

UMA CONCEPÇÃO DEFLACIONISTA DA RACIONALIDADE NA ACÇÃO

Carlos Eduardo Evangelisti Mauro

TESE DE DOUTORAMENTO EM FILOSOFIA

AGOSTO, 2009

FACULDADE DE LETRAS DA UNIVERSIDADE DO PORTO

Tese apresentada para cumprimento dos requisitos necessários à obtenção do
grau de Doutor em Filosofia, realizada sob a orientação científica de

Sofia Gabriela Assis de Moraes Miguens Travis

Apoio financeiro da FCT, com a bolsa de doutoramento (2006-2009) SFRH / BD / 23393 / 2005

Apoio financeiro do Programa ALBAN, com a bolsa de doutoramento (2004/2005) E04D047286BR

Para Camila, Clara, Pedro, Rosa Maria e Godofredo

AGRADECIMENTOS

Esta é uma parte difícil de escrever, pois são muitas as pessoas que, de alguma forma, contribuíram para que eu pudesse escrever esta tese. De entre algumas formas possíveis, decidi fazer alguns agradecimentos especiais que dizem respeito ao trabalho mais directo, ligado à própria tese.

Começo por agradecer à minha orientadora Professora Doutora Sofia Miguens Travis, não só pela orientação específica, mas também pela amizade que caracterizou todo o processo de orientação. Sem a sua orientação paciente, não teria chegado ao final desta tese. Agradeço profundamente pela confiança e pela lealdade que pautou a nossa relação durante estes últimos anos. Gostaria também de agradecer formalmente às várias oportunidades académicas que me foram oferecidas durante o doutoramento – em quase todos os casos a Professora Sofia estava envolvida. Portanto, deixo aqui o meu profundo agradecimento à orientadora Professora Doutora Sofia Miguens Travis.

Quero agradecer especialmente ao Professor Doutor João Alberto Pinto, que paciente e organizadamente, como lhe é peculiar, me orientou em muitos momentos da tese. Agradeço profundamente a generosidade com que se dedicou à leitura e ao debate da minha tese.

Agradeço à Susana Cadilha, minha companheira em alguns artigos e em alguns projectos científicos nos últimos quatro anos. A Susana é a companheira de

trabalho que muitos querem, mas que somente alguns sortudos como eu têm o privilégio. Agradeço as suas sempre produtivas e competentes observações.

Agradeço à Professora Doutora Maria Clara Dias pelo interesse em meu trabalho e pelo sempre produtivo e estimulante debate acerca dos temas desta tese. Agradeço também ao Professor Doutor José Maria R. Ramos pela presença sempre importante na minha trajectória académica desde há, pelo menos, 15 anos. Agradeço ao Professor Doutor Joshua Knobe pelo estímulo ao meu trabalho na filosofia experimental e pelos comentários que fez acerca da minha abordagem à acrasia, que, de forma directa, me levaram ao item 3.4.2 desta tese. Agradeço, também, ao Professor Doutor Henrique Vailati pela amizade e pelo carinho que me dispensou em todos estes anos, como amigo e como director da faculdade onde eu leccionei entre os anos de 1999 e 2004.

Agradeço à Ana Gonzalez do Departamento de Filosofia pelo apoio e pela competência com que sempre geriu as questões que me diziam respeito dentro do Departamento. Agradeço também à Daniela Oliveira do Instituto de Filosofia pelo apoio geral que me tem dado – tem sido essencial em muitos dos projectos em que estou envolvido neste Instituto.

Agradeço ao Departamento de Filosofia da Faculdade de Letras da Universidade do Porto e ao Instituto de Filosofia por me terem dado todas as condições necessárias para escrever a tese. Agradeço à biblioteca da FLUP por me ter acolhido durante estes anos e especialmente ao seu director, o Dr. João Emanuel Cabral Leite.

Agradeço a todas as pessoas dos serviços da FLUP que de alguma forma foram importantes nestes anos, especialmente à Rafaela (biblioteca), Carla Amaral, Carla Augusto, ao Sr. Júlio e às colaboradoras do GuiPaul Bar.

Agora, coloco aqui uma lista de todas as pessoas e instituições que de algum modo me ajudaram com a amizade, com comentários, com recursos financeiros, enfim, a todos estes sou muito grato. Se não faço comentários específicos para cada uma destas pessoas e instituições é porque se o fizesse teria que incorporar, pelo menos, mais 50 páginas de agradecimentos e mesmo assim correria o risco de escrever mais, ou menos, acerca de algumas destas pessoas. Portanto, esta lista refere-se a pessoas e instituições importantes e que contribuíram, de alguma forma, para esta tese.

Pessoas (em ordem alfabética): Alberto Ferreira, Álvaro Miguens, António C. Ferreira, António Leiria, Arminda Sousa, Beatriz H. M. Mauro, Bento Dias, Camila E. M. Ferreira, Cecília Lemann, Charles Travis, Clarissa Flueeler, Cláudia Belmonte, Cristina Miguens, Desirrée Turchetto, Diana Miguens, Filipa Mendes, Frederico Mauro, Iracema Salgado, Isabel Almeida, José Antônio Mauro, José Carlos Rodrigues, José E. M. Mauro, José Meirinhos, Luís Alberto Machado, Luísa Oliveira, Manuela Miguens, Manuela Teles, Mariana Lemann, Marília Severino, Mônica Mauro, Oliver Flueeler, Quinzé, Regina H. E. M. Ferreira, Reinaldo (Reimar), Rosilda Portas, Roxo de Almeida, Rui Vieira da Cunha, Teresa Torres, Thiago Mauro, Tomás M. Carneiro, Urs Flueeler, Zé(s) Martins. Instituições: Primeiros Passos, UPorto, FLUP, Departamento de Filosofia, Instituto de Filosofia, MLAG, FCT, FLAD, FAAP, Programa Alban, FAPESP, CAPES.

RESUMO

ABSTRACT

UMA CONCEPÇÃO DEFLACIONISTA DA RACIONALIDADE NA ACÇÃO

A DEFLATIONARY CONCEPTION OF RATIONALITY IN ACTION

Carlos Mauro

Nesta tese, propõe-se e defende-se um modelo funcionalista da racionalidade na acção. Neste modelo, a racionalidade tem a função de produzir a acção através de estados mentais como crenças, desejos ou intenções. Para isto, defende-se: (i) uma abordagem realista intencional; (ii) a necessidade de a filosofia da acção e a teoria da escolha racional utilizarem a psicologia popular como fonte/base principal de suas abordagens teóricas e aplicadas. Por fim, a tese procura produzir algum diálogo entre a teoria da escolha racional e a teoria da acção da filosofia, argumentando, por um lado, que ambas têm a mesma base psicológica popular e, por outro lado, que uma explicação da acção sedimentada sobre a noção de escala de preferências (caso da teoria da escolha racional) pode ser discutida e objectada em termos filosóficos.

In this thesis, it is proposed and defended a functionalist model of rationality in action. In this model, rationality is responsible for producing the action through mental states such as beliefs, desires or intentions. Thus, it is defended: (i) an intentional realistic approach; (ii) the need that the philosophy of action and the theory of rational choice have to use folk psychology as the primary source/basis of their theoretical and applied approaches. Finally, this thesis aims at producing some level of dialogue between the theory of rational choice and the theory of philosophy of action, defending, on the one hand, that both have the same folk psychological basis and, on the other hand, that an explanation of the action established on the notion of scale of preferences (in the case of the theory of rational choice) can be discussed and refuted in philosophical terms.

ÍNDICE

INTRODUÇÃO	1
CAPÍTULO 1: ACÇÃO E RACIONALIDADE: BASES E PRESSUPOSTOS TEÓRICOS	10
1.1 QUÃO <i>FOLK</i> PODEM OU DEVEM SER OS FILÓSOFOS?	11
1.1.1 PSICOLOGIA POPULAR	11
1.1.2 PSICOLOGIA POPULAR, INTUIÇÕES E A TEORIA DA ACÇÃO (EXPLICAÇÃO E PREVISÃO)	19
1.1.3 EFICÁCIA DAS EXPLICAÇÕES DA ACÇÃO BASEADAS NA PSICOLOGIA POPULAR	26
1.2 ACÇÃO HUMANA: DO QUE ESTOU A FALAR?	32
1.3 RACIONALIDADE NA ACÇÃO (RA): DO QUE ESTOU A FALAR?	42
1.3.1 RA NA ECONOMIA	50
1.3.1.1 PREFERÊNCIAS, ANÁLISE CUSTO-BENEFÍCIO, RESTRIÇÃO ORÇAMENTAL, ESCOLHA E RACIONALIDADE	57
1.3.2 RA: DA PSICOLOGIA POPULAR PARA FILOSOFIA	80
1.3.2.1 COMO AS PESSOAS COMUNS EXPLICAM AS ACÇÕES?	80
1.3.2.2 FILOSOFIA DA ACÇÃO: COMO É QUE D. DAVIDSON EXPLICA AS ACÇÕES	86
CAPÍTULO 2: MODELO FUNCIONALISTA DA RACIONALIDADE NA ACÇÃO	94
CAPÍTULO 3: ACRASIA – UM CASO DE IRRACIONALIDADE NA ACÇÃO?	104
3.1 INTUIÇÕES COMUNS ACERCA DO CONCEITO	105
3.2 [BREVEMENTE] PLATÃO, ARISTÓTELES E D. DAVIDSON ACERCA DA POSSIBILIDADE DA ACRASIA	111
3.2.1 PLATÃO	113
3.2.2 ARISTÓTELES	118
3.2.3 D. DAVIDSON	125
3.3 OBSERVAÇÕES ACERCA DA FRAGILIDADE PSICOLÓGICA E CONCEPTUAL DA ACRASIA COMO UM FENÓMENO DA RACIONALIDADE NA ACÇÃO	135
3.3.1 JUÍZOS DE 1A. E DE 2A. ORDENS	136
3.3.2 DELIBERAÇÃO, DECISÃO E ACRASIA	143
3.4 UM EXPERIMENTO ACERCA DO CONCEITO POPULAR DE ACRASIA E AS SUAS IMPLICAÇÕES FILOSÓFICAS	150
3.4.1 FILOSOFIA EXPERIMENTAL [X-PHI]: QUESTÕES BÁSICAS	150
3.4.2 EXPERIMENTO: CONCEPÇÃO VULGAR DE ACRASIA	175
3.4.2.1 RAZÕES PARA UM EXPERIMENTO SOBRE A ACRASIA	175
3.4.2.2 EXPERIMENTO, RESULTADOS E MÉTODOS	180
3.4.2.3 AVALIAÇÃO FINAL DO EXPERIMENTO	197
OBSERVAÇÕES FINAIS	200
REFERÊNCIAS BIBLIOGRÁFICAS	204

INTRODUÇÃO

O tema deste trabalho é a racionalidade na acção, i.e. aquilo que circunda os raciocínios práticos conducentes a acções particulares de agentes reais.

Um objectivo central do trabalho é propôr um modelo da racionalidade na acção, a que se chamará 'funcionalista', querendo com isso dizer que ele não envolve à partida qualquer componente normativo-prescritiva, qualquer imperativo específico acerca do que se deve fazer (Capítulo 2 – Modelo funcionalista da racionalidade na acção).

A aplicação fundamental deste modelo da racionalidade na acção considerada no trabalho é uma interpretação da acrasia (Capítulo 3 – Acrasia - um caso de irracionalidade na acção?). Essa interpretação da acrasia está subjacente à proposta deflacionista acerca da racionalidade na acção que dá o título ao trabalho. A acrasia ou fraqueza da vontade é um tema tratado por muitos filósofos ao longo da história da filosofia: se foi Aristóteles que a nomeou assim, é possível considerar que já Platão definiu uma posição perante o problema, defendendo a ideia de que

apenas a ignorância impede os humanos de agirem como devem agir. Na filosofia contemporânea a grande referência para o tratamento da acrasia é Donald Davidson: é Davidson quem define a acção acrática como aquela em que o agente age intencionalmente contra o seu próprio melhor juízo, sendo a acrasia o protótipo de irracionalidade na acção. Se, intuitivamente, a acrasia, assim definida, parece ser algo que todos já experimentámos, pretende-se no entanto mostrar neste trabalho que a concepção de acrasia generalizadamente aceite na filosofia contemporânea, e que se procurará caracterizar, resulta de uma projecção de elementos normativo-prescritivos de racionalidade, i.e. precisamente daquilo que se pretende evitar com o modelo da racionalidade na acção aqui apresentado e defendido.

Os assuntos acima mencionados – a racionalidade na acção e a acrasia – poderiam ser tratados de uma forma puramente imanente à filosofia. Deles se ocupam hoje filósofos da acção, filósofos morais, filósofos da mente e outros. No entanto, e não apenas porque a minha formação inicial é em Economia, esta tese não pretende situar-se estritamente no campo da Filosofia mas no espaço de um diálogo entre Filosofia e Economia. Vários capítulos e sub-capítulos procuram explorar aspectos dessa relação (1.1 Quão folk podem ou devem ser os filósofos? 1.3.1 Racionalidade na Economia, Cap. 2 Modelo Funcionalista da Racionalidade na Acção, Conclusão).

A Filosofia e a Economia encontram-se desde logo no conceito de acção. É comum separar o campo de acção dos economistas do de outros cientistas sociais, como os sociólogos, chamando à acção que interessa aos economistas ‘acção racional’. Muito depende, de qualquer forma, aquilo que se faz em Economia, de

dois conceitos que estão em pleno terreno filosófico: a acção e a racionalidade. No entanto, do ponto de vista do economista, o carácter filosoficamente problemático desses conceitos não é necessariamente aparente. Ora, se aquilo que neste trabalho é defendido acerca de acrasia está certo, as presunções de racionalidade usualmente concedidas às evocações da racionalidade dos agentes em Economia terão quer ser revistas ou pelo menos mais cuidadosamente consideradas. A suposição por trás desta tese é a seguinte: uma coisa é o uso de modelos como a Teoria da Escolha Racional no trabalho científico corrente, outra coisa é a suposição de que assim estão a ser ditas coisas profundas e verdadeiras acerca de como as coisas são com agentes reais, erguendo a partir daí em 'autoridade social' a voz da Economia. A interpretação da acrasia avançada neste trabalho é um ponto de apoio para algum cepticismo quanto a tal presunção. Há muitas formas morais de viver vidas humanas, individual e colectivamente – a evocação da 'acção racional' não deve servir como pretexto para uma resposta única, uma resposta prescritivo-normativa, que será de facto uma presunção substancial de alguém, uma presunção encapotada, acerca do que se deve fazer, nomeadamente acerca do que se deve fazer em sociedades. No núcleo desta tese está a ideia de um contributo da Filosofia para a Economia com o seguinte perfil: assuntos que podem parecer nos nossos dias assuntos técnicos e de especialistas, e como tal de alguma forma circunspectos e inobjectáveis, território quase exclusivo de economistas, estão na verdade em pleno território da 'filosofia moral', i.e. envolvem decisões substantivas acerca do que deve ser feito que não podem apoiar-se num mero apelo à racionalidade. 'Filosofia Moral' foi um termo utilizado por Adam Smith para

a Economia, e foi, de resto, algo como a convicção de que há boas razões para manter em mente esta definição que me trouxe da Economia para a Filosofia.

Cabe aqui uma palavra acerca da abordagem escolhida para tratar a aplicação principal do tópico da racionalidade na acção, a acrasia. Essa abordagem foi a chamada ‘filosofia experimental’. No quadro da filosofia experimental, elaborei e conduzi uma experiência acerca de acrasia da qual pretendo dar uma interpretação filosófica (3.4 Um experimento acerca do conceito popular de acrasia e suas implicações filosóficas). Tenho consciência do carácter polémico da filosofia experimental, que é uma forma recente das relações que têm vindo a ser trabalhadas entre a filosofia e as ciências da cognição nas últimas décadas. A filosofia experimental é no entanto uma componente importante desta tese e do caso nela defendido acerca de racionalidade na acção e acrasia. No fundo desta incursão pela filosofia experimental encontra-se uma discussão complexa – mais complexa do que este trabalho procurará abarcar – acerca do que se entende por método em Filosofia. Pretendo contribuir para essa discussão mais pelo exemplo da investigação aqui apresentada do que pela tematização directa. Certamente muita coisa haverá a dizer contra o movimento recente da filosofia experimental, nascido das relações da filosofia com a ciência cognitiva. Não deixei de seguir essa controvérsia nem de tentar compreender as reticências de tantos filósofos perante a filosofia experimental: aqui procuro no entanto compreendê-la a ela própria, nas suas mais legítimas pretensões, e testar-lhe a produtividade no âmbito específico da racionalidade na acção. Esta é obviamente uma parte da tese em que recorro à minha formação em métodos empíricos e matemáticos de investigação (não posso de resto deixar de agradecer a cientistas da UP, nomeadamente Rita Gaio, da FCUP,

pelo seu auxílio no uso de métodos estatísticos sofisticados, especificamente o *modelo logit ordinal*). Tenho a convicção de que o uso de tais métodos não está de forma alguma fechado aos filósofos – a história da filosofia aí está para o provar. A questão interessante é saber se resultados assim obtidos de algum modo contribuem para a compreensão de problemas filosóficos, no caso presente os problemas da racionalidade na acção. Esta parte da tese pretendeu ser uma incursão não completamente modesta, na medida em que pretendi manter-me a par de todas as discussões acerca de métodos e experiências em curso noutros lugares, mas é uma incursão de um recém-iniciado pelo território da filosofia experimental. Tendo lançado nos últimos meses um blogue de filosofia experimental em Português, pude verificar em primeira mão a forma como o assunto excita e entusiasma estudantes de Filosofia e outros interessados.

