no comportamento e na decisão individual. Por exemplo, estudos empíricos demonstram que existem dois tipos de expectativas que podem prever o nosso próprio comportamento: aquilo que esperamos que os outros façam (expectativas empíricas) e aquilo que pensamos que os outros pensam que nós devíamos fazer (expectativas normativas). Se estes tipos de expectativas entram em conflito, as pessoas farão aquilo que acham os outros também fariam, na mesma situação²⁵⁸. O psicólogo Herbert Kelman defende a existência de três processos de alteração de atitudes devido à influência social: conformidade (o sujeito aceita a influência, não porque acredita no seu conteúdo, mas porque espera alcançar uma reação favorável de outras pessoa ou grupo), identificação (o sujeito aceita a influência, porque espera estabelecer ou manter uma relação satisfatória com outra pessoa ou grupo. Este comportamento está associado à relação desejada) e internalização (o sujeito executa o comportamento influenciado

²⁵⁷ BRATMAN, «Practical Reasoning and Acceptance in a Context», art. cit., p. 13.

²⁵⁸ BICCHIERI, Cristina e XIAO, Erte. «Do the right thing: but only if others do so», *Journal of Behavioral Decision Making*, 22 (2009) 191–208.

devido ao conteúdo do comportamento induzido, que está de acordo com o seu sistema de valores. Ou seja, este comportamento está integrado na sua individualidade)²⁵⁹.

Portanto, considerando que os agentes estão integrados e agem socialmente, parece pertinente aludir de novo a Michael Bratman, que presta atenção ao papel complexo das intenções e planificação na agência humana. Intencionar, não é apenas intencionar agir, é também planificar, ou seja, estabelecer intenções estáveis e direcionadas para o futuro.

Por outro lado, o fenómeno de intencionar agir, não é apenas o resultado do raciocínio prático: é também um elemento fundamental de *input* de informação dos agentes. Compreender as intenções e os planos, é pois essencial para compreender a natureza da agência e uma chave para caracterizar a diferença crucial entre raciocínio prático, relacionado com o que fazer, e raciocínio teórico, relacionado com o que acreditar.

O modelo BDI – *Belief-Desire-Intention*, desenvolvido por Michael Bratman, é uma proposta acerca da natureza da racionalidade prática, centrado no fenómeno de intencionar agir. O modelo separa duas capacidades distintas: a capacidade de selecionar um plano e a capacidade de implementação de planos. De acordo com o BDI, o agente é capaz de fazer o balanço entre o tempo que gasta na deliberação acerca dos planos e a implementação dos planos²⁶⁰. As capacidades de planeamento básico, suportam a organização temporal da nossa agência e enraízam duas capacidades práticas humanas fundamentais: as nossas capacidades de auto governação e formas importantes de sociabilidade²⁶¹.

Michael Bratman explora a agência humana na forma em que esta se relaciona com auto determinação, auto governação e autonomia. Para o autor, as questões centrais à Filosofia da Ação, dizem respeito às estruturas básicas da agência humana e alargou o seu trabalho a partir de questões relativas à natureza dos planos e intenção para um fenómeno específico: intenção partilhada e agência partilhada.

²⁵⁹ KELMAN, Herbert. «Compliance, identification, and internalization: Three processes of attitude change» *Journal of Conflict Resolution*, 2 (1) (1958) 51-60.

²⁶⁰ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 10-11.

²⁶¹ MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., p. 91.

Segundo Michael Bratman e ao contrário do que pensam os incompatibilistas, a vontade não é algo alheio à ordem causal: criaturas com vontade – *will* – são criaturas com certas estruturas psicológicas, nomeadamente capacidade de impor uma estrutura no seu pensamento e ação. A questão é entender essas estruturas e o autor, entende ser crucial compreender a natureza da planificação, bem como as normas associadas. O ato de escolher, ou decidir, são formas que usualmente usamos para delinear os planos nas nossas mentes.

As intenções envolvem compromisso e é esse compromisso que torna possível a persistência dos planos, e que torna possíveis novos planos, baseados naquilo em que o agente já está comprometido. Por outro lado, as intenções, ao contrário dos desejos, estão sujeitos a requisitos de racionalidade. Em conjunto com as crenças, as intenções têm a função de criar um modelo consistente do futuro do agente, para esse agente²⁶².

Os seres humanos são não apenas agentes com propósitos, mas também agentes capazes de reflexão acerca da sua motivação. Formamos planos e políticas prévias, organizadoras das nossas atividades ao longo do tempo e vemos-nos como agentes que persistem ao longo do tempo, que iniciam, desenvolvem e completam atividades e projetos que se prolongam temporalmente. Assim, qualquer teoria razoável acerca da ação humana terá que tem em conta estas três vertentes: a nossa capacidade de reflexão, planificação e conceção de agência temporalmente estendida²⁶³.

As políticas de auto organização são compromissos que, em contexto apropriado, podem ajudar a determinar de que forma se posiciona um agente perante determinada motivação. Com isto, não pretende dizer que estas políticas não são passíveis de revisão racional, porque o sujeito pode, por vezes, reanalisar reflexivamente a sua posição²⁶⁴.