O uso da filosofia experimental como metodologia traz-me a um outro aspecto que foi muito importante para mim neste trabalho. Pretendi ver este trabalho não apenas como um contributo da Filosofia para a Economia mas também como um contributo da Economia para a Filosofia, para que os filósofos não deixem de ter uma consciência clara de que os assuntos que se discutem em Filosofia – em particular quando discutem a natureza da acção e da racionalidade na acção – são assuntos de relevo, assuntos fundamentais e fundadores, para a problemática dos economistas. Tanto quanto a Economia é concebida como baseada sobre uma teoria da acção racional é exactamente isto que acontece. Assim, e isto justifica a inclusão de partes mais técnicas na tese (nomeadamente 1.3.1 RA na Economia), em termos metodológicos, para além do excuro pela filosofia experimental, encontra-se neste trabalho uma exposição e uma análise do quadro teórico básico

da racionalidade na acção correntemente utilizado em Economia, envolvendo questões tais como preferências, transitividade de preferências, análise custo-benefício, etc. Sem a justaposição deste quadro de uso corrente em Economia às discussões dos filósofos sobre racionalidade na acção o meu campo de problemas não ficaria caracterizado. Procurei escrever essas partes do trabalho da forma mais legível possível para pessoas de outras disciplinas – se os tópicos têm importância, eles devem ser traduzíveis.

Relativamente a este contributo de dupla direcção Economia-Filosofia e Filosofia-Economia que se joga nos problemas que interessam e, portanto, nesta tese, não posso deixar de dizer que o meu trabalho dos últimos anos no projecto *Conversations on Practical Rationality*, ao pôr-me em contacto directo com muitos dos autores que no panorama contemporâneo exploram estas relações disciplinares – por exemplo Daniel Hausman, George Ainslie, Alfred Mele, Hugh McCann, Joshua Knobe, John Broome, Michael Bratman, Michael Smith – foi extremamente importante para mim. Evidentemente é uma área específica da filosofia da economia que me interessa. Em filosofia da economia não se discute apenas questões epistemológicas acerca da natureza do conhecimento produzido pelos economistas, em termos de causalidade e explicação, questões ontológicas acerca da natureza das entidades evocadas pelos economistas ou questões acerca da possibilidade de previsão, do uso de estatística ou da natureza de regularidades – subjacente a todas as discussões está a questão da racionalidade na acção e é essa que constitui o foco dos meus interesses.

Concluo com uma recapitulação da ordenação dos Capítulos e partes da tese e com um resumo breve, que retomarei na Conclusão, das razões por que este assunto me interessa.

Começo, no Capítulo 1, por caracterizar bases e pressupostos teóricos envolvidos nas questões da acção e da racionalidade tal como elas serão aqui discutidas. Isso conduz-me a ligar a questão da acção com a *folk psychology*, ou psicologia popular, e também a identificar o problema do uso das nossas intuições comuns para pensar sobre questões relativas à natureza da acção e da racionalidade na acção (1.1, 1.1.2, 1.1.3). Procuro ainda caracterizar a questão da explicação da acção a partir da psicologia popular (1.1.3). De acordo com a psicologia popular crenças e desejos explicam acções – essa ideia é uma referência central do trabalho. Procuro em seguida (1.2) restringir o âmbito de aplicação da minha proposta central, o modelo funcionalista, à acção humana. Em 1.3. introduzo tratamentos da racionalidade na acção e da psicologia popular, tanto quanto esta envolve a explicação da acção acima identificada, provindos do exterior da Filosofia, nomeadamente da Economia e da Psicologia.

No Capítulo 2 exponho o modelo funcionalista da racionalidade na acção que pretendo defender, e que pretende, como já afirmei, não envolver à partida qualquer componente normativo-prescritiva, qualquer imperativo específico acerca do que se deve fazer. Este é um capítulo simples e que visa a descrição de um modelo geral da racionalidade na acção. É, portanto, um modelo que está em um nível teórico mais abstracto e que deverá ser sofisticado, em trabalhos futuros, a partir de casos mais específicos, a exemplo do caso da acrasia estudado no Capítulo 3.

No Capítulo 3 passo ao ponto aplicado central da tese, a questão da acrasia. Após fazer um levantamento das intuições comuns acerca do conceito, faço uma incursão pela história da Filosofia, pretendendo deixar claras as interpretações, que considero fundamentais, de Platão, Aristóteles e Donald Davidson (3.2). No ponto 3.3 discuto a fragilidade psicológica e conceptual da noção de acrasia e procuro avançar alternativas. No ponto 3.4 passo à investigação realizada no quadro da filosofia experimental, após uma breve discussão da natureza desta. Em 3.4.2.2 encontram-se todos os dados que recolhi e com que trabalhei. Concluo com uma avaliação final da experiência, a que se segue uma Conclusão geral do trabalho.

Quanto às razões por que este assunto me interessa, elas são as seguintes: o assunto da racionalidade na acção interessa-me porque concluí que não poderia ficar fechado na Economia aceitando simplesmente os pressupostos acerca de racionalidade na acção de que se partia. Tornou-se necessário olhar directamente para a natureza da racionalidade na acção. Penso no entanto que o interesse do assunto não está indexado aos meus próprios interesses pessoais: o assunto da racionalidade na acção é interessante desde logo porque todos nós vivemos em sociedades nas quais alguns economistas agem por vezes como os advogados de uma racionalidade cuja natureza não se compreende muito bem, se olharmos directamente para o agente real individual. Enquanto estudante de Economia, autores como Adam Smith, Ludwig von Mises e sobretudo Max Weber, com o seu olhar filosófico sobre os tópicos tratados, marcaram-me muito. Foi um pouco essa marca que, revestida de autores e problemas mais directamente contemporâneos, pretendi explorar nesta tese.

Espero que o que se segue possa levar pessoas de várias disciplinas a reconsiderar o que se entende por agentes racionais, acções racionais e irracionais, intenção e deliberação e vários outros tópicos ligados à racionalidade na acção. Espero sobretudo que passemos a pensar duas vezes antes de considerar certos fenómenos como fenómenos de irracionalidade na acção.

Por fim, uma observação. Utilizei, nesta tese, fragmentos dos seguintes artigos meus:

Capítulo 1: Mauro 2002, 2003, 2005, 2006a, no prelo *a*, no prelo *b*.

Capítulo 3: Mauro 2006a, 2009b, 2009d.

CAPÍTULO 1

ACÇÃO E RACIONALIDADE: BASES E PRESSUPOSTOS TEÓRICOS

1.1 QUÃO *FOLK* PODEM OU DEVEM SER OS FILÓSOFOS?

A pergunta ‘Quão folk devem ser os filósofos?’ contém em si dois temas metodologicamente importantes para este trabalho: 1º) a relação entre a psicologia popular¹ (PP), as intuições e a explicação da acção (filosofia da acção); 2º) a eficácia e a eficiência, em termos aplicados, das explicações da acção baseadas na PP.

De forma a poder tratar destes dois temas (em 1.1.2 e em 1.1.3), começo por tentar responder à seguinte pergunta básica: Sobre o que estou a falar quando me refiro à psicologia popular?

1.1.1 PSICOLOGIA POPULAR

Quando falo de psicologia popular estou a falar da capacidade que todos nós temos, enquanto agentes, para explicar e prever as nossas próprias acções, assim como as acções dos outros, evocando estados mentais como crenças e desejos. O vocabulário conceptual da PP dá forma à chamada linguagem comum. Esta, por sua vez, parece ser partilhada pelos agentes, que a utilizam, em termos práticos, para explicar e prever acções.

A PP tem sido definida como uma *estrutura conceptual* e/ou uma *rede de princípios* (muitas vezes implícitos) utilizadas pelas pessoas comuns para entender, explicar e prever os seus próprios comportamentos e estados mentais,

¹ Tradução do termo em inglês folk psychology.

assim como os das outras pessoas. Von Eckardt (2004), num texto de referência, diz que a PP, neste sentido, consiste, no mínimo, em: a) um conjunto de práticas atributivas, explicativas e preditivas; b) um conjunto de noções ou conceitos utilizados nessas práticas.

Outra caracterização padrão é a de Baker (2001), que caracteriza a PP como tendo, principalmente, os seguintes significados:

(1) Psicologia de senso comum que explica o comportamento humano em termos de crenças, desejos, intenções, expectativas, preferências, esperança, medos e assim por diante; (2) uma interpretação de tais explicações quotidianas como parte de uma teoria popular, formada por uma rede de generalizações, empregando conceitos como crença, desejo, e assim por diante.

Estes significados podem ser considerados, de certa forma, significados consensuais mínimos. Em (1), a PP está caracterizada como uma estrutura explicativa e conceptual inscrita no vocabulário popular (comum). Em (2), as explicações populares são interpretadas como parte de uma teoria popular formada por generalizações tais como ‘uma acção é causada por crenças e desejos’, ‘crenças e desejos são estados mentais do agente’.

O filósofo Shaun Nichols (2002a), de maneira bastante coloquial, tenta mostrar o alcance e a relevância quotidiana da psicologia popular na explicação e na previsão das acções, como se vê a seguir:

De maneira geral, a psicologia popular é simplesmente a informação que as pessoas comuns têm sobre a mente. Embora o âmbito da psicologia popular seja, assim, muito vasto, os debates contemporâneos da psicologia popular na filosofia e na ciência cognitiva têm-se centrado, em grande parte, na área da psicologia popular que orienta a previsão e a explicação das acções. Uma grande parte do interesse por esta parte da psicologia popular deriva do papel central que esta desempenha na nossa vida quotidiana. Previsões e

explicações “populares psicológicas” abundam nas nossas vidas, como, por exemplo, quando tentamos descobrir o que o bebé quer, ou o que os seus colegas pensam sobre o seu trabalho, ou o que a sua esposa fará se chegar tarde em casa. (...) É tão disseminado o papel da psicologia popular em nossas vidas que Jerry Fodor (1987, p. xii) observou que, caso a psicologia popular se revelasse um grande engano, seria “a maior catástrofe intelectual na história de nossa espécie.

As palavras de Fodor, citadas por Nichols, trazem à tona uma questão importante acerca da PP, o debate que envolve uma polémica entre os realistas acerca da psicologia popular e os seus críticos. Os realistas, como é o caso de Fodor, acreditam que ao explicar e prever as acções das pessoas com base em estados mentais, como crenças e desejos, pode-se assumir: a) a existência de entidades como estados mentais; b) que estes estados são as causas das acções.

Segundo Carruthers e Botterill (1999, p. 24-26), os críticos do *realismo da psicologia popular* (RPP), como, por exemplo, Paul Churchland (1988), têm em comum a rejeição da afirmação de que a psicologia popular nos compromete com a existência de estados mentais causalmente eficazes, como as crenças e os desejos.

Essa é uma questão que está necessariamente por trás do presente trabalho. O RPP constitui a referência para a abordagem do problema da explicação da acção e da racionalidade na acção. De facto, aquilo que será defendido nos capítulos 2 e 3 parte do RPP.

Fodor (1987) deixa clara a sua posição quando, no prefácio de *Psychosemantics*, escreve o seguinte:

Em sua maior parte, este livro é uma defesa da psicologia das crenças e dos desejos e está escrito principalmente para filósofos. Sem dúvida alguma, ainda que se possa conceber outras versões do Realismo Intencional, tenho a tendência em apostar numa de suas

espécies, a chamada Hipótese da Linguagem do Pensamento – ou, de uma forma alternativa, Teoria Representacional da Mente.

O realismo intencional² referido por Fodor é uma outra maneira de se referir ao RPP. Mais à frente, quando abordar a questão acerca da relação entre a psicologia popular e a filosofia da acção, voltarei a esta citação e explicarei em mais detalhes este género de realismo.

Entre outras qualidades (se é que podemos chamá-las assim), os seres humanos descrevem e explicam as suas próprias mentes e acções. Somos *self-explainers* e *self-understanders* (Scott-Kakures, 2006). Como foi acima mencionado, quando tentamos compreender e/ou explicar as nossas próprias mentes e acções, assim como as dos outros, evocamos uma linguagem que, para nós, faz sentido, de tal modo que caso contrário talvez não houvesse compreensão nem explicação. Essa linguagem comum possui um vocabulário vasto que parece ser partilhado pelos seres humanos. Seria interessante imaginarmos como seria um mundo similar ao nosso no qual as pessoas não partilhassem estes conceitos. Seria possível algum tipo de coordenação ou de cooperação? Esta pergunta é particularmente relevante se considerarmos que tal estratégia se tem mostrado extremamente importante em termos evolucionistas. Mesmo que sejam palavras diferentes em línguas naturais diferentes parece haver uma partilha massiva desses conceitos que se referem, por exemplo, aos estados mentais básicos como as crenças e os desejos.

Além de crenças e desejos, o vocabulário conceptual da psicologia popular inclui intenções e razões. Quando observamos outras pessoas aceitamos, com certa

² Sobre este tema ver: Sofia Miguens (2008) *Será que minha mente está dentro da minha cabeça? Da ciência cognitiva à filosofia*. Porto: Campo da Letras. Parte 1, capítulo2.

facilidade, que elas têm intenções e que agem motivadas por razões, sem que, em princípio, isso tenha algum significado filosófico profundo – neste nível as palavras “intenção” e “razões” não são conceitos filosóficos. As pessoas *não* dizem comumente “P fez X por uma crença Γ e por um desejo Φ ”, ou, então, “P fez X pela razão δ ”. Mas dizem, no mesmo sentido, “P fez X porque Θ ”, ou, então, “ Θ levou P a fazer X”. Descrevemos as acções dessa maneira, dando a essa descrição de senso comum o estatuto de explicação. Novamente, não tem que haver aqui só por isso algum significado filosófico importante. Quando alguém S pede a P que explique a acção X que P realizou há 5 minutos, P provavelmente dirá algo como “fiz X porque Θ ”. Isto parece-nos bastante razoável, isto é, intuitivamente parece correcto dizer que estas frases explicam a acção (num sentido vulgar de ‘explicar’ a acção). Mesmo os filósofos quando estão longe da filosofia dizem coisas como “Ana beijou Filipa porque queria conquistá-la” num sentido a que intuitivamente chamaríamos causal e explicativo e não somente descritivo. Neste exemplo, não seria de todo contra-intuitivo dizer que o desejo de conquistar Filipa, aliado à crença de que beijá-la faria isto acontecer, causou o beijo de Ana em Filipa.

Ora, passando agora da explicação comum da acção para a explicação da acção como problema, verificamos que grande parte das pessoas que se dedicam profissionalmente à explicação da acção e/ou à utilização de explicações da acção para intervenções de natureza social e/ou económica, utiliza um vocabulário, mais ou menos técnico, mais ou menos bem definido, e, de modo geral, constrói descrições com status de explicação científica que seguem, na maioria dos casos, o seguinte modelo básico: “P fez X porque Θ ”, ou previsões do tipo “P fará X porque Θ em T_n ”. São exemplos disto frases deste género: “O banco central americano

reduziu a taxa de juros porque a inflação subiu Z%", ou então, "O banco central americano vai reduzir a taxa de juros porque a inflação cairá Z% no próximo trimestre".

Mesmo num discurso científico social como é o da Teoria Económica, este vocabulário é utilizado para explicar e/ou descrever determinadas acções individuais e colectivas. Mais à frente voltarei à Teoria Económica de maneira mais técnica, mas agora é importante falar um pouco acerca do uso da linguagem mentalista comum neste discurso científico, discurso que é, em princípio, aceite com reservas, mas que goza de um "status científico superior", por exemplo, se comparado ao da filosofia da acção.

Os economistas costumam, por exemplo, utilizar expressões como: "expectativa racional" e "conjunto de preferências". A primeira é utilizada, resumidamente, para descrever e explicar o comportamento humano que é também *causado*, segundo os economistas, pelas *expectativas racionais* que alguns agentes têm acerca do raciocínio e do comportamento de outros agentes. Neste caso, as pessoas tentam prever as acções dos outros, bem como prever as reacções e as consequências das suas próprias acções. Isto faz-se através da imputação de racionalidade aos outros agentes, que, por sua vez, devem possuir crenças e desejos (preferências) que definem, em conjunto, as acções futuras. A segunda é utilizada, resumidamente, para descrever e explicar o comportamento humano através do conjunto de preferências que pode, em última instância, ser análogo a algo denominado "conjunto de desejos".

O que é aqui fundamental para os meus propósitos de análise é o seguinte. Como Daniel Hausman (2006, p. 45), um dos principais filósofos da economia contemporâneos, é categórico ao afirmar,

A Teoria da Escolha Racional, que domina a Economia, deriva de uma teoria comum da escolha humana, que tem sido chamada “psicologia popular. (...) Uma forma de aprimorar a explicação da acção da psicologia popular é substituir a noção não comparativa de desejo pela noção comparativa de preferência.

Noutras palavras, a Teoria Económica necessita absolutamente da noção comparativa de preferência para poder produzir comparações consistentes entre cabazes de bens, que, por sua vez, permitem a construção de curvas de indiferença (como já foi dito, isto será visto em maior detalhe à frente, ainda neste capítulo). Ora, essa noção de preferência não é mais do que o substituto da noção não comparativa de desejo da psicologia popular.

Em suma, a PP é inevitável para a Economia. Ela é a porta de entrada dos economistas ao tema da explicação da acção e da racionalidade na acção. Coloca-se portanto a questão de saber que tipo de integração será possível entre as formas pelas quais os economistas e os filósofos explicam a acção e concebem a racionalidade na acção, e esta questão passa necessariamente pela psicologia popular e pelo seu estatuto. Abaixo, cito mais uma vez o filósofo da economia Daniel Hausman (Mauro, no prelo) para mostrar como parece ser plausível defender a forte ligação entre a Economia e a PP.

O modelo da acção racional dos economistas está muito no espírito da psicologia popular, mas é muito mais simples. O economista começa por supor que os agentes já completaram a quase totalidade das suas deliberações e que conseguiram decidir, à luz

das suas crenças, através das suas intenções, planos e desejos, formar um ranking das alternativas, levando em consideração tudo aquilo que é relevante para a avaliação e a escolha. Se este ranking se estende aos objectos imediatos de escolha, então não há nada para o economista fazer excepto dizer que os agentes escolhem qualquer alternativa viável que está no topo dos seus rankings. Além de descrever as propriedades formais de tais classificações, não há nada a dizer sobre a deliberação ou a racionalidade. O que dá aos economistas algo a fazer e impede que o seu modelo de acção racional se torne trivial é que os economistas abordam uma parte do problema da formação das preferências. Eles supõem que as preferências sobre os resultados das escolhas são dadas, mas não as preferências sobre os objectos imediatos de escolha. Estas preferências são derivadas de preferências sobre resultados e de crenças (ou probabilidades subjectivas) sobre as consequências dos objectos imediatos de escolha. Assim, existe algum espaço para a deliberação no refinamento que os economistas fazem da psicologia popular, mas a deliberação foi reduzida ao cálculo.

No item 1.3.1, será, em termos gerais, descrita a teoria da escolha racional e esta citação voltará a ter utilidade, e será, então, melhor explicada. Por agora, quero chamar a atenção para o facto de que, como disse Hausman, os economistas supõem que os agentes deliberam à luz de suas crenças e que fazem isto através das intenções, dos planos e dos desejos. O vocabulário utilizado é claramente aquele que define a psicologia popular e que, como veremos no item 1.3.2, é também determinante na filosofia da acção.

A linguagem comum, mesmo que revestida de alguma pele científica, é aqui utilizada de maneira clara, no seu sentido mais simples. Mesmo que determinados cálculos matemáticos, ou determinados métodos estatísticos, sejam utilizados também com a intenção de dar um aspecto neutro e correcto do ponto de vista da linguagem científica *strictu senso*, são a linguagem comum e a psicologia popular que estão na base das práticas teóricas. A Teoria Económica não teria sido possível sem conceitos como preferências, expectativas, aversão, crença, gosto, desejo, etc.

1.1.2 PSICOLOGIA POPULAR, INTUIÇÕES E A TEORIA DA ACÇÃO (EXPLICAÇÃO E PREVISÃO)

Tendo definido a PP e dito algo acerca do seu alcance e importância, passo à pergunta seguinte, que corresponde ao primeiro tema desta secção (1.1): Qual é a relação entre a psicologia popular, as intuições e a explicação da acção formulada pela filosofia da acção?

Parece ser incontroverso que sem o apelo às intuições e sem a utilização da psicologia popular seria impossível produzir qualquer teoria da acção a partir da filosofia. De onde mais viriam os conceitos envolvidos? Neste caso, como noutros na filosofia, os filósofos, de um ponto de vista metodológico, não apenas utilizam as suas próprias intuições como dependem delas e das intuições que têm acerca das intuições das outras pessoas. No caso dos conceitos relativos à acção, o quadro de referência de tais intuições é a psicologia popular.

Quando o filósofo utiliza o termo *acção* está já a atribuir um sentido especial a um acontecimento físico que podemos chamar, de modo mais neutro, comportamento, isto é, está a utilizar a noção psicológica popular de acção que traz consigo a noção de intenção. A filosofia da acção, de modo geral, diferencia a acção de meros acontecimentos físicos, como é o caso de um comportamento, recorrendo à noção de intenção. Isto é, um comportamento é uma acção se e somente se puder ser descrito em termos intencionais. Para isso deve ser possível identificar, pelo menos, um desejo e uma crença relevante que, em conjunto, expliquem a acção. Dá-se a este processo o nome de racionalização (esta é a terminologia davidsoniana).

Em suma, quando fala de acção o filósofo está já, de algum modo, comprometido com um vocabulário derivado da PP e também, de algum modo, comprometido com o RPP. Se não fizer isto, passará a fazer outro tipo de investigação que pode, por exemplo, ser uma investigação neurocientífica ou behaviorista. Porém, neste caso, não estará investigando a acção e sim acontecimentos físicos.

Não é possível falar da acção sem atribuir ao comportamento (aos movimentos físicos de corpos) determinadas propriedades que podem, resumidamente, ser chamadas de mentalistas. É condição necessária para que um comportamento seja considerado uma acção o poder ser descrito de forma intencional, recorrendo-se a, pelo menos, um desejo e uma crença relevante.

Ao aceitarmos que filósofos e economistas utilizam, como porta de entrada e como meio para a formulação de intuições acerca da acção, a psicologia popular, teremos que reconhecer a necessidade que todos estes teóricos têm de ser “folk”, isto é, de se guiarem pela psicologia popular, mesmo que isso seja feito de modo não ingénuo.

Procuremos agora relacionar esta constatação com o tema deste trabalho. Este trabalho é basicamente sobre a racionalidade na acção, a própria acção, e a acrasia. Se no primeiro caso aceito que os filósofos cumprem um papel essencial e, de certa maneira, principal na delimitação ou definição da racionalidade, no segundo caso penso que é razoável procurar relacionar o trabalho do filósofo com uma rede mais abrangente de áreas do conhecimento que tratam da acção. Nesse caso, seriam importantes as ciências cognitivas, especialmente a psicologia, a neurociência cognitiva e as ciências sociais, especialmente a economia. O terceiro

caso, o caso da acrasia, é o mais curioso, pois vem sendo tratado exclusivamente pelos filósofos, nos termos definidos no capítulo 3. Alguns poderiam dizer que a psicologia e a psiquiatria tratam deste tema quando estudam a compulsão e a obsessão. Creio, no entanto, que isso é um erro, e um erro revelador, já que no caso da acrasia, pressupõe-se que o sujeito age livre e intencionalmente e nos casos da compulsão e da obsessão não é esse o caso. No entanto, de um modo geral, os filósofos da acção não consideram estes casos como casos de acrasia.

A acrasia é algo que acontece *no* agente real. O agente pode, ele próprio, dar-nos pistas sobre o que de facto pode ser o fenómeno. Ora, se isso é assim, por que devemos produzir definições e modelos explicativos da acrasia sem utilizarmos a observação directa?

Uma das hipóteses a ser testada nesse trabalho é a de que a acrasia é um problema exclusivamente moral - no âmbito de uma discussão acerca de racionalidade e irracionalidade nada pode ser dito sobre ela. Ora, a filosofia tem produzido nos últimos 40 anos uma teoria hegemónica (não única) acerca da acrasia que exclui qualquer explicação no âmbito moral (veja-se 3.2). Assim, para testar a minha hipótese procurarei afastar-me desta teoria hegemónica e investigar como os agentes reais identificam o problema da acrasia (veja-se 3.4).

Voltemos então à questão acerca de intuições, em particular intuições acerca de acção e de possíveis acções irracionais como a acrasia. Já afirmei acerca das intuições de filósofos e de não-filósofos que sem elas não parece possível fazer filosofia da acção. Mas será que intuições são absolutamente fiáveis para ajuizarmos acerca da natureza de um fenómeno, por exemplo de um fenómeno como a acrasia? Em particular, será que as intuições dos filósofos têm algum tipo

de privilégio? Os dados utilizados pelos filósofos ou vêm da ciência, ou vêm da sua própria experiência no mundo, ou podem também ser produzidos pelo próprio filósofo através de métodos experimentais (filosofia experimental – veja-se secção 3.4). Dizer imediatamente que o filósofo consegue, por exemplo, afastar-se de maneira relevante da sua natureza de agente e da sua prática de agir para explicar a acção humana em termos gerais é extremamente suspeito. À partida, os investigadores da acção humana são eles próprios agentes, com uma história complexa, intensa e variada como tal. Parece ser impossível que um filósofo fale da acção humana sem utilizar a sua própria experiência como fonte de, pelo menos, uma parte das suas intuições acerca do tema da acção.

Em suma, parece ser impossível que os filósofos, que são eles próprios agentes, não utilizem a psicologia popular, assim como parece ser muito difícil e, talvez, pouco produtivo, tentar fazer da filosofia da acção uma área de investigação sem apelos às intuições. As intuições na filosofia da acção têm como quadro de referência a psicologia popular, e quando um filósofo apela às intuições estará a utilizar, por um lado, o vocabulário “folk” e, por outro, a psicologia popular como fonte para as intuições. Por exemplo quando um filósofo trata da questão da causalidade e da responsabilidade moral e reivindica uma posição de vinculação, afirmando que “se um agente A é moralmente responsável pelo evento e, então, A realizou uma acção, ou omissão, que causou e”, ele está a utilizar claramente a psicologia popular para dar conta de um problema filosófico moral³.

³ Esta intuição, por exemplo, está em discussão na filosofia experimental. Alguns experimentos têm colocado em cheque a correcção desta intuição. Sobre isto ler: Julia Driver, Joshua Knobe, Ben Fraser e John Deigh, em Sinnott-Armstrong, 2008.

O tema da importância, necessidade e legitimidade do apelo dos filósofos às suas próprias intuições é um tema central desta tese. No entanto, é necessário clarificar a posição que será aqui assumida. Há actualmente uma disputa entre (Liao, 2008): a) aqueles que acreditam nas intuições como evidência, b) os que aceitam as intuições como evidência, desde que sejam experimentalmente testadas, c) aqueles que são contrários à utilização das intuições como evidência. Insiro-me claramente no grupo (b), no entanto, acredito que no início de um trabalho filosófico as intuições, mesmo não testadas, desempenham um papel essencial, principalmente na boa formulação do problema. Para não comprometer o fluxo do texto, remeto para a secção 3.4.1 um tratamento mais profundo deste assunto.

Recapitulando: parece difícil, para não dizer impossível, não utilizarmos, por um lado, as nossas intuições e aquelas que julgamos serem as intuições dos outros acerca da acção e da racionalidade e, por outro lado, um vocabulário, que deriva da psicologia popular, que faça sentido e que represente tais intuições. Assim, assume-se explicitamente que o conjunto de conceitos que constitui a referência deste trabalho tem origem na PP e que o nível explicativo aqui em causa aceita como explicativos os termos derivados de um discurso mentalista, pelo menos similar ao que se encontra na PP.

Quem pensa que esse mentalismo é útil e que pode, de facto, nos levar a algum conhecimento importante acerca da acção humana, terá que dar conta, mesmo que de maneira simples, de algumas objecções que são difíceis de serem respondidas (se é que podem sê-lo) não apenas pelos dilemas que colocam, mas porque também são totalmente pertinentes.

O trecho abaixo, retirado de Baum (2009, p. 105), sintetiza um feixe de críticas que normalmente está presente no debate acerca da explicação da acção através da psicologia popular, ou, pelo menos, através de seus conceitos mais importantes, como acontece no caso da filosofia da acção.

(...) representações mentais de eventos futuros são exemplos de mentalismo, vítimas de todos os problemas que discutimos (...) Onde está essa intenção interna? De que ela é feita? Como essa porta aberta fantasmagórica poderia causar o meu comportamento de abrir o trinco? Como é que uma representação interna do Moby Dick poderia causar a sua procura? Isso não é uma explicação; serve apenas para obscurecer os factos relevantes sobre o ambiente: abrir o trinco normalmente leva a uma porta aberta, e procurar uma mercadoria normalmente produz a mercadoria. Esses factos naturais explicam o comportamento sem nenhuma necessidade de introduzir intenções internas.

Não é o foco deste trabalho produzir uma resposta filosófica profunda a esta crítica; uma boa quantidade de autores e de artigos tratam de maneira competente deste tema – Fodor (1987), por exemplo. No entanto, é importante reiterar que tudo o que se vai fazer aqui parte do pressuposto de que coisas como intenções, crenças, e desejos são estados mentais que, num determinado sentido, explicam a acção. Se assim não fosse, não se poderia presumir, como aqui faço, que podemos, nesse nível explicativo, construir bons modelos, modelos que possam ser utilizados na explicação de acções passadas e na previsão de acções futuras.

Os eliminativistas, de modo geral, supõem que se um dia o conhecimento acerca do funcionamento do cérebro for de tal modo completo que possamos compreender e prever o que compreendemos e prevemos com a utilização da psicologia popular, poderemos prescindir dos conceitos vulgares de crença, desejo, etc. Deixaremos de falar sobre o desejo X de S e passaremos a falar de um processo

cerebral X de S, que não precisará ser descrito em termos mentalistas. Porém, quando disser “processo cerebral X” estarei a falar sobre algo muito mais preciso e específico do que “desejo X”. Segundo os referidos eliminativistas, se um dia isto acontecer de modo massivo, de tal forma que a nossa comunicação corrente (vulgar) ocorra a partir deste novo vocabulário, será plausível dizer que não mais utilizaremos o vocabulário da psicologia popular. Passar-se-ia assim a produzir modelos explicativos da acção a partir deste novo vocabulário e a psicologia popular desapareceria.

Este debate trava-se na filosofia da mente e está longe de estar resolvido, ou mesmo estagnado. Dizer que poderemos um dia eliminar a psicologia popular é algo extremamente polémico e para muitos não parece algo plausível. É esse o meu caso. Não se desenvolverá, no entanto, no âmbito deste trabalho nenhuma defesa forte de uma posição não eliminativista. O nível explicativo do trabalho ficará aqui justificado pelo apelo à tradição da filosofia da acção, como área legítima da filosofia, que dá ao investigador um conjunto de conceitos específicos, que presumem que o eliminativismo não colhe. Não se discutirá portanto mais profundamente a opção por considerar o eliminativismo implausível. Por outro lado, é possível defender a ideia de que uma investigação filosófica que pretende ser útil nos termos das ciências sociais tem que, pelas razões acima apontadas, partir e ser realizada no âmbito da PP.

Assim sendo, como já foi dito, neste trabalho serão utilizados conceitos como crenças, desejos, intenção, acção, racionalidade, decisão, escolha, preferência etc. Muitos destes conceitos são estados mentais que, por sua vez, não podem ser manipulados e/ou tratados cientificamente (num sentido forte). São conceitos

naturalmente imprecisos, mas são eles que parecem captar uma boa parte dos fenómenos que me interessam.

1.1.3 EFICÁCIA DAS EXPLICAÇÕES DA ACÇÃO BASEADAS NA PSICOLOGIA POPULAR

O segundo tema desta secção é a possibilidade de utilização prática, ou aplicabilidade, das explicações da acção e da concepção de racionalidade produzidas pela filosofia, especialmente pela filosofia da acção.

A aplicabilidade da investigação não é um pré-requisito para se fazer filosofia da acção. Não parece adequado, nem legítimo, cobrar do filósofo da acção uma aplicação imediata, ou mesmo remota, da sua investigação. No entanto, dada a maior interdisciplinaridade contemporânea da investigação, parece razoável defender a legitimidade e a crescente importância dada à aplicabilidade do trabalho filosófico, *sem que isto signifique* uma condição necessária para a investigação filosófica. São cada vez mais comuns os artigos e livros publicados em parceria com filósofos. Os economistas, os psicólogos, os físicos, os biólogos, os neurocientistas, os antropólogos, os juristas, entre outros, já reconheceram a profundidade que a análise filosófica pode conceder aos seus estudos e procuram com muita frequência e regularidade estabelecer parcerias com filósofos e com grupos de investigação filosófica.

Além desta vertente, há também, apesar de isto ser mais comum na academia anglo-saxónica, a procura de uma formação interdisciplinar, isto é, tornaram-se mais comuns filósofos que têm uma primeira formação em outras áreas, normalmente numa daquelas acima citadas. É importante também ressaltar a importante inserção da filosofia nos centros de ciência cognitiva, principalmente nos Estados Unidos. Existe um importante número de filósofos que se dedicam, a partir da filosofia, às ciências cognitivas. Muitos filósofos com este perfil publicam em revistas científicas que têm como alvo primordial investigadores de outras áreas que não a filosofia.

Esta ligação horizontal da filosofia às outras áreas científicas produz naturalmente, em alguns casos, uma modificação nos objectivos sociais da investigação, sem que isto configure algum tipo de 'imperativo moral'. Quando um filósofo, por exemplo, discute o problema da causalidade e da responsabilidade moral pode fazer isto de maneira contida na filosofia, como também pode fazê-lo de modo a interagir directamente com os debates em curso no direito, ou pode ter a ambição de modificar a própria legislação que enquadra as relações sociais.

Para evocar o caso que me é mais próximo, no caso da economia e das políticas públicas esta relação é bastante produtiva. Por exemplo, a Teoria da Escolha Racional, apesar dos inúmeros debates, parece servir de base para várias ciências sociais além da própria economia, como, por exemplo, a gestão (privada e pública), a sociologia e a ciência política. Assim sendo, neste âmbito, a filosofia da acção cumpre papéis importantes quando analisa os pressupostos e as conclusões propostas pela teoria da escolha racional e quando produz a sua própria teoria da acção, teoria da acção esta que envolve desde a questão metafísica básica sobre o

que é uma acção até questões relacionadas com a explicação da acção e concepções de racionalidade envolvidas.

Tendo em vista essa forma alternativa de fazer filosofia, parece pertinente que um filósofo com este perfil reflecta sobre a eficácia da sua investigação, em termos aplicados. Uma explicação da acção, ou mesmo um modelo, será tanto mais útil à intervenção social ou individual – no sentido, por exemplo, da melhoria da qualidade de vida individual e da melhoria dos processos colectivos de escolha e de acção – quanto mais for compreensível através de várias disciplinas. Para isso será necessário utilizar termos partilhados por todos. A minha convicção é que a psicologia popular é suficientemente comum e segura para ser utilizada como porta de entrada e como meio para a produção de modelos explicativos da acção. Esta posição está bem representada em Fodor (1987, p. 20), na seguinte passagem:

Para dizer mais uma vez, a questão é que a teoria graças à qual obtemos este extraordinário poder preditivo é precisamente a velha psicologia de senso comum das crenças e dos desejos. Ela é aquela que nos diz, por exemplo, como inferir as intenções das pessoas a partir dos sons que pronunciam (se alguém emite em forma de palavras “chegarei ao aeroporto no voo das três da tarde”, então, *ceteris paribus*, tentará chegar ao aeroporto no voo das três da tarde). Também nos diz como inferir o comportamento das pessoas a partir das suas intenções (se alguém tem a intenção de chegar ao aeroporto no voo das três, então, *ceteris paribus*, se comportará de uma forma tal que o que acontecerá será que chegue a este lugar a esta hora, deixando à margem falhas mecânicas e actos divinos). Tudo isto não funciona apenas com as pessoas cuja psicologia conhecemos intimamente: seu melhor amigo, ou a sua esposa. Funciona com *completos estranhos*, com pessoas que não conheceria se tropeçasse nelas. Funciona não só em condições laboratoriais, nas quais se pode controlar as variáveis que interactivam, como também e, dizendo a verdade, predominantemente, em condições de campo nas quais tudo o que se sabe acerca das fontes de variação é o que nos diz acerca delas a psicologia de senso comum”

Se levarmos a sério o que Fodor diz, devemos também levar a sério a ideia de que podemos produzir, através do uso automático que fazemos da psicologia popular enquanto agentes humanos, modelos aplicados que partam das intuições psicológicas populares. Esta é a minha justificação, que considero forte, para apostar neste nível explicativo e nesta estratégia de investigação.