Os planos, intenções e políticas, oferecem enquadramentos que estruturam a deliberação e o raciocínio prático, enquadramentos esses que têm uma estabilidade característica ao longo do tempo. Moldam o raciocínio prático, por vezes ao estabelecer quais as opções a serem consideradas, outras vezes ao estabelecer o que deve ser considerado como

²⁶² MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*, op. cit., pp. 11-13.

²⁶³ BRATMAN, Michael. *Structures of Agency*. Oxford University Press. New York, 2007, p. 21.

²⁶⁴ BRATMAN, *Structures of Agency*, op. cit., p. 36.

favorecendo ou desfavorecendo opções consideradas. Nestes enquadramentos, as nossas intenções, planos e políticas, ajudam a constituir e suportar formas intertemporais de coordenação e organização, quer ao nível da vida individual de cada agente, quer ao nível de vários agentes. Estas formas de organização, intrapessoais e interpessoais, tornam-se centrais à nossa capacidade de atingir objetivos complexos e temporalmente estendidos. Fica delineado um modelo de uma importante variedade de fenómenos sociais que, devido ao seu papel, podem ser plausivelmente entendidos como formas partilhadas de valorização²⁶⁵.

Espero não ter sido excessiva a alusão a Michael Bratman, mas de facto a dimensão social de cada agente, não poderá ser desconsiderada no que respeita à agência humana, racional ou irracional.

Por outro lado, estando em causa agentes, que desempenham ações no mundo, parece tornar-se impossível não considerar a responsabilidade de cada agente, em relação às ações que produz, sejam elas racionais ou irracionais. E, segundo a perspetiva de Alfred Mele, o auto controlo será necessário, para mediar essas ações motivadas.

Parece-me que é numa linha de complementaridade com o auto controlo, que é trabalhado o conceito de mecanismo, levado em conta por Jon Elster ou Vasco Correia, mecanismos estes que funcionam como estratégias cognitivas mediadoras e /ou controladoras da ação. A compreensão dos mecanismos envolvidos na ação, bem como das estratégias cognitivas, podem desempenhar um papel fundamental, não só na compreensão das ações irracionais, mas também na compreensão da integração cognitiva de modos de ação racionais, que permitem a agência racional.

A noção de mecanismo parece tornar-se central em duas dimensões; para a compreensão da irracionalidade motivada, de forma a compreender como é possível o desempenho da ação acrática ou a ocorrência do auto engano, mas também como uma forma de equilíbrio entre o plano prático (ações do agente) e o plano cognitivo (formação e retenção de crenças). Pois será através desses mecanismos, que o agente poderá adquirir métodos de raciocínio e auto controlo, favorecedoras da ação e pensamento racional.

²⁶⁵ BRATMAN, *Structures of Agency*, op. cit., pp. 283-286.

Parece ainda de ressaltar que, no caso particular de agência irracional, a análise da temática não precisa de sustentar-se num apriorismo, pelo qual cada teórico constrói a sua argumentação, segundo a sua própria intuição Filosófica.

E, de facto, parece ser nessa linha que Alfred Mele sustenta o seu pensamento. O seu livro *Free. Why Science hasn't disproved Free Will*, debruça-se precisamente na análise de experiências científicas acerca da ação livre. Nessa obra descreve as experiências de forma entendível, e analisa cada uma delas, para corroborar, argumentativamente, a sua tese de que nenhuma das experiências em análise, coloca de fora a existência do livre arbítrio.

Por outro lado, parece frutífera a integração de uma visão interdisciplinar nas análises Filosóficas, método a que Alfred Mele recorre, ao analisar experiências levadas a cabo por psicólogos e analisadas ao longo da sua obra. Aliás, no âmbito dos seus trabalhos do projeto *The Big Question in Free Will*, Alfred Mele edita um livro de carácter interdisciplinar. Conforme sugere o próprio título: *Surrounding Free Will. Philosophy. Psychology. Neuroscience*, esta obra, analisa o livre arbítrio a partir de várias perspetivas: neurociência, psicologia social, cognitiva e do desenvolvimento e filosofia tradicional e experimental.

Cada área do saber, aporta conceitos e abordagens que permitem uma reflexão e conclusões sustentadas, concertadas e complementares, acerca de uma mesma temática, permitindo a sua compreensão, como um todo.

Com este trabalho e a partir da perspectiva de Alfred Mele, espero ter conseguido estabelecer uma relação entre os fenómenos da acrasia e do autoengano, considerando que ambos os fenómenos derivam da irracionalidade motivada, sugerindo que a acrasia, vista como irracionalidade prática que se revela nas ações, poderá derivar do auto engano, visto como irracionalidade cognitiva, responsável pela formação de juízos enviesados, que o agente considera nas suas deliberações. De facto, a perspectiva motivacional sugere uma influência dos desejos e emoções na faceta cognitiva, gerando o auto engano e, esses mesmos desejos, afetam os nossos juízos avaliativos, através dos quais escolhemos entre opções que nos parecem fiáveis.