Podemos pensar num exemplo simples. Podemos supor que uma boa explicação da acção tende a dar-nos recursos suficientes para a construção de um modelo eficaz de previsão de acções futuras. Provavelmente, este modelo não explicará nada próximo da totalidade dos comportamentos verificados. No entanto, provavelmente, explicará uma parte relevante. Se, por exemplo, pensarmos na área sensível e controversa das políticas públicas poderemos perceber o quanto pode ser relevante qualquer incremento na compreensão e na previsão das acções. Só será possível agir em termos colectivos se construirmos modelos que possuam aderência à realidade linguística e mental dos agentes, de tal forma que as variáveis a serem variadas sejam reconhecíveis, tanto pelos *policy makers* como pelos indivíduos. Se um governo pretende intervir, por exemplo, num problema socio-comportamental como o tráfico de drogas, ou mesmo sobre algum determinado tipo de terrorismo local, precisará compreender as causas das acções das pessoas envolvidas. Neste caso, segundo uma perspectiva RPP, estaremos a tratar com crenças, desejos e intenções que causam a acção. Se pensássemos em lidar com estas questões sem apelar aos referidos estados mentais, teríamos que dizer que estes comportamentos foram ou serão causados pelos processos cerebrais X e Y, ou numa perspectiva behaviorista pelos estímulos W e T e pelas recompensas Q e K.

Para defender correctamente a posição de que o RPP pode dar resultados mais eficazes nas políticas públicas, ou mesmo em todas as ciências sociais aplicadas, se comparado, por exemplo, com o behaviorismo, seria necessário escrever uma nova tese e utilizar algum tipo de medida que foge ao trabalho filosófico. Por esta razão, aqui, mais uma vez, posiciono-me acerca de um tema paralelo, mas que é importante neste trabalho.

Uma linha de pensamento simples e que pode justificar, pelo menos, em termos superficiais este posicionamento é a seguinte. Não há dúvida de que existem e que são utilizados nas políticas públicas (ciências sociais aplicadas) modelos behavioristas. No entanto, em princípio, se considerarmos seriamente a forma como vulgarmente explicamos e prevemos as nossas acções, teremos que aceitar que mesmo uma política de origem behaviorista seria interpretada em termos mentalistas, isto é, não se foge da PP, ela actua automaticamente e confere um sentido intencional-causal a qualquer comportamento que pareça ser intencional, mesmo que ele responda a estímulos e a recompensas. Se isto for plausível, como parece ser, a utilização de teorias RPP parece ser o caminho mais directo e mais eficaz para atingir resultados aos níveis individual e colectivo superiores aos das políticas baseadas em modelos behavioristas.

Assim, e levando em consideração aquilo que até aqui introduzi nesta secção, defendo que o filósofo pode e deve ser “folk” em alguns sentidos específicos, sob pena de não ter conceitos e instrumentos adequados para trabalhar. Porém, não são todas as áreas da filosofia que necessitam de grandes doses de PP, como precisam a filosofia da acção, a filosofia da mente, a filosofia moral, a filosofia política e a filosofia das ciências sociais. Nestas áreas, os filósofos precisam da PP

por duas razões básicas: 1) a capacidade instrumental da PP, que foi mencionada nesta secção, e dá ao filósofo um repertório vasto e rico de conceitos a serem investigados e utilizados no trabalho filosófico; b) a aplicabilidade da investigação filosófica, que também foi mencionada nesta secção, e que é algo a ser desejado pelo filósofo, principalmente para aqueles ligados a algumas das áreas acima mencionadas.

Definida a minha posição acerca do nível explicativo e do quão folk poderei, ou deverei, ser neste trabalho, passo, na próxima secção, a definir os limites dentro dos quais procurarei trabalhar o conceito de acção.

1.2 ACÇÃO HUMANA: DO QUE ESTOU A FALAR?

Passo a definir o que se entende por acção humana neste trabalho. A definição proposta não é ela própria uma hipótese a ser testada, já que o trabalho não se insere primariamente no debate metafísico da filosofia da acção que procura responder à questão “O que é uma acção?”. Trata-se de uma definição de certa forma pragmática cujo objectivo é delimitar a abrangência e aplicação mais directa das propostas deste trabalho, uma vez que é necessário definir os limites dentro dos quais se pode em princípio aplicar a explicação da acção e o modelo da racionalidade na acção aqui propostos. O problema da natureza da acção é um problema metafísico de extrema importância e que tem sido o foco de muitos trabalhos em filosofia da acção, mas, como foi dito, não é um dos objectos de investigação neste trabalho. O meu objectivo é apenas um modelo de racionalidade na acção que dê conta, minimamente, de uma explicação que seja eficaz, porque cumpre os seus objectivos práticos – explicação e previsão.

Para tentar realizar esta tarefa, parece ser útil procurar quais são as condições necessárias para que um comportamento específico possa contar como uma acção, além da já referida abordagem mentalista da acção (RPP). Passo a explicitar o que, neste trabalho, contará como acção humana, através de uma lista de condições.

Condições Necessárias

C1: Ser Humano

C1.1: Ser humano vivo, sem morte cerebral

C1.2: A partir do tempo de vida em que surgem coisas como crenças, desejos ou intenções

A condição C1 parece ser demasiado evidente e intuitiva para merecer sequer ser mencionada. Porém, se é óbvio que para falarmos da acção humana precisamos do ser humano como um elemento essencial e básico, não é óbvio que o comportamento humano possua alguma particularidade que o distinga do comportamento de um símio, de um cão, de um computador, ou de um termómetro, por exemplo. Isso, no fundo, aponta para a possibilidade de ser, ou não, legítimo dizer algo como “acção humana”. Isto é, aponta para a dúvida acerca do que pode ser específico no comportamento humano que o distinga de comportamentos dos outros sistemas e constitua um caso único de ‘acção’. Por razões metodológicas, esse assunto não será aqui tratado, mas é importante dizer que por acção neste trabalho refiro-me apenas ao comportamento intencional levado a cabo por um ser humano. Essa é uma definição acerca do objecto dessa investigação e não uma resposta ao problema metafísico da natureza da acção. Parece desnecessário incluir qualquer defesa da possível especificidade do humano, pois, para os objectivos aqui presentes isso não teria função alguma.

Para qualificar C1, podemos perguntar: Quais são as condições que melhor definem um conjunto determinado de características dos seres humanos que serão o objecto deste modelos? Podemos elaborar perguntas básicas como:

C1.1 Estamos falando de seres humanos vivos que incluem, ou não, seres humanos com morte cerebral ou mesmo, por exemplo e em caso limite, seres humanos decapitados com movimentos corporais reflexos?

C1.2 Estamos falando de seres humanos nascidos, ou poderemos considerar o embrião e o feto como seres humanos sobre os quais este modelo deverá também incidir?

C1.3 Estamos nos referindo a seres humanos que têm controlo consciente, eficaz e efectivo sobre os seus movimentos corporais, ou podemos, por exemplo, considerar que essa tentativa de explicação da acção poderá incluir acções de um ser humano com uma patologia motora, psiquiátrica, ou mesmo neurológica – além de casos complexos como, por exemplo, o dos sonâmbulos, ou daqueles que estão sob efeito de uma droga alucinogénea?

Temos, natural e intuitivamente, a tentação de produzir modelos que sejam válidos para todas as classes e para todas as condições humanas. É provável que em algum dia isso seja possível. Agora, parece-me preferível restringir o objecto sob pena de não chegarmos a lado algum. Alguém poderá dizer que uma boa explicação da acção humana deveria envolver todas as classes e condições acima referidas (1-3) para ser considerada uma explicação forte da acção humana. Partindo do pressuposto, de certa forma intuitivo, de que isso não é possível, chamemos, então, *explicação mínima* àquela que restringe classes e condições – por exemplo, aquela a ser desenvolvida aqui. Essa ambição já é suficientemente grande para esse trabalho ser rotulado como “ambicioso”.

Não me parece que conseguiremos determinar, de maneira não pragmática, as classes e condições humanas sobre as quais uma explicação forte da acção humana deve incidir. Sendo assim, pretendo de maneira justificada escolher determinadas classes e condições, tendo sempre em vista as possibilidades, restrições e limites metodológicos da filosofia e a esperada utilidade social de uma explicação desse género.

Começo por (C1a): Parece intuitivo dizer que, pelo menos, a maioria de nós, seres humanos adultos, está inclinada a pensar em acções humanas como algo que depende de um processo de deliberação que, independentemente do tempo que dura, necessita de um agente e de uma intenção que possa ser explicada apelando à mente do agente, ou à mente de um observador. Levando em consideração essa forte intuição, mas admitindo que não se trata de um argumento, parece-me pouco razoável incluir, como objecto dessa explicação específica da acção, pessoas com morte cerebral, ou decapitados com movimentos reflexos, por exemplo.

Cabe aqui ressaltar que, de facto, é muito difícil defender que um ser humano com morte cerebral, ou, mesmo, um ser humano decapitado possa agir de maneira intencional. No entanto, imagino que poderá haver objecções do tipo: não é difícil imaginar um agente que, ainda com a cabeça “colada” ao corpo produz a intenção de, por exemplo, “fechar a mão” para demonstrar coragem, milésimos de segundo antes de ser decapitado. Assim, a mão se fecha mesmo depois de a cabeça estar separada do corpo. Bom, se isto for possível, este poderia ser um exemplo de uma acção intencional “sem cabeça”. De facto, algo desse género pode ser dito, mas não vejo nenhuma relevância, nem dificuldade em afirmar que esse fechar da mão

esquerda foi intencional, mesmo que o agente estivesse “sem cabeça” no instante em que a acção foi executada.

E se pensarmos no caso de um ser humano decapitado que mesmo com um frágil funcionamento cerebral e com um resto mínimo de consciência segue por um milésimo de segundo um pássaro, ou o movimento de uma coisa qualquer no mundo? Podemos dizer que esse movimento é uma acção que cumpre a condição C1.1? Se essa história for possível, a resposta será sim – esse movimento dos olhos, ou de qualquer outro que seja possível levar a cabo apenas com a cabeça, cumpre a condição C1.1.

Vejamos C1.2: Parece necessário afirmar, por questões metodológicas, que seres humanos não nascidos e com menos de 4 anos devem ficar fora do campo de abrangência destes modelos explicativos da acção e da racionalidade na acção. Desta forma, consideraríamos apenas os seres humanos vivos a partir do tempo de vida em que surgem coisas como crenças, desejos ou intenções⁴.

Independentemente do que eu penso acerca da relação entre a intencionalidade e o tempo de vida do ser humano, parece-me importante estabelecer um critério, um limite para a aplicação e para o teste do modelo. Isto é importante para que se possa testar intuições dentro dos limites em que o modelo se aplica. Isto é, é preciso dar tanto aos filósofos como aos cientistas condições para testar a explicação da acção aqui desenvolvida. Vamos supor, de maneira absurda e no limite, que o modelo é perfeito, mas que não possuímos instrumentos e estratégias suficientes para testá-lo em crianças com dois dias de vida, ou no feto, ou no embrião, e vamos também supor que todos estes são seres humanos. O que

⁴ Para uma visão geral acerca do desenvolvimento cognitivo e dos estados mentais, sugiro a leitura de Goswami (2008).

aconteceria com um tal modelo? Seria refutado por falta de aderência a uma parte importante do campo explicativo. Parece ser pouco razoável propor um modelo explicativo que não possa ser testado, nem por uma experiência mental básica, nem em seres humanos nessas condições. É disso que procuro escapar com essas restrições.

Pretende-se aqui elaborar uma explicação tal que possa ser testada mesmo que a abrangência não seja a totalidade dos seres humanos. Talvez seja uma maneira simplista de pensar, mas parece-me defensável dizer que se uma explicação da acção é boa para uma mulher de 20 anos deverá ser razoável para um homem de 40 anos, mas talvez não seja totalmente eficaz para uma menina de seis meses, dado o que sabemos acerca da mente dos neonatos ou das crianças muito novas (até aos 4 anos). Não sabemos tanto assim acerca do nosso desenvolvimento para podermos afirmar categoricamente quando surgem as crenças, os desejos e a intencionalidade, ou se simplesmente sempre estiveram lá, ou, no limite, se aquilo a que chamamos crenças e desejos, por exemplo, esteve sempre lá com atribuições funcionais e adaptativas diferentes daquelas verificadas em crianças a partir dos 4 anos.

Outras questões podem ser relevantes, como, por exemplo: o comportamento de um sonâmbulo que organiza a sua mesa de trabalho enquanto dorme é considerado uma acção? Um indivíduo que utiliza uma droga alucinógena e salta numa piscina sem água agiu num sentido aqui relevante? Será que podemos dizer que todas as acções humanas dependem de uma acção corporal perceptível? Pense no caso de um homem que está na margem do rio e vê uma pessoa a afogar-se. Ele,

após alguns instantes de deliberação afirma não ter a intenção de ajudar aquela pessoa. Portanto não faz nenhum movimento. Isso não é uma acção/omissão?

Além destes casos, existem, certamente, muitos outros que são, pelo menos, tão difíceis quanto estes. Para responder de maneira sistemática e relevante às perguntas que surgem a partir destes casos, seria necessário produzir uma teoria da acção que fosse um sistema explicativo complexo e com a ambição de ser completo.

Ciente da contribuição à qual me proponho, não posso dedicar esforços adicionais a esta investigação paralela. Por isso, talvez, a maneira mais simples de resolver esta necessidade de limitação do campo seja enumerar alguns requisitos para que uma acção possa ser incluída no conjunto das acções explicadas pelo modelo a ser proposto, além da condição de que as acções se predicam a seres humanos (C1).

Requisitos

- R1) São acções individuais - acções que ocorrem num momento, ou durante um período de tempo, num local determinado de cada vez.
- R2) São externas (em comparação a algo como acções mentais, tais como efectuar uma subtracção ou lembrar-me do que almocei ontem)
- R3) Podem ou não incluir movimento corporal (acção, ou omissão no sentido vulgar)
- R4) São positivas - produzem alguma mudança no mundo.
- R5) São intencionais.

Estes requisitos⁵ são simples, intuitivos e auto-explicativos, mas podem conter algumas armadilhas. Para minimizar os possíveis efeitos destas, comentarei brevemente cada um dos requisitos.

O requisito R1 tenta diferenciar acções que são constituídas de várias acções e acções que são únicas. Se considerarmos que construir uma casa é uma acção, poderemos desmembrá-la em várias outras acções, como: construir a sala, a casa de banho, o terraço, os alicerces, o telhado, etc. Podemos dizer que comprar os azulejos e contratar o empreiteiro também são acções que constituem a acção “construir uma casa”. No entanto, temos também que considerar que comprar os azulejos da cozinha é uma acção individual que está ligada a uma outra acção de fundo que provoca, de algum modo, o desejo de comprar os azulejos e dá condições para que uma determinada crença se associe a este desejo.

De facto, podemos cair na tentação de dizer que todas as acções são, de algum modo, dependentes de outras acções e de desejos de ordem mais elevada, como, por exemplo, o desejo de ser saudável que me faz desejar comer salada no almoço que, por sua vez, me faz ter o desejo de ir ao restaurante W naquele momento. Enfim, aqui é importante perceber que aquilo que forma, ou provoca, um determinado desejo e/ou uma crença não está em jogo. A partir do momento que um comportamento é percebido como o objecto a ser explicado, é aquele comportamento específico que estará em jogo. Isto é, aquele comportamento terá que ter, para ser uma acção, uma explicação intencional que nos leve a crenças e a desejos específicos, sejam eles quais forem.

⁵ Estes requisitos seguem parcialmente Lagier (2001).

Os requisitos R2 e R3 podem trazer algum problema. Quando digo que são externas, não estou a dizer que algum movimento corporal deve ser percebido pelo observador, obrigatoriamente. Se eu observo S a tomar café, posso isolar o comportamento característico e explicá-lo, considerando-o, em princípio, uma acção nomeada como “tomar café”. No entanto, se eu observo S a calar-se diante de uma ofensa proferida directamente à sua mãe, posso intuitivamente isolar este comportamento e explicá-lo, considerando-o, em princípio, uma acção nomeada como “calar-se”. Nos dois casos, utilizei a expressão “em princípio” porque o observador tende a dizer que a acção foi X, ou Y, para iniciar o processo de investigação, mesmo que depois tenha que reformular a caracterização da acção.

O requisito R4 é bastante intuitivo, mas deve ser considerado em conjunto com R5. Não são todos os movimentos corporais que modificam o mundo, que podem ser considerados acções, principalmente no contexto do modelo que aqui será apresentado. Por exemplo, o balançar das pernas quando estamos a ler não seria considerado um movimento do corpo que poderíamos caracterizar como uma acção intencional. Poderíamos fazer isto apenas se o indivíduo dissesse que aquele balançar de pernas era, por exemplo, um sinal secreto para o seu comparsa do outro lado da sala, ou se o observador percebesse, por exemplo, que aquele movimento pudesse ser explicado por razões (crenças e desejos). Parece claro que a linha que separa um simples evento de uma acção intencional pode ser muito ténue. Tendo, de facto, a considerar quase impossível a tarefa de produzir uma definição precisa e infalível. Porém, como estou preocupado com a grande maioria das acções (pelo menos aquelas que são, de algum modo, relevantes, de acordo com o que foi dito anteriormente acerca da utilidade desta investigação), não me

parece adequado dar demasiada importância a casos limítrofes que poderão ser investigados posteriormente e que não implicam nenhuma forte objecção aos modelos que serão aqui apresentados.

O requisito R5 é aquele que será mais bem descrito ao longo desta tese. No entanto, cabe aqui, resumidamente, dizer que conta como uma acção intencional um comportamento que pode ser explicado através de razões. Estas são caracterizadas por estados mentais específicos, como são os casos das crenças, dos desejos ou das intenções. Em contraste, podemos pensar, como foi dito em R3, em eventos que não conseguimos explicar por meio de razões, como o mexer dos dedos dos pés quando estamos assistindo televisão, por exemplo. É claro que estes movimentos têm explicações, mas estas não são caracterizadas por razões.

Resumidamente, a condição C1 (1.1 e 1.2) e os requisitos R 1-5 delimitam aquilo a que passarei a chamar de acção daqui para a frente, neste trabalho.

1.3 RACIONALIDADE NA ACÇÃO: DO QUE ESTOU A FALAR?

A ideia de *racionalidade* está, de modo generalizado, presente no nosso quotidiano, seja implícita seja explicitamente. É plausível pensar que os seres humanos sempre tiveram mecanismos de escolha e de motivação para a acção, qualquer coisa como ‘mecanismos de racionalidade’. No entanto, a investigação acerca da racionalidade ligada à escolha e à acção, a própria ideia de ‘racional’ como um adjectivo frequentemente utilizado no discurso popular, parece tornar-se mais frequente quanto mais nos aproximamos da actualidade. Isto é muito claro quando pensamos no desenvolvimento relativamente recente da filosofia da acção, da teoria da escolha racional, da própria teoria microeconómica e mesmo da psicologia (num sentido amplo). É preciso lembrar também que provavelmente⁶ nem todas as culturas dão a importância que nós, ocidentais americanos e europeus, damos à racionalidade na acção, ou valorizam da mesma forma a qualificação de uma acção ou de um agente como racional ou irracional. Esta constatação não tem relevância directa para o argumento da tese, e por isso permanecerá como uma intuição não investigada, não sendo, muito embora, desprovida de interesse.

Mas que queremos de facto dizer quando dizemos que uma acção, ou alguém, é racional ou irracional? Falamos frequentemente sobre acções racionais e irracionais nas típicas “conversas de boteco”, quando, por exemplo, alguém diz que S agirá de maneira racional ao decidir-se pelo namorado *T* e não pelo namorado *W*, pelas razões *r_S*, ou, então, quando alguém *P* diz que é irracional um clube de futebol

⁶ Utilizei a palavra “provavelmente” porque não pretendo utilizar, ou realizar neste trabalho, nenhuma comparação intercultural que pudesse dar-nos alguma resposta mais fiável.

(C) pagar “tanto dinheiro” por um jogador, pelas razões r^p . Fala-se também sobre a racionalidade nas conversas mais inocentes, como, por exemplo, quando se acusa alguém de ser irracional por ter enveredado pelo caminho “incerto” da música e não pelo caminho “certo” da medicina. Acho que todos nós já ouvimos algo deste género: “é irracional comprar um carro amarelo”. A qualificação ‘irracional’ pode aqui significar que comprar um carro amarelo pode ser estranho, ou esquisito ou incomum por causa da cor, como pode também querer dizer que a acção é irracional porque esta cor diminui muito o valor futuro do carro.

Resumidamente, as pessoas comuns utilizam as qualificações ‘racional’ e ‘irracional’ para adjectivar acções que parecem produzir resultados mais (ou menos) aceitáveis, de um ponto de vista consequencialista, ou, então, também, acções que representam meios mais, ou menos, aceitáveis, de um ponto de vista instrumental. Essas formas de falar de ‘racional’ e ‘irracional’ não estão assim tão afastadas de algumas concepções filosóficas acerca do conceito. No entanto, neste trabalho, o termo racionalidade será utilizado de uma maneira específica, que procurarei elucidar, já que é acerca desse sentido de racionalidade que pretendo produzir um modelo funcionalista. Nesta tese, o conceito a ser tratado refere-se ao raciocínio prático particular e não a considerações de tipo generalista. Explico em seguida o que entendo por considerações de tipo generalista. Podemos pensar em duas formas básicas e muito simples de pensar acerca da racionalidade na acção: o particularismo e o generalismo. De acordo com o Particularismo (a), uma acção que é racional/irracional num conjunto de circunstâncias não é necessariamente racional/irracional em outro. De acordo com o Generalismo (b) uma acção que é

racional/irracional num conjunto de circunstâncias tem que ser racional/irracional em qualquer conjunto de circunstâncias.

Por exemplo, dizer “comprar um carro amarelo é irracional” é admitir alguma forma de generalismo, mesmo que seja apenas “um modo informal de falar”. Por outro lado, ao dizer “se S comprar um carro amarelo será irracional pelas razões particulares r^s ” é admitir que existem razões particulares r_s que tornam a acção de “comprar um carro amarelo” irracional para S (r^s). As razões r são razões do agente e estão circunscritas num tempo e num conjunto específico de conteúdos dos estados mentais do próprio agente.

Não é difícil compreender e defender a ideia de que a maioria das pessoas parte de uma concepção generalista sobre a racionalidade para formar juízos acerca das acções de outros. Quantas vezes não ouvimos coisas como “o sadomasoquismo é irracional”, ou “fumar é irracional”, ou “estudar filosofia é irracional”? O que está em jogo nestes casos é a defesa simplista do modo de agir inverso ao criticado. Esta concepção generalista não é o que está aqui em causa. No entanto, é importante ressaltar que o uso da noção de racionalidade, seja na filosofia, na psicologia, ou na economia, está frequentemente ligado a exemplos do tipo dos acima mencionados, que têm, na minha opinião, uma natureza claramente normativo-prescritiva.

Escrever sobre a racionalidade pode ser uma tarefa difícil tendo em conta que (quase) todos nós temos uma concepção daquilo que conta como racional e não estamos muito dispostos a desfazermos-nos facilmente dela. As noções de ‘racional’ e ‘irracional’ são extremamente úteis para classificar acções e, desse modo, ajudam-nos a mapear o mundo social, e de resto também a prever as acções dos

outros numa interacção estratégica – que é um tipo comum e importante de relação social. Falta saber, obviamente, se os dois casos são idênticos. Todos nós podemos dar exemplos de acções que consideramos racionais e podemos dizer que “conhecemos” definições “objectivas” da racionalidade na acção, que podem conduzir a afirmações curiosas tais como “ser racional é agir de acordo com o bem comum”, ou “a racionalidade é o melhor meio de decidir sobre o que fazer”, ou “ser racional é agir de maneira auto-interessada”.

Mas vejamos os problemas que podemos estar a ignorar. Lembro-me de uma aula de *análise organizacional* no mestrado em gestão pública e governo (que tinha no corpo discente gestores, antropólogos, sociólogos, economistas, entre outros cientistas sociais, além de alguns políticos profissionais) em que o debate era sobre o orçamento participativo, isto é, um modelo de definição de orçamentos públicos no qual a população local pode intervir nas decisões acerca das prioridades no uso dos recursos públicos. Nesse dia, um aluno, que era vereador numa cidade próxima de São Paulo, ficou absolutamente perplexo quando foi exposto a um exemplo no qual a população preferiu em T_1 pavimentar a rua ao invés de realizar as obras de saneamento básico – obras que, normalmente, são realizadas abaixo da pavimentação. Disse-nos algo assim: “Isso é um absurdo! Qualquer pessoa percebe que é irracional pavimentar a rua antes de colocar os tubos/canos de esgoto e de água potável...”. De facto, de um ponto de vista técnico parece ser muito pouco defensável a decisão (colectiva) da população. Naquele instante, ouve-se a voz do professor, um escocês divertido que, felizmente, desafiava constantemente as nossas intuições. Disse-nos algo assim: “Não, não ... não foi uma decisão irracional ... as pessoas moram no local há mais de 20 anos e, obviamente, produziram

soluções não ideais para o problema do saneamento básico – deram algum jeito e convivem com essa solução que produz algum equilíbrio. No entanto, não conseguem andar pelas ruas não pavimentadas, seja num dia seco, ou num dia de chuva, sem ficarem sujos com a terra, ou com a lama, respectivamente. A preferência colectiva é essa, por mais absurda que pareça. Seria uma decisão irracional se fossem contra essa preferência”.

Esse exemplo é importante, pois mostra-nos a fragilidade e a imprecisão do uso comum do conceito de ‘racional’. Existem muitas definições e muitos usos pouco precisos do conceito de racionalidade na acção, ou, de maneira alternativa, racionalidade na escolha. Talvez seja ingénuo pretender criar um modelo de racionalidade na acção que venha a ser partilhado por todos, no entanto o meu propósito, que me parece razoável e alcançável, é precisamente propor uma descrição da racionalidade na acção que seja neutra e que possa servir como terreno comum para definições e modelos de ordens mais elevadas. É a isso que vou chamar ‘modelo funcionalista’. Até agora tenho falado de acções de indivíduos como sendo racionais ou irracionais e é a estas que em primeiro lugar o modelo se aplicará. No entanto, quando falamos de preferências individuais, temos um conjunto específico de problemas para resolver, mas quando falamos de preferências e intenções colectivas, temos um outro conjunto caracterizado principalmente pelo problema da agregação dessas preferências. No item seguinte, por uma questão de ordem, descreverei o *Paradoxo de Condorcet* que trata da agregação das preferências individuais.

Voltando ao modelo que pretendo propôr: o intuito desta tese é propor o embrião de um modelo funcionalista da racionalidade na acção, afastado de

elementos morais e normativos-prescritivos. Queria propôr uma analogia básica para explicitar o estatuto desse modelo: a ideia é comparar a racionalidade com o sangue no corpo humano. Não é possível existir um ser humano sem sangue, assim não podemos dizer que existem seres humanos não-sanguíneos (irracionais, na analogia) – ou têm sangue, ou não são seres humanos. A racionalidade assim descrita é uma condição necessária, mas não suficiente⁷ para que um ser seja humano. A racionalidade na acção tem a função de produzir intenções e/ou acções tendo como matéria-prima diversos inputs, sejam eles subjectivos (internos ao agente), sejam externos ao agente (intersubjectivos ou objectivos). A racionalidade não é algo a que as ideias de *bom* e de *mau* se aplicam - assim como não se aplicam ao sangue. Pretendo, portanto, dar à racionalidade um estatuto funcional comparável ao estatuto do sangue no corpo humano.

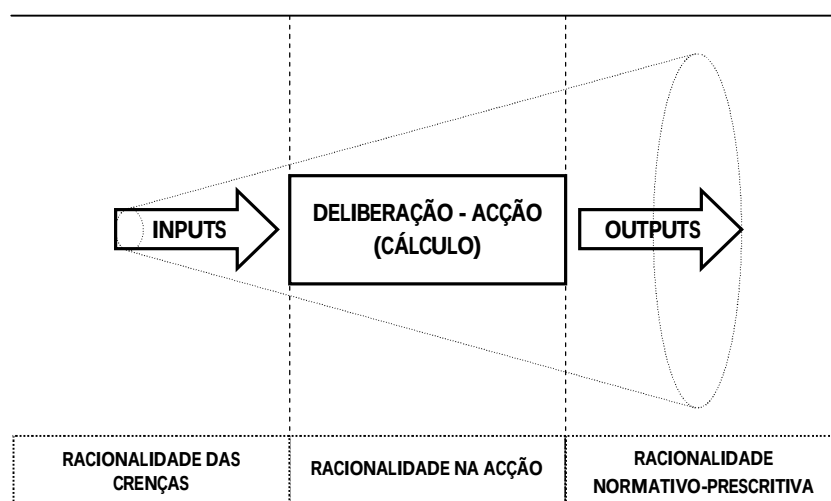
Como pretendi esquematizar na Figura 1, à frente, a racionalidade na acção tem a propriedade de sintetizar/processar inputs (desejos, crenças, por exemplo) transformando-os em outputs (acções ou pensamentos que podem ou não levar a acções). A racionalidade na acção participa no cálculo deliberativo e não como uma instância avaliativa externa. Da mesma forma, na racionalidade na acção a formação racional, ou não, das crenças (de um ponto de vista normativo-prescritivo) não é levada em consideração: as crenças e os desejos entram como dados não avaliativos no sistema de cálculo.

Não é preciso entrar em muitos detalhes para percebermos que, se esta for de facto a função da racionalidade na acção, não será possível qualquer tipo de generalismo sobre a racionalidade, tanto de um ponto de vista instrumental, como

⁷ Não vou abordar a questão da racionalidade nos animais não humanos, mas parece-me claro que a definição aqui proposta parece obrigar-me a afirmar que os animais não humanos podem ser racionais.

consequencialista. Cada ser humano conta com um conjunto específico de conteúdos mentais e irá sintetizar os inputs de maneiras diferentes de acordo com o seu próprio sistema racional. A ideia cultural de uma racionalidade externa ao agente (normativa-prescritiva) pode, por vezes, acomodar as diferenças ao produzir ou moldar os sistemas racionais individuais de acordo com pressupostos definidos dentro de teorias normativas-prescritivas da racionalidade, como, por exemplo, a teoria da escolha racional (TER).

Figura 1: Racionalidade na Acção – Fluxograma Básico



Isto é, podemos verificar que muitas acções humanas são derivadas da aprendizagem de modelos normativo-prescritivos de acção, e que, portanto, muitos desses modelos passam a gozar de certo estatuto descritivo, mesmo que, claramente, não sejam descritivos. Por outro lado, muitos desses modelos, como é o caso da teoria da escolha racional, são extremamente úteis, pois parecem de alguma forma dar conta de acções colectivas relevantes para as ciências sociais.

Nesse sentido, podemos começar por “desdiabolizar” a teoria da escolha racional, admitindo nomeadamente que ela é extremamente útil, pois dá conta de muitos fenómenos colectivos. A questão é que ela não é ainda o instrumento adequado para explicar e prever a acção individual. A vantagem da TER é ser robusta no sentido formal e permitir uma série de previsões e de simulações que outras abordagens não permitem. Por exemplo, a TER permite que utilizemos a teoria dos jogos como uma ferramenta de análise do comportamento do agente económico, mas permite também que possamos prever o comportamento político, como é frequentemente o caso na ciência política contemporânea. Na Sociologia alguns modelos analíticos utilizam a TER como uma ferramenta essencial para compreender a acção social.

O presente trabalho procura um caminho que não se afaste completamente da TER, mas que incorpore elementos mais sensíveis na explicação e previsão da acção individual. O modelo resultante desse intuito será desenvolvido no capítulo 2. Antes de avançar nesse sentido, retomarei alguns pontos essenciais do tratamento do tema da racionalidade na economia, com base na TER, na filosofia e na psicologia popular.

1.3.1 RACIONALIDADE NA ACÇÃO NA ECONOMIA

A Economia é uma ciência social e assim sendo trata principalmente de questões do nível agregado ou colectivo ou social. Existem, no entanto, várias formas de conceber a produção de 'explicações sociais', isto é, a natureza das explicações dos fenómenos sociais. Os quatro principais tipos básicos de explicação dos fenómenos sociais são o atomismo, o holismo radical, o anti-reducionismo, e o individualismo metodológico, como mostra o Tabela 1 (Levine, 1989, p.59).

Tabela 1 - Tipos de Explicação de Fenómenos Sociais

		Propriedades e relações entre entidades sociais agregadas são irredutivelmente explicativas	
		<i>Sim</i>	<i>Não</i>
Relações entre indivíduos são explicativas	<i>Sim</i>	Anti-reducionismo	Individualismo metodológico
	<i>Não</i>	Holismo radical	Atomismo

Os quatro tipos acima referidos, diferem quanto ao que consideram exclusivamente explicativo. Uma boa e útil definição do que são as entidades sociais agregadas é a de Levine (1989, p. 59):

Entidades sociais agregadas incluem, por exemplo, sociedades, grupos, classes, organizações, nações, comunidades. Essas entidades possuem propriedades (por exemplo, taxas de inflação, formas institucionais, modelos de distribuição de renda) e realizam

uma variedade de relações entre si (por exemplo, relações entre sindicatos e empresas, entre nações, entre classes colectivamente organizadas). Os indivíduos também têm tanto propriedades (crenças, hábitos, recursos) tanto quanto entram numa diversidade de relações com outros indivíduos (relações de parentesco, de patrão-empregado etc.).

Vejamos, agora, uma descrição resumida de cada um dos tipos de explicação dos fenómenos sociais. O *atomismo* não aceita nenhum tipo de explicação que venha da *relação entre indivíduos*, nem entre *entidades sociais agregadas*. Para um atomista, somente são explicativas as entidades individuais constituídas de modo plenamente não relacional. Sendo assim, a explicação da acção seria redutível apenas ao indivíduo. O atomismo sustenta, ainda, que os factos ligados às relações sociais afectam as acções apenas em virtude de influenciarem, atomisticamente, os pensamentos individuais. (Levine, 1989, p.60).

O anti-reducionismo aceita explicações micro-fundamentadas mas defende a irredutibilidade das propriedades e relações entre entidades sociais agregadas. Resumidamente, os anti-reducionistas entendem que há explicações micro-fundamentadas, mas que, porém, há outras que não são redutíveis ao indivíduo. Para os anti-reducionistas não existe a necessidade, que existe para os individualistas, da redução das explicações sociais ao indivíduo.