Considerar um agente que, atuando de forma não constrangida, seja dotado de desejos, vontades e mecanismos de motivação e auto controle, permite desvelar a possibilidade da autonomia e responsabilidade moral pelas ações desse agente.

A dimensão do auto controle, relaciona-se com a autonomia individual e centra-se na compreensão dos mecanismos envolvidos na ação, para compreender as ações racionais e irracionais. Através do auto controle, o agente pode intervir no seu próprio processo motivacional, gerando o balanço de motivação favorável à ação continente. Identificar mecanismos capazes de aprimorar os métodos de raciocínio, podem colocar cada agente no caminho da racionalidade.

Salientando os fatores contextuais, a dimensão de auto controle permite que o agente possa ajustar o seu comportamento a contextos práticos, mesmo na presença de um contexto cognitivo independente do contexto externo.

Assim, espero ter conseguido estabelecer a relação entre irracionalidade motivada e agência, através da lente de Alfred Mele, bem como ter aberto o caminho para a possibilidade de sustentação e alicerce do comportamento, baseado na racionalidade do sujeito, considerando um contexto de livre arbítrio e responsabilidade moral.

Referências bibliográficas

AINSLIE, George, *Break-Down of Will*, Cambridge University Press, Cambridge 2001.

American Psychiatric Association, *DSM-5: Manual de Diagnóstico e Estatística das Perturbações Mentais*, (5ª Ed.), Climepsi Editores, Lisboa 2014.

BICCHIERI, Cristina e XIAO, Erte. «Do the right thing: but only if others do so», *Journal of Behavioral Decision Making*, 22 (2009) 191–208.

BRATMAN, Michael, *Structures of Agency*. Oxford University Press. New York, 2007.

BRATMAN, Michael, «Practical Reasoning and Acceptance in a Context», *Mind*, New Series, Vol. 101, No. 401 (Jan., 1992) 1-15. DOI: 10.1093/mind/101.401.1.

CADILHA, S, MIGUENS, (coord). *Ação e Ética. Conversas sobre racionalidade prática*. Edições Colibri 2011.

CORREIA, Vasco, «Os limites da racionalidade: auto-engano e acrasia» *Disputatio*, Vol. III, No. 28 (2010).

CORREIA, Vasco. «Contextual Debiasing and Critical Thinking: Reasons for Optimism» *Springer Science + Business Media Dordrecht*, (2016).

DAMÁSIO, António, *O Sentimento de Si*, Publicações Europa-América, Lda., Mira-Sintra 2004.

DAMÁSIO, António, *O Livro da Consciência*, Temas e Debates 2010.

DAVIDSON, Donald, *Essays on Actions and Events*, Oxford University Press Inc., New York 2001.

ELSTER, Jon. *Alchemies of the Mind. Rationality and the Emotions*. Cambridge University Press, United Kingdom 1999.

ELSTER, Jon. *Sour Grapes. Studies in the Subversion of Rationality*. Cambridge University Press, United Kingdom 1983.

ELSTER, Jon. *Ulysses Unbound*. Cambridge University Press, United Kingdom 2000.

HARE, Richard, *Freedom and Reason*, Oxford University Press, New York 1965.

KELMAN, Herbert. «Compliance, identification, and internalization: Three processes of attitude change» *Journal of Conflict Resolution*, 2 (1) (1958) 51-60.

KRAUT, Richard, «Aristotle's Ethics», *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/sum2017/entries/aristotle-ethics/>](https://plato.stanford.edu/archives/sum2017/entries/aristotle-ethics/).

MAURO, C, MIGUENS, S. CADILHA, S. Ed. *Conversations on Human Action and Practical Rationality*. Cambridge Scholars Publishing 2013.

MELE, Alfred, *Aspects of Agency. Decisions, Abilities, Explanations, and Free Will*, Oxford University Press, New York 2017.

MELE, Alfred, *Free. Why science hasn't disproved Free Will*, Oxford University Press, New York 2014.

MELE, Alfred, *Backsliding. Understanding Weakness of Will*, Oxford University Press, New York 2012.

MELE Alfred, *Effective Intentions. The Power of Concious Will*. Oxford University Press, New York 2009.

MELE, Alfred, *Free Will and Luck*, Oxford University Press, New York 2006.

MELE, Alfred, *Motivation and Agency*, Oxford University Press, New York 2003.

MELE, Alfred, *Self-Deception Unmasked*, Princeton University Press, New Jersey 2001.

MELE, Alfred, *Autonomous Agents*, Oxford University Press, New York 1995.

MELE, Alfred, *Springs of Action. Understanding Intentional Behaviour*, Oxford University Press, New York 1992.

MELE, Alfred, *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*, Oxford University Press, New York 1987.

MELE, Alfred, «Akrasia, Self-control, and Second-Order Desires», *Noûs*, 26, nº 3 (1992) 281-302. DOI: 10.2307/2215955.

MIGUENS, Sofia, *Compreender a Mente e o Conhecimento*. Mlag Discussion Papers, Porto 2009.

STROUD, Sarah, «Weakness of Will», *The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2014/entries/weakness-will/>>.