O *holismo radical*, ou *colectivismo metodológico*, ao contrário do individualismo metodológico, tem como elemento central na explicação da acção social as propriedades e relações entre entidades sociais agregadas, nos termos acima descritos. Para os colectivistas metodológicos, as relações entre os indivíduos são geradas pela “operação do todo”. Portanto, “não se trata, apenas, de que o ‘todo é mais que a soma das partes’, mas que o todo é a causa exclusiva e as partes são

meros artefactos, ainda que constituídos a partir de relações sociais” (Levine, 1989, p.61). Sendo assim, a intencionalidade na acção passa a ser uma propriedade do *todo*, e não mais dos indivíduos, ou seja, estes são apenas agentes de forças sociais elas próprias dotadas de intenção/intencionalidade, como, por exemplo, as classes sociais.

O colectivismo metodológico assume a existência de leis gerais de ordem ou de desordem social que regulam a acção individual (Orenstein, 1998, p.159). Elster escreve que “o colectivismo metodológico parte da premissa que existem entidades supra-individuais que são anteriores ao indivíduo na ordem de explicação. A explicação procede das leis destas entidades, enquanto as acções individuais são derivadas deste agregado. (Elster, 1989, p.6 – cf. Orenstein, 1998, p. 159)

O *individualismo metodológico*, assim como o atomismo, concebe a explicação da acção social como sendo redutível ao nível individual. Não, porém, ao indivíduo isolado, mas ao conjunto de relações entre os indivíduos. O individualismo metodológico supõe que as acções sociais *intencionais* dos indivíduos, baseadas nas suas crenças e desejos, produzem ordem ou desordem social, como consequência *não intencional* da combinação das decisões individuais. As propriedades relacionais não são redutíveis a propriedades atomísticas. Assim, para o individualista metodológico, a explicação do fenómeno social encontra-se essencialmente nas relações entre indivíduos.

A acção individual intencional acima referida pode ser estratégica ou paramétrica. Na acção estratégica cada actor pondera a acção dos demais na sua

própria acção. Na acção paramétrica a acção de cada actor é decidida a partir de uma estrutura fixa de informações.

Dada a superioridade explicativa das relações entre os indivíduos, em detrimento das relações entre entidades sociais agregadas e da redução a propriedades atomísticas, pode-se entender que no individualismo metodológico o *todo* visto como a soma das acções individuais não relacionais é menor do que o *todo* entendido como a combinação de decisões individuais estratégicas (no sentido acima descrito), que considera as relações entre os indivíduos.

A Teoria Económica tem como base das suas explicações de fenómenos sociais o individualismo metodológico. É assim por exemplo, no caso dos modelos explicativos de propriedades de entidades sociais agregadas tais como a taxa de juros, a inflação, etc. É também por isso que o agente económico individual é extremamente relevante para os economistas. Na Economia, a acção humana e a racionalidade na acção são elementos centrais nos modelos positivos e normativos que os economistas têm produzido desde Adam Smith (mesmo que de maneira não formal, mas nem por isso menos interessante). Muitos dos mais recentes prémios Nobel de Economia têm tratado directa, ou indirectamente, desses temas, como são os casos de Daniel Kahneman, Vernon Smith, Robert Aumann, Thomas Schelling, Amartya Sen, John Harsanyi, John Nash ou Richard Selten.

O agente económico racional é um dos pilares da TER, a qual, por sua vez, é um dos pilares da Teoria Económica. No entanto, podemos perguntar-nos se não faria mais sentido dizer que uma determinada acção pode ter sido racional, ou irracional, ao invés de dizer que o agente económico é racional. Por outro lado, a verdade é que quando os economistas afirmam que o agente económico é racional

isso não significa que acreditem, de facto, que as pessoas são completamente racionais (por toda a vida).

Daniel Hausman (Mauro, no prelo) sintetiza assim a questão:

Apesar de os economistas mais ou menos identificarem a acção (escolha) com a acção racional, eles não estão comprometidos com a visão contra-intuitiva de que não existe algo como a acção irracional. As fontes mais comuns da acção irracional são as crenças irracionais, as preferências irracionais, ou uma falha da ligação que deve existir entre a acção e as crenças e os desejos do próprio agente.

O pressuposto de que o agente económico é racional tem um papel central na Teoria Económica porque permite que sejam construídos modelos comportamentais que explicam uma parte das acções humanas, especialmente as acções colectivas. É claro que os agentes económicos não são completamente racionais, de acordo com a definição de racionalidade da Economia. No entanto, é esse pressuposto que permite que a TER confira robustez metodológica à Teoria Económica (TE) e à Teoria dos Jogos, por exemplo. Isso é de alguma forma análogo àquilo que nós, pessoas comuns, fazemos todos os dias quando imputamos racionalidade uns aos outros, quando, por exemplo, associamos a boa condução de um carrinho no supermercado à capacidade racional do outro de tomar decisões aceitáveis de acordo com um conjunto esperado de acções, como não jogar o carrinho contra uma prateleira, ou contra o meu carrinho – segundo uma concepção popular de racionalidade. É certo que verificamos no dia-a-dia que as pessoas agem muitas vezes de uma tal maneira que poderíamos classificar de irracional, à luz de uma teoria popular da racionalidade. No entanto, conseguimos

excelentes resultados esperando que todas as pessoas ajam de maneira racional, nesse sentido popular.

Olhemos agora para um outro aspecto da consideração dos agentes como racionais. Pressupor que os agentes são maximizadores das relações entre preferências, é o que os economistas chamam de racionalização da escolha.

Ariel Rubinstein (2006, p.20) define assim a questão da racionalização:

Os economistas eram, muitas vezes criticados por pressuporem que os agentes⁸ maximizam a relação entre preferências. A resposta mais comum a esta crítica é que nós realmente não precisamos deste pressuposto. Tudo que temos de assumir é que o comportamento do agente pode ser descrito *como se* ele estivesse maximizando a relação entre preferências.

A ideia popular de que a TE classifica categoricamente as pessoas como racionais e irracionais leva a situações curiosas nas quais são confrontadas formas de viver, gostos, desejos, sentimentos, enfim, todo tipo de elementos que não são discutíveis à luz da teoria económica (do que discutem os economistas). São muitos os exemplos desse género de confronto, mas pode pensar-se num exemplo simples como aquele em que são confrontados formas ou estilos de vida. Por exemplo, não foram poucas as vezes que ouvi de amigos artistas coisas como “esse mundo é muito racional e esse é o mal do mundo”, ou “a racionalidade torna-me menos criativo”. Ouvi também de amigos que trabalham, por exemplo, como gestores de bancos, ou indústrias, coisas como “esses artistas/ecologistas são todos irracionais”. Ou, então, o clássico confronto entre os alunos que gostam de matemática e os alunos que não gostam e preferem, por exemplo, a literatura.

⁸ decision makers

Quase sempre os primeiros são considerados por muitos como sendo os mais racionais e os segundos são considerados os menos racionais. Terão essas ideias acerca da racionalidade alguma repercussão na TE? Não, de forma alguma. Um economista nunca diria que comprar um carro alemão é sempre uma acção racional, porque o carro é alemão, ou porque há algum juízo avaliativo positivo associado intrinsecamente aos carros alemães. Há muitas variáveis em jogo e, claro, comprar um carro alemão pode ser uma acção irracional perante a TE.

A Teoria Económica do Consumidor (TEC) é um dos alicerces da TE, pois é através dela que o economista produz conhecimento acerca do comportamento humano, dado o facto das explicações económicas serem derivadas de uma posição individualista metodológica. Os elementos centrais da TEC são a restrição orçamental e as preferências dos consumidores. Os economistas supõem que os consumidores escolhem o melhor cabaz de bens que podem adquirir. Isto é, escolhem a partir de um ranking de preferências dentro de um conjunto de oportunidades específico. Como “melhor cabaz” entende-se o cabaz preferido dado o ranking de preferências do consumidor. Como “cabaz de bens que podem adquirir” entende-se o cabaz possível dado o conjunto de oportunidades. Tratarei aqui apenas de alguns aspectos das preferências dos consumidores. No entanto, algumas definições anteriores são importantes.

Quando se lê TEC deve-se lembrar, como dito há pouco, que é nesta área da TE que são desenvolvidas as investigações acerca da acção humana. Quando se lê ‘consumidor’ deve-se entender ‘agente humano’. A TEC trata especificamente de agentes cujas escolhas envolvem inevitavelmente restrição orçamental e cabazes de bens. Porém, podemos alargar essa perspectiva pensando na restrição

orçamental como sendo uma restrição de meios, sejam eles quais forem, e os cabazes de bens como sendo quaisquer elementos que possam ser compreendidos como sendo opções a serem consideradas num processo de deliberação. É assim que, e apesar de tratar do agente económico, a Economia produziu uma série de conceitos que podem ser úteis aos filósofos que tratam da acção e da racionalidade. O objectivo aqui é expor um ponto muito específico acerca de uma possível analogia entre alguns conceitos da TEC e da Teoria da Acção desenvolvida na Filosofia. Nesse sentido, não se fará uma revisão crítica da TEC. Mesmo assim, torna-se necessária uma pequena introdução sobre a relação existente entre a restrição orçamental, as preferências e a escolha.

1.3.1.1 PREFERÊNCIAS, ANÁLISE CUSTO-BENEFÍCIO, RESTRIÇÃO ORÇAMENTAL, ESCOLHA E RACIONALIDADE

PREFERÊNCIAS

As preferências são ordenadas segundo a sua desejabilidade, isto é, de acordo com a TEC, somos capazes de ordenar as preferências de tal forma que uma preferência seja estritamente preferível à outra, ou que sejam indiferentes.

No discurso corrente, mesmo entre os economistas, é comum dizer-se que um cabaz é preferível a outro, sem a necessidade de dizer se é fraca ou estritamente preferível. De facto, assume-se que quando X é preferível a Y , X é estritamente preferível a Y ($X \succ Y$), por outro lado, assume-se que quando X é debilmente

preferível a Y ($X \succ Y$) e Y é debilmente preferível a X ($Y \succsim X$), X é indiferente a Y ($X \sim Y$).

Quando se diz que o sujeito S prefere estritamente X a Y ($X \succ Y$), diz-se que S deseja mais o cabaz X (x_1, x_2) do que Y (y_1, y_2). Esse é um conceito que deve ser prático, no sentido em que se para S $(x_1, x_2) \succ (y_1, y_2)$, S escolherá X se tiver oportunidade de escolher.

Podemos considerar, do ponto de vista filosófico, a frase acima como um princípio que rege a acção, isto é, como um princípio que define o sentido da escolha/acção a partir da desejabilidade/preferência associada a cada uma das opções.

(PE) Princípio da Escolha/Acção: *Se, para S , $X \succ Y$, e S for livre, S escolherá X se tiver oportunidade de escolher.*

Obviamente, o conteúdo específico de cada ranking de preferências é diferente entre indivíduos. Na Economia, não está em jogo se uma pessoa prefere cheeseburger ou sushi para o jantar, mas sim se a sua acção condiz com as suas preferências. Segundo os economistas, apesar das diferenças entre os indivíduos, é possível dizer que grande parte das relações de preferências tem algumas características em comum. Alguns economistas, como por exemplo Hal Varian (1999), fazem referência a três “axiomas” básicos acerca das relações entre as preferências – essas relações seriam: A1) *completas*, A2) *reflexivas* e A3) *transitivas*. Outros autores, como, por exemplo, Mas-Colell, Whinston e Green (1995), enfatizam A1 e A3.

Para os objectivos desta tese, parece-me adequada a abordagem de Robert Frank (2006) que tem a qualidade de ser, na minha opinião, mais próxima do que realmente pensam os economistas. Pode-se, então, dizer que são quatro as propriedades das relações de preferências, que conferem determinada consistência às escolhas dos agentes⁹. Essas quatro propriedades que são responsáveis pela atribuição da característica de consistência são os seguintes (Frank, 2006):

P1: As relações entre as preferências são **completas** pois o consumidor deve ser capaz de comparar dois cabazes quaisquer. Isto é, o cabaz X poderá ser preferível ao cabaz Y , ou o cabaz Y poderá ser preferível ao cabaz X , ou o agente poderá ser indiferente entre os dois cabazes. Isto é, dado qualquer cabaz X e qualquer cabaz Y , podemos supor que $(x_1, x_2) \succcurlyeq (y_1, y_2)$, ou $(y_1, y_2) \succcurlyeq (x_1, x_2)$, ou as duas coisas. Resumidamente, a relação completa permite ao agente poder escolher entre duas, ou mais, opções.

P2: As relações entre as preferências são **transitivas** pois supõe-se que se um cabaz X é preferível ao cabaz Y e o cabaz Y é preferível ao cabaz W , logo, o cabaz X será preferível ao cabaz W . Isto é, $(x_1, x_2) \succcurlyeq (y_1, y_2)$ e $(y_1, y_2) \succcurlyeq (w_1, w_2)$, então, $(x_1, x_2) \succcurlyeq (w_1, w_2)$.

P3: As relações entre as preferências são do tipo “**quanto mais, melhor**”. Isto é, se tudo se mantiver constante o agente sempre preferirá uma quantidade maior a uma menor de um bem. Isso, resumidamente, significa que dada uma situação em que estou satisfeito com uma

⁹ Os economistas não afirmam que os agentes agem assim, sem excepção. Utilizam este recurso metodológico para poder criar um modelo explicativo do comportamento do agente económico que seja útil às decisões económicas dos consumidores, empresas e governo.

quantidade (possível) do bem X , dadas as minhas restrições, um aumento da quantidade desse bem é melhor do que a sua diminuição.

P4: A propriedade da **convexidade** refere-se à característica geral dos agentes de preferirem combinações de bens, ao invés dos extremos. Isto é, se, de alguma forma, ficamos satisfeitos com 8 unidades de w e 0 unidade de t [Cabaz $A=(8,0)$], ou 0 unidade de w e 8 unidades de t [Cabaz $B=(0,8)$], ficaremos ainda mais satisfeito com 4 unidades de w e 4 unidades de t , isto é, preferimos o cabaz $C=(4,4)$, ou qualquer outra mistura entre os bens w e t .

Vejamos, então, as principais questões envolvidas em cada uma das quatro propriedades.

A primeira propriedade, a das preferências completas, define que os agentes são capazes de escolher, isto é, são capazes de comparar dois cabazes quaisquer. É difícil imaginar uma objecção a este pressuposto no campo económico. Mesmo que um indivíduo se encontre numa situação hipotética em que tem que escolher entre duas opções excludentes, por exemplo, água e comida, quando está sedento e faminto ao mesmo tempo, terá condições de ordenar preferências (mesmo que esta seja uma escolha difícil). No entanto, parece estranho afirmar que o agente pode ser fortemente indiferente entre dois cabazes e mesmo assim agir. Um caso simples, em que, em princípio, deveríamos ficar indiferentes é aquele em que nos oferecem dois copos de água (X,Y) , aparentemente iguais, quando estamos sedentos. Para bebermos a água, temos que escolher um dos dois copos oferecidos. Poderíamos associar quantidades idênticas de utilidade aos copos com água, isso

parece plausível à luz de A1, pode-se dizer que poderíamos associar a mesma desejabilidade. No entanto, quando agimos, pegamos num copo específico e não no outro. Para agir intencionalmente, tivemos que escolher entre o copo X e o copo Y. Por exemplo, podemos: (a) escolher o copo da esquerda ou o da direita porque o copo X (esquerda) é preferível ao copo Y (direita); (2) pegar o copo mais, ou menos, iluminado; (3) pegar o copo que chegou primeiro; (4) pegar o que está mais à frente; (5) pegar o que parece mais fresco; enfim, são muitas as razões possíveis, desde que sejam razões que explicam a minha acção intencional, isto é, desde que signifiquem que, por mais absurda que seja, eu agi por uma, ou por outra, razão. Enfim, mesmo de modo parcialmente inconsciente, preferimos um copo ao outro, logo, no fim das contas, um terá que ser estritamente preferível ao outro para que uma acção seja levada a cabo, caso contrário terá sido um evento e não uma acção intencional. Portanto, a relação completa parece plausível se considerarmos que para um indivíduo agir é necessário que a relação entre as preferências seja estrita e caso seja uma relação de indiferença, o indivíduo não agirá, de facto.

Cabe aqui uma ressalva importante. Os economistas não distinguem estritamente 'escolha' de 'acção'. Presume-se que a acção corresponde à escolha. No entanto, como veremos, a distinção entre (acto de) escolha e acção é importante para o modelo defendido neste trabalho, principalmente se a escolha estiver directamente associada ao conceito de decisão. Pode dizer-se que de maneira geral, na filosofia da acção, se considera que uma decisão leva a uma intenção e que esse estado mental leva, de algum modo, o sujeito a agir.

A segunda propriedade, a da transitividade, é uma hipótese sobre a conduta dos indivíduos na escolha e não uma afirmação puramente lógica (Varian, 1999). É

importante lembrar que, segundo este pressuposto, as relações entre os cabazes de consumo X , Y e W são binárias, portanto, os cabazes são comparados dois a dois.

Uma objecção possível é a seguinte: imagine um agente que tem que escolher em T_1 entre as tortas de limão X e de abacaxi Y e escolhe a de limão X , isto é $X \succ Y$. Em T_2 , tem que escolher entre as tortas de abacaxi Y e de banana W e escolhe a de abacaxi Y , isto é $Y \succ W$. Logo, através do axioma da transitividade, conclui-se que o agente, em T_3 , preferirá a torta de limão X à de banana W , isto é $X \succ W$. Teríamos então $X \succ Y \succ W$, em que $X \succ W$. Vamos supor, porém, que o agente real preferiu em $T_{3+\Delta}$ a torta de banana W à de limão X , isto é, $W \succ X$. Teríamos, portanto, uma situação em que $X \succ Y \succ W$, em que $W \succ X$. Concluir-se-ia, assim, que o agente levou a cabo uma acção irracional, no sentido da TER. Porém, tal conclusão pode ser refutada. Poderiam ser várias e complexas as razões, porém, basta uma simples razão eficaz, isto é, uma razão que leva o agente à acção, para que esta objecção seja plausível. Imaginemos que, em algum momento anterior a $T_{3+\Delta}$, este agente esteve numa situação análoga e escolheu a torta de limão X ao invés da torta de banana W ($X \succ W$), e que ocorreu, associada a esta escolha, uma desgraça qualquer. O agente, então, produziu uma associação negativa a esta situação de decisão entre X e W , suficientemente forte para levá-lo a inverter a preferência numa situação análoga e posterior. Sendo assim, o indivíduo preferiria de facto a torta de banana W à de limão X , quando colocadas numa situação de decisão posterior à trágica experiência. Desta forma, nada o impediria de preferir racionalmente a torta de banana W à de limão X em T_{3+n} .

O meu exemplo pode fazer algum sentido quando pensamos na transitividade ao longo do tempo, como algo mais parecido com “escolhas” não contraditórias no tempo.

Porém, será que podemos mesmo falar de contradição no tempo e de transitividade no tempo, sem parecer que estamos a cobrar do agente algo que não é possível, isto é, uma estabilidade de preferências que não parece ser real?

De facto, a transitividade e mesmo a contradição são conceitos que têm aplicação num instante específico, isto é, em T_1 , $X \succ Y \succ W$, em que $X \succ W$. Se isso é verdade e em T_1 eu tenho que escolher entre X e W , escolherei sem dúvida alguma X , caso contrário não será uma acção intencional. Os economistas diriam que seria uma acção irracional, no entanto, parece impossível dizer que o agente escolheria W e não X em T_1 , se $X \succ Y \succ W$. O que podemos dizer, no limite, é que as condições para um observador externo perceber a escala de preferências desse agente não foram adequadas e que, portanto, a escala de preferências é, de facto, outra, por exemplo, $W \succ X \succ Y$, em que $W \succ Y$.

É curioso como não percebemos como somos incrivelmente pouco exigentes ao dizer que somos capazes de agir intencional e livremente contra a nossa própria escala de preferências, sem que isso signifique, de facto, uma mudança nessa escala.

A terceira propriedade refere-se à ideia, razoavelmente popular, de que preferimos uma quantidade maior de um bem (preferido) a uma quantidade menor. Essa propriedade gera, obviamente, algumas objecções imediatas, como, por exemplo, quando pensamos numa situação em que bebemos uma quantidade maior de água do que aquela que é suficiente para matar a sede. Ficamos

satisfeitos com dois copos de água, mas se bebermos sete copos podemos passar mal. Para os economistas, um exemplo desse tipo diz respeito a dificuldades práticas, como um problema de autocontrolo (não conseguir parar de beber água), ou o de não conseguir armazenar o bem para consumo futuro (não ter um recipiente onde armazenar a água para mais tarde). A ideia que de facto está por trás parece bastante intuitiva. Se pudéssemos armazenar o máximo possível de bens que gostamos, mantendo tudo constante, não seria razoável dizer que isso produziria algum mal ao agente. No entanto, normalmente associamos uma quantidade maior de um determinado bem a alguns custos adicionais que já caracterizam uma modificação nos factores em jogo. Por exemplo, se pudéssemos entrar num supermercado e levar para casa tudo o que quiséssemos, sem que isso significasse um aumento na factura, provavelmente ficaríamos tentados a levar o máximo possível dos bens que gostamos e que são mais caros, por exemplo, vinho, chocolates especiais, carnes, especiarias, etc. Teremos, no entanto, que pensar nas restrições que temos que respeitar, como, por exemplo, a nossa capacidade de armazenar a comida, que conta com a utilidade negativa de termos 30 caixas de vinho em nossa sala e 20 caixas de chocolate em nosso quarto. Enfim, podemos rapidamente pensar em restrições práticas que limitariam a nossa acção de levar connosco o máximo possível de bens preferidos. Porém, o que os economistas dizem não é o contrário disso. O “quanto mais, melhor” verifica-se apenas quando tudo continua constante. A partir do momento em que os factores são modificados (como a utilização da casa e a incapacidade de armazenagem), deixamos de preferir aquele bem e passamos a preferir “mais espaço”, por exemplo.

A quarta propriedade, a convexidade, representa a intuição, também popular, de que as pessoas preferem combinações de bens à exclusividade de um bem sobre o outro. Se gostamos de arroz (X) e de feijão (Y), mas somos indiferentes aos dois (X é fracamente preferível a Y e Y é fracamente preferível a X), preferimos comprar uma combinação qualquer que inclua uma porção de arroz e uma porção de feijão, ao invés de comprarmos apenas arroz, ou apenas feijão. Isso significa que a escolha melhor possível é aquela que possui quantidades diferentes de zero dos dois bens, em detrimento da exclusividade de um só bem – que também estaria dentro da curva de preferências do indivíduo (veja-se mais à frente no item “restrição orçamental e escolha”).

Acerca das preferências individuais há ainda outra questão que é importante ressaltar. Mesmo que tenhamos um grupo social cujos indivíduos tenham preferências racionais e portanto transitivas (de acordo com a TER), poderemos verificar preferências colectivas intransitivas, portanto irracionais. O Paradoxo de Condorcet mostra-nos precisamente isso: a forma como preferências individuais transitivas podem levar-nos a preferências colectivas intransitivas. Utilizando um exemplo simples, cujo desenho é já conhecido, mostrarei como funciona o referido paradoxo¹⁰.

Considere-se um concurso musical, onde amadores cantam para obterem notoriedade. Pense-se, por exemplo, no American Idol com três jurados. O primeiro é bastante conservador e não gosta de artistas inovadores. Vamos chamá-lo conservador (C). O segundo é moderado e às vezes parece ser conservador, mas também gosta de alguma inovação. Vamos chamá-lo moderado

¹⁰ Para outro exemplo semelhante, veja-se Fiani, 2006, p. 27-29.

(M). O terceiro é vanguardista, gosta/defende a inovação musical (a qualquer preço) e repele o conservadorismo. Vamos chamar-lhe vanguardista (V). Os três jurados têm preferências transitivas e cada um tem direito a um voto.

No último dia da competição, são os jurados que escolhem o vencedor. São três os finalistas: “Luís Brochado”, “Maria Saidinha” e “MC Bomba”. O primeiro é conservador e cantou com uma grande voz uma canção como “deve ser cantada” segundo C. A esse candidato damos o nome de LB. A segunda concorrente é uma cantora segura e que cantou um fado em ritmo de bossa nova. Agradou imenso a M. A esta candidata damos o nome de MS. O terceiro candidato caiu nas graças de V, cantando uma música inclassificável em termos rítmicos. A esse candidato damos o nome de MB.

As preferências dos jurados podem ser assim descritas:

Tabela 2: Preferências dos Jurados

C	LB > MB > MS
M	MS > LB > MB
V	MB > MS > LB

O jurado C prefere LB a todos os outros, no entanto, se não for LB a ganhar, prefere que MB ganhe a competição, pois acredita que esse candidato não terá a mínima possibilidade de se transformar num ídolo nos 12 meses seguintes. Assim sendo, provará a sua tese de que bons cantores são aqueles que cantam como ele defende que devem cantar.

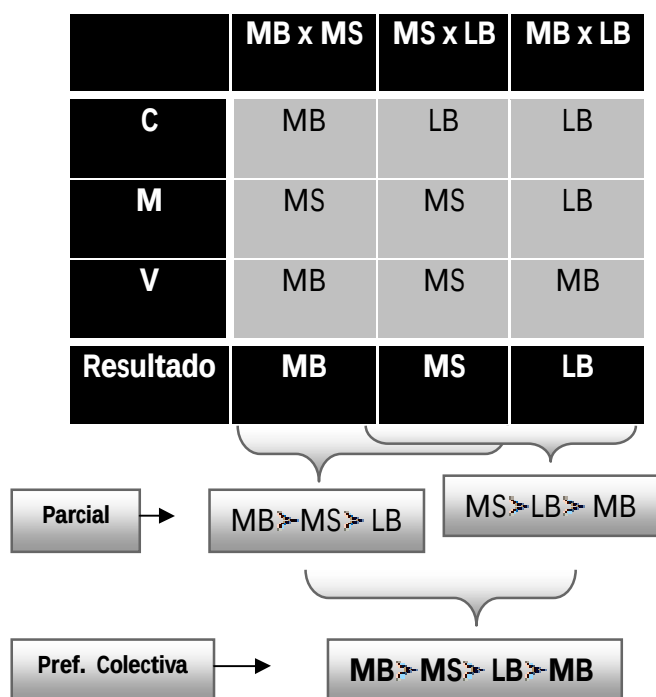
Para M, o melhor candidato é MS. Seguindo a sua tendência ligeiramente conservadora, M coloca em segundo lugar na ordem de preferências LB e em terceiro MB.

Para o jurado V, não há dúvidas, MB é de longe o melhor candidato. A contragosto, se fosse necessário votaria em MS e em último em LB.

Será que é possível afirmar que existe uma escala de preferências do júri, isto é, será que existe uma relação transitiva (racional) entre as preferências do júri? Para tentarmos descobrir o que se passa, vamos supor que os candidatos vão a confronto em pares.

Seguindo a Tabela 2, podemos perceber os seguintes resultados:

Tabela 3: Preferências Colectivas



Dadas as preferências individuais transitivas, verificamos os seguintes resultados: Na primeira disputa [MB–MS], o resultado foi MB > MS. Na segunda

disputa [MS–LB], o resultado foi $MS \succ LB$. Portanto, o resultado dessas duas primeiras disputas dá-nos a seguinte ordem de preferências: $MB \succ MS \succ LB$. No entanto, na última disputa [MB–LB], o resultado foi $LB \succ MB$. Portanto, os três resultados agregados dão-nos a relação entre preferências intransitiva: $MB \succ MS \succ LB \succ MB$.

Esse exemplo mostra-nos, de maneira intuitiva, aquilo que o Paradoxo de Condorcet pretende demonstrar. Vimos, então, que podemos ter preferências intransitivas ao nível agregado, mesmo que ao nível individual as preferências sejam transitivas.

ANÁLISE CUSTO-BENEFÍCIO

Para os economistas, além destes pressupostos com objectivos descritivos, o elemento normativo central para a tomada de decisão é a análise custo-benefício. Isto é, quando um economista é consultado por um agente acerca de se deve ou não fazer X , o economista produzirá a resposta através da análise custo-benefício. O princípio é simples: deve realizar-se uma acção X somente se os benefícios de realizar esta acção, $B(x)$, excedem seus custos, $C(x)$. Ou seja:

- (1) Se $B(x) > C(x)$, faça x ;
- (2) Se $B(x) < C(x)$, não faça x .

Pense na seguinte situação. Está a assistir a um bom filme pela televisão, deitado no sofá, coberto por uma manta, numa noite fria, e pensa: “Quero comer

pipocas. Devo levantar-me, ir à cozinha, preparar as pipocas e voltar para o sofá, ou devo deixar para outro momento?”. Esta é uma situação de decisão em que a análise custo-benefício é claramente percebida. Farei pipocas somente se $B(\text{fazer pipocas}) > C(\text{fazer pipocas})$.

Os economistas afirmam que quase sempre é possível traduzir os custos e os benefícios de uma acção para um referencial monetário. Neste exemplo, pode calcular-se $C(x)$ e $B(x)$ da seguinte forma. Se alguém lhe oferecesse 1 euro para se levantar, fazer pipocas e voltar para o sofá, você aceitaria? Provavelmente diria que não. Porém, se lhe oferecessem 100 euros para levar a cabo a mesma acção, provavelmente aceitaria. Isto significa que o seu preço de reserva está entre 1 euro e 100 euros. O preço de reserva é o montante mínimo que o levaria a levantar do sofá. Digamos que após vários exercícios deste tipo descobrisse que o seu preço de reserva é de 10 euros, isto é, $C(x) = 10$ euros. Então, $B(x)$ deverá ser maior que 10 euros para a acção ser realizada. Para determinar $B(x)$, podemos proceder da mesma maneira até encontrar o preço de reserva correspondente. Digamos que, naquele momento, está disposto a pagar, no máximo, 5 euros para alguém preparar as pipocas para você. Assim sendo, $B(x) = 5$ euros, teríamos $B(x) < C(x)$, portanto, não deveria realizar a acção¹¹.

Este cálculo não é realizado desta forma explícita e sistemática pelo agente real. Os economistas sabem disto. No entanto é assim proporcionada uma boa estratégia analítica para medir os possíveis custos e benefícios de uma acção. Os economistas assumem, assim, que mesmo que o agente real não faça tais cálculos

¹¹ Para um exemplo análogo, veja-se o exemplo 1-1 do capítulo 1 de Frank (2006).

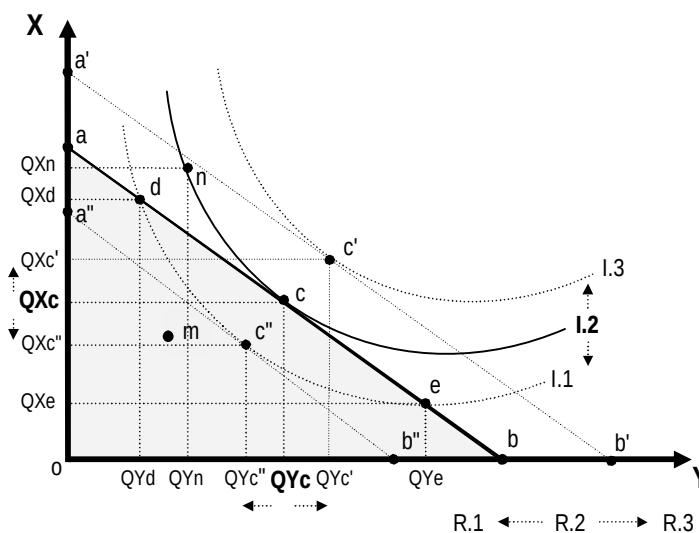
exactamente como foi descrito acima, a economia pode prever muitas acções se partir do princípio de que os agentes agem sempre deste modo.

RESTRIÇÃO ORÇAMENTAL E ESCOLHA

Outro elemento essencial é a restrição orçamental, que pode ser descrita como o conjunto de possibilidades de consumo do agente, pois representa as combinações possíveis entre cabazes de consumo dentro dos limites orçamentais do agente. Isto é, representa as combinações possíveis que um agente pode gerar entre os bens que quer consumir dado o seu rendimento disponível. Por exemplo, José ganha cinco salários mínimos e tem que equilibrar os seus gastos de modo a garantir, pelo menos, a sua sobrevivência biológica, além de garantir seu acesso a alguns bens que considera importantes, tais como a educação, a saúde e a diversão. José constrói, então, um orçamento com 30 itens, distribuindo os seus recursos financeiros de maneira a satisfazer suas preferências de consumo. Portanto, neste exemplo, o agente deve ter alguma consciência acerca das suas preferências e do seu conjunto de possibilidades de consumo. Este conjunto é representado na Figura 2 pelo triângulo AOB (área cinzenta), onde a recta R.2 representa a sua restrição orçamental. Isto significa que dentro deste conjunto José pode fazer qualquer combinação de consumo, já que todas estarão dentro dos seus limites orçamentais, isto é, dentro das suas possibilidades. Neste caso, os pontos *a*, *b*, *c*, *d*, *e*, *a''*, *b''*, *c''*, *m*, são exemplos destas combinações possíveis. Já os pontos *a'*, *b'*, *c'*, *n*, são combinações que estão fora do referido conjunto de oportunidades. Sabemos,

portanto, quais são as combinações possíveis, mas, não sabemos, ainda, qual é a combinação que maximiza as preferências. Para isto, teremos que cruzar o conhecimento que temos acerca da restrição orçamental e das preferências de José.

Figura 2: Curva de Indiferença e Restrição Orçamental



As preferências são representadas por curvas denominadas curvas de indiferença. Estas curvas indicam quais são as combinações entre cabazes de bens que maximizam as preferências do agente dado o seu ranking de preferências. Na Figura 2, as curvas $I.1$, $I.2$, $I.3$ são exemplos destas curvas. O ponto n representa uma combinação onde José fica satisfeito com uma determinada quantidade do cabaz X (QXn) e com outra do cabaz Y (QYn), mas está fora do seu conjunto de possibilidades. De outra maneira, ele também fica satisfeito no ponto d , que está dentro do seu conjunto de possibilidades, com a combinação entre as quantidades QXd e QYd . Portanto, a curva de indiferença representa o ranking de preferências do agente na escolha entre cabazes de bens.

Podemos, agora, descobrir a combinação que maximiza a utilidade esperada dados os recursos financeiros de José. Basta relacionar a recta de restrição orçamental R.2 com as possíveis curvas de indiferença. Encontraremos, então, um ponto, numa destas curvas, que tangencia R.2 e que maximiza o uso dos recursos dadas as preferências de José. Na Figura 2, o ponto *c* representa a melhor combinação possível entre os cabazes X e Y dentro do seu conjunto de possibilidades, enquanto QY_c e QX_c representam as quantidades que maximizam a utilidade da sua escolha.

No instante 1, José revela uma restrição orçamental R.2, uma curva de indiferença I.2 e as quantidades QY_c e QX_c . Podemos supor, por exemplo, que o seu rendimento disponível, no instante 2, aumentou de cinco para seis salários mínimos, diminuindo a sua restrição orçamental e aumentando o seu conjunto de possibilidades. Assim, haveria um deslocamento de R.2 para R.3, onde uma outra curva de indiferença (I.3) seria utilizada e outro ponto (*c'*) maximizaria suas preferências com as quantidades $QY_{c'}$ e $QX_{c'}$. Inversamente, caso no instante 2 o seu rendimento diminuísse para quatro salários mínimos, utilizaríamos para análise a recta R.1, a curva I.1, o ponto *c''* e as quantidades $QY_{c''}$ e $QX_{c''}$.

Há muitas explicações adicionais essenciais acerca deste tema, relativas, nomeadamente, a mudanças parciais nos preços, mudanças no rendimento do agente, mudanças de preferências, etc. Enfim, há toda uma área teórica desenvolvida que não pode ser resumida em poucas linhas¹². Para os objectivos deste trabalho, as explicações que introduzi são suficientes, apenas. O meu objectivo é que o leitor que não é economista, ou tão-pouco filósofo ligado à teoria

¹² Sugiro para aqueles que querem, de maneira introdutória, estudar mais acerca da TEC: a) os capítulos 2, 3, 4, 5, 7, 10 e 12 de Varian (1999); b) os capítulos 1, 3, 6, 7, 8 e o apêndice do cap. 3 Frank (2006).

da decisão, possa compreender os elementos básicos da TEC que são importantes para a leitura do resto do trabalho.

RACIONALIDADE E IRRACIONALIDADE

A hipótese central da TE, mais especificamente da análise microeconómica, é que os indivíduos são racionais. Mesmo não havendo consenso sobre o que significa exactamente dizer que os indivíduos são racionais, há, pelo menos, a confiança de que é útil em termos analíticos pressupor que os indivíduos agem de acordo com os seus rankings de preferências e que as relações entre essas preferências seguem, de algum modo, na maioria dos casos, os quatro princípios anteriormente descritos.

Há que acrescentar o seguinte. Na economia, pode-se basicamente abordar a racionalidade através de dois conceitos: (1) racionalidade sob o ponto de vista dos *objectivos imediatos*; (2) racionalidade sob o ponto de vista *do interesse próprio*. A abordagem dos objectivos imediatos (AOI) como padrão de racionalidade implica que o agente será racional quando for eficiente ao cumprir os objectivos que tiver definido para si próprio, no momento da acção. Uma observação relevante é perceber que quando os economistas falam sobre objectivos definidos no momento da acção, não estão a problematizar o processo que leva à acção, isto é, não estão a considerar se essa decisão/definição acerca de objectivos imediatos produz uma acção, ou uma intenção. Em termos filosóficos isso é muito importante, como veremos nos capítulos 2 e 3. Por enquanto, faço aqui apenas o registo dessa questão.

Nessa abordagem da racionalidade não é importante investigar se os objectivos imediatos fazem algum sentido através de algum critério. Não importa, por exemplo, se o agente tem uma preferência por acções auto-destruidoras, ou se prefere ter prejuízo a ter lucro. O que é relevante saber é se o agente procura agir da maneira mais eficiente (quanto aos seus recursos) em busca da realização desses objectivos.

Por outro lado, a abordagem do interesse próprio (AIP) pressupõe que, de modo geral, a motivação para a acção é congruente com os interesses estritamente materiais do agente. Nesse sentido, acções altruístas, ou justas por serem desejadas, entre outros tipos, não são consideradas acções racionais.

Boa parte dos economistas aceita a objecção de Stigler comumente denominada “problema do óleo do cárter”. Segundo Stigler, se assistimos a um indivíduo a beber o óleo usado do cárter do seu próprio carro, em seguida a agonizar e a morrer, poderemos concluir que aquele indivíduo gostava muito daquele óleo. Segundo Stigler, a AOI leva-nos a explicar qualquer acção, por mais bizarra que seja, depois dela ter sido levada a cabo. Assim, em princípio, essa abordagem permite uma explicação para tudo, mas, exactamente por ser tão abrangente, passa a não explicar nada – essa é uma frase comum aos críticos da AOI.

Esse problema do óleo do cárter não me parece convincente. Como veremos mais à frente. De um certo ponto de vista filosófico, a acção de beber o óleo do cárter pode ser explicada por razões que de facto explicam a acção. O sujeito pode não gostar de óleo, mas pode, por exemplo, querer suicidar-se da maneira mais dolorosa possível. Pode até querer satisfazer um desejo por uma espécie de

martírio. Enfim, há muitas explicações possíveis. No entanto, se aqueles que aceitam a objecção do óleo do carácter estiverem a dizer que basta um suposto gosto/ desejo por óleo de carácter para levar o indivíduo a agir, na minha opinião estarão a cometer um erro grave. Aceitar que as pessoas agem por objectivos imediatos não parece dizer que as pessoas têm apenas uma ordem de preferências ou desejos. A pessoa pode ter a forte preferência por óleo de carácter se comparada ao sumo de abacaxi, mas se souber que o óleo mata e se tiver uma preferência mais básica por viver a morrer, certamente o desejo pelo óleo será transformado num desejo ocioso, isto é um desejo que existe, mas que não levará o indivíduo à acção. Essa objecção do óleo do carácter parece-me muito básica em termos psicológicos. Claro que aqui o debate está a ser resumido, mas, de forma geral, os economistas têm produzido critérios para medir a racionalidade dos agentes nas acções a partir de pressupostos comportamentais que são, na minha opinião, normativo-prescritivos. É o caso dos custos afundados, por exemplo. Ser normativo-prescritivo não será um problema se o que queremos é criar uma régua para medir a racionalidade nas acções, a partir de critérios explicitamente normativo-prescritivos. No entanto, se quisermos, como é o meu caso, propor um modelo descritivo da racionalidade teremos problemas em aceitar tais critérios. Passo ao exemplo acima citado, o dos custos afundados, para fundamentar este ponto de vista. Este é um exemplo de irracionalidade do tipo que os economistas classificam como sendo derivado de limitações cognitivas. Caso o agente ultrapasse essas limitações agirá, certamente, de acordo com os princípios da teoria.

Um dos princípios do modelo de escolha racional é aquele que afirma que os custos afundados devem ser ignorados quando o indivíduo decide. Os custos

afundados são aqueles que já não são mais recuperáveis no momento da decisão. Vejamos o seguinte exemplo: João e Maria vão a um rodízio de carne. João é daquele tipo que adora carne e come quase até passar mal. Maria é um pouco diferente, come e gosta de carne, mas troca a carne por um bom prato de legumes e salada. João, convida Maria para sair - ainda não são namorados, mas estão quase lá. Vão ao cinema e João diz que está com muita vontade de ir ao Vitelão, o melhor rodízio de carne da cidade. Maria, ainda querendo agradar ao provável futuro namorado, aceita a proposta. Como pretendem ser um casal não machista, cada um paga a sua parte da conta – o preço por pessoa nesse restaurante é igual a uma semana de almoço no restaurante onde Maria almoça normalmente. Naquele dia, Maria não está com muita vontade de comer carne. Na verdade, preferia comer um prato de legumes cozidos e uma boa salada ao invés de carne, a especialidade do Vitelão. Essa alimentação leve (legumes e salada) custaria 5 euros num outro local, mas no Vitelão as pessoas pagam 50 euros na entrada e podem comer o que e quanto quiserem, inclusive legumes cozidos, salada, carne, sobremesa, etc.. O que deve Maria fazer? Muitos dirão: “ela deveria desistir e dizer que não quer ir com João”. Mas isso não é mais uma possibilidade, pois ela já aceitou a proposta, estão no restaurante e ela tem que decidir o que comer, isto é, ela já pagou os 50 euros. Em muitos casos deste tipo, as pessoas, mesmo preferindo não comer a carne, acabam comendo, para “maximizar o dinheiro gasto”, ou para “amortizar o investimento realizado”, como alguns dizem. No entanto, no sentido da TER estão redondamente enganadas. Mesmo tendo pago 50 euros na entrada, Maria deveria ignorar esse custo, pois este é um custo afundado, isto é, ela pode fazer o que quiser, mas não o recupera mais. Ela deveria pensar com mais tranquilidade acerca

do que mais lhe apetece fazer, uma vez que já pagou – deve ignorar os custos afundados. Nesse caso, ela deveria pedir um bom prato de legumes e de salada, enquanto João come os dois quilos de carne que está disposto a comer. Maria agiria de maneira racional, de acordo com a TER, se, mesmo pensando estar a “desperdiçar” 45 euros (50 euros que pagou na entrada – 5 euros que pagaria pelos legumes/salada noutra local) comer apenas aquilo que, de facto, lhe apetece. Se Maria, como acontece muitas vezes com muitas pessoas, escolhesse comer muita carne para “fazer valer o seu investimento”, estaria a piorar a sua situação, pois além de pagar os 50 euros estaria a consumir o que não quer, ou aquilo que prefere menos.

Outro exemplo típico é o seguinte: o que deveria fazer uma empresa que depois de ter investido 10 milhões de euros num projecto descobre que este não é economicamente viável? Deveria ignorar os custos afundados (10 milhões) e esquecer o tal projecto. No entanto, e por mais incrível que pareça, há gestores que têm muita dificuldade para desistir de um projecto nessa situação. Pensam, “já gastamos 10 milhões...vamos perder esse dinheiro?”. A resposta dada pela TER é: sim. Na verdade, esse investimento conta como custos afundados e, por isso, não há nada a fazer para recuperá-los. Seguir em frente, ou desistir não produzirá nenhum efeito sobre os 10 milhões. Nesse instante, o gestor deveria pensar em não perder dinheiro no futuro, dado o resultado da análise do projecto.

Apesar de serem dois casos de dificuldade em ignorar custos afundados, o primeiro caso que descrevi está mais ligado ao tema desta tese, pois a acção de Maria está num nível individual, mesmo se, de alguma forma, existe uma interacção estratégica com João. Nesse caso, parecem existir dois, ou mais níveis de

raciocínio prático que estão em jogo e que podem levar Maria a agir de maneira racional, mesmo que a sua explicação da acção não pareça sê-lo, num primeiro momento.

Ignorar os custos afundados é um dos princípios que regem as acções racionais a partir da abordagem dos interesses próprios. Neste caso, o agente deve levar em consideração o melhor resultado de um ponto de vista económico. Apesar disso, podemos dizer que o melhor resultado pode ser aquele que satisfaz os seus desejos e as suas crenças de 1ª ordem, mesmo que não sejam aqueles que levam ao melhor resultado económico, que podem ser desejos de 2ª ordem¹³. Honrar os custos afundados pode ser um desejo de 1ª ordem, mesmo para a pessoa que conhece a ineficiência económica inerente a isso. Por exemplo, honrar custos afundados pode ser um desejo que representa uma maneira estratégica de, em alguns casos, honrarmos compromissos ou escolhas que possam levar-nos a uma experiência que consideramos útil. Por exemplo, um caso comum de irracionalidade derivada da não ignorância dos custos afundados é o da compra de um sapato apertado. Podemos supor que S estava em viagem e comprou um bonito e muito caro par de sapatos, mas estava com pressa e não o experimentou, comprou o seu número usual, arriscando-se. Chegando em casa calça-o e percebe que está apertado. S pensa: “custou um mês do meu salário e vou jogá-lo fora?”. De facto, é difícil ignorar os custos afundados nessa situação, mas era o que S deveria fazer. No entanto, S utiliza o sapato como uma forma de auto-punição, isto é, uma forma de

¹³ Esta distinção será desenvolvida nos capítulos 2 e 3. Resumidamente e de modo incompleto, esta distinção refere-se a desejos de duas ordens: a primeira é constituída pelos desejos, conscientes, e não conscientes, que de facto levam o agente a agir; a segunda é constituída por desejos, conscientes, que não coincidem, necessariamente, com os desejos de primeira ordem – estes desejos de segunda ordem não levam o agente a agir.

punir-se a ponto de nunca mais cometer o mesmo erro¹⁴. O curioso é que ainda parece irracional, não? Será que não basta ter perdido o dinheiro e não poder utilizar o sapato? Isto é, será que S já não tem os elementos necessários para incorporar a aprendizagem e nunca mais comprar sapatos sem os experimentar? Provavelmente não! S pode ser um tipo de pessoa, razoavelmente comum, que prefere aprender a partir de um certo sofrimento e por isso impõe-se determinado comportamento, em princípio, irracional. A questão aqui é saber quais conjuntos de desejos e de crenças vamos considerar definidores ou motivadores da acção racional – os de 1ª, ou os de 2ª ordem. Nos capítulos 2 e 3 essa questão será abordada – tentar-se-á argumentar que os desejos e crenças de 1ª ordem são os definidores/motivadores da acção e que, portanto, são estes que devem contar para a definição da racionalidade na acção.

¹⁴ Sobre uma discussão acerca dos custos afundados e da acrasia, veja-se Zilhão, 2009.

1.3.2 RACIONALIDADE NA ACÇÃO: DA PSICOLOGIA POPULAR À FILOSOFIA

Esta secção pretende (re)afirmar que a psicologia popular é a porta de entrada dos filósofos no campo da explicação da acção e na definição da racionalidade na acção (RA) (na secção 1.1, este tema já foi superficialmente tratado), mostrando a continuidade entre a PP e a teoria da acção. Pretende-se além disso descrever a explicação standard da acção e a concepção de racionalidade que está em causa na filosofia da acção.

Embora existam diferentes formas de explicar a acção e de definir a RA, é possível sintetizar os pontos centrais básicos a partir das concepções de acção e de racionalidade de Donald Davidson, que se reportam eles próprios à PP. Estes pontos são partilhados por grande parte dos filósofos da acção. Contemporaneamente, Davidson é a referência para as teorias da acção e da racionalidade. Concordando, ou não, com ele, impõe-se tê-lo como ponto de partida.

1.3.2.1 COMO É QUE AS PESSOAS COMUNS EXPLICAM ACÇÕES?

Todos nós, filósofos, ou não filósofos, explicamos e prevemos as nossas acções e as acções dos outros (veja-se secção 1.1). Fazemos isto de modo natural, não recorrendo a nenhum manual, ou a nenhuma teoria científica ou filosófica da acção

ou da racionalidade. Esta característica é presumivelmente central no sucesso adaptativo da nossa espécie. Conseguimos ainda ter crenças acerca das nossas crenças e acerca das crenças dos outros, assim como sobre os desejos e as intenções. Conseguimos prever mudanças discretas dos desejos dos outros, assim como explicar determinadas mudanças a partir de uma história causal das razões, ou de uma história dos estados mentais do outro. Enfim, somos possuidores de uma habilidade particular que, específica da nossa espécie ou não, nos permite interagir com outros agentes de modo estratégico.

São vários os teóricos que procuram investigar e sistematizar as explicações do comportamento das pessoas comuns. Para servir como base para a análise da explicação popular do comportamento, escolhi o trabalho do psicólogo Bertam Malle (2001 e 2004). Bertram Malle é um psicólogo com uma relação bastante próxima com a filosofia – escreveu, por exemplo, artigos importantes nesta área com o filósofo Joshua Knobe (1997 e 2002), além de ter constantemente em conta filósofos e teorias filosóficas da acção.

Ao investigar a natureza das explicações populares, Malle (2004) identificou as seguintes características essenciais, expressas nos seis postulados das explicações populares do comportamento (folk behavior explanations) que propõe:

(P1) As pessoas distinguem entre eventos comportamentais intencionais e não intencionais.

(P2) Para os eventos comportamentais considerados intencionais, as pessoas utilizam um dos seguintes três modos de explicação: *explicações por razões* (ER), *história causal das explicações por razões* (HCER), ou *explicações pela capacidade* (EC).

- (P3) As ER são aquelas explicações do comportamento que citam as razões do agente para ter a intenção de agir, ou para agir intencionalmente. As razões são conceptualizadas como estados mentais do agente à luz das quais este forma uma intenção para agir.
- (P4) As HCER citam factores que estabelecem o plano de fundo dos motivos da acção, contribuindo assim para explicar o contexto da acção. As HCER não são elas próprias razões, portanto o agente não forma uma intenção à luz ou em razão de factores históricos causais.
- (P5) As EC não clarificam por que razão um agente tem a intenção de agir, mas antes como é possível que a intenção se tenha transformado numa acção realizada com sucesso. Essas explicações pressupõem as razões e a intenção do agente e citam os factores que habilitam a acção para ser realizada como foi intencionada.
- (P6) Para os eventos comportamentais considerados não intencionais, as pessoas utilizam apenas um modo de explicação – explicações causais. Estas explicações citam causas “mecânicas” – isto é, factores que provocaram o evento não intencional sem o envolvimento de raciocínio subjectivo ou da racionalidade.

Vejamos, agora, alguns destes postulados em mais detalhes. Em P1, Malle afirma que as pessoas distinguem entre comportamentos intencionais (acções na terminologia filosófica) e comportamentos não intencionais. Isto parece bastante intuitivo e é a base de tudo o resto acerca da explicação e da previsão das acções. Se não conseguíssemos distinguir entre comportamentos intencionais e não intencionais não teríamos condições para explicar, nem para prever as acções – ao falarmos de comportamentos intencionais já estamos dentro do campo da psicologia popular. Se não conseguíssemos distinguir entre estes dois tipos de

comportamentos, só nos restaria explicar todos os comportamentos através da física ou, no limite, através do behaviorismo. Como vimos na secção 1.1, isto não é plausível.

Em P2, Malle afirma que as pessoas utilizam três tipos básicos de explicações psicológicas populares: explicações por razões (ER), explicações pela história causal das razões (HCER), ou explicações por factor possibilitador (EC). Vejamos agora cada uma delas.

P3 corresponde às ER. Esta é a principal explicação da PP e é aquela que serve de base e origem para a teoria da acção formulada na filosofia. Esta explicação evoca as crenças, os desejos e as intenções. Basicamente, a ideia é que a acção é causada pelas crenças e pelos desejos (estados mentais) através da formação de intenções que levam o agente a agir num determinado sentido e não em outro. Um exemplo simples é o seguinte: João cumprimentou José porque tinha o desejo de ser simpático e tinha a crença que cumprimentar José o levaria a ser simpático. João, assim, formou a intenção de cumprimentar José. João dá a mão a José.

P4 refere-se às HCER. Este tipo de explicação dá-nos uma ligação explicativa às razões e às suas próprias histórias causais, citando factores antecedentes que de algum modo tenham produzido as razões efectivas das acções. Exemplos típicos deste género de explicação são aqueles que trazem à tona a condição cultural, ou económica, do agente, assim como apelam a causas ligadas à infância, ou a algum acontecimento passado, entre outras possibilidades. Quando alguém diz “ José bateu em seu filho Joaquim porque apanhou muito quando era criança”, está a explicar a acção de “bater” através do apelo à história da possível razão que de

facto teria levado José a bater em seu filho, que seria, no final das contas, constituída por, pelo menos, uma crença e um desejo.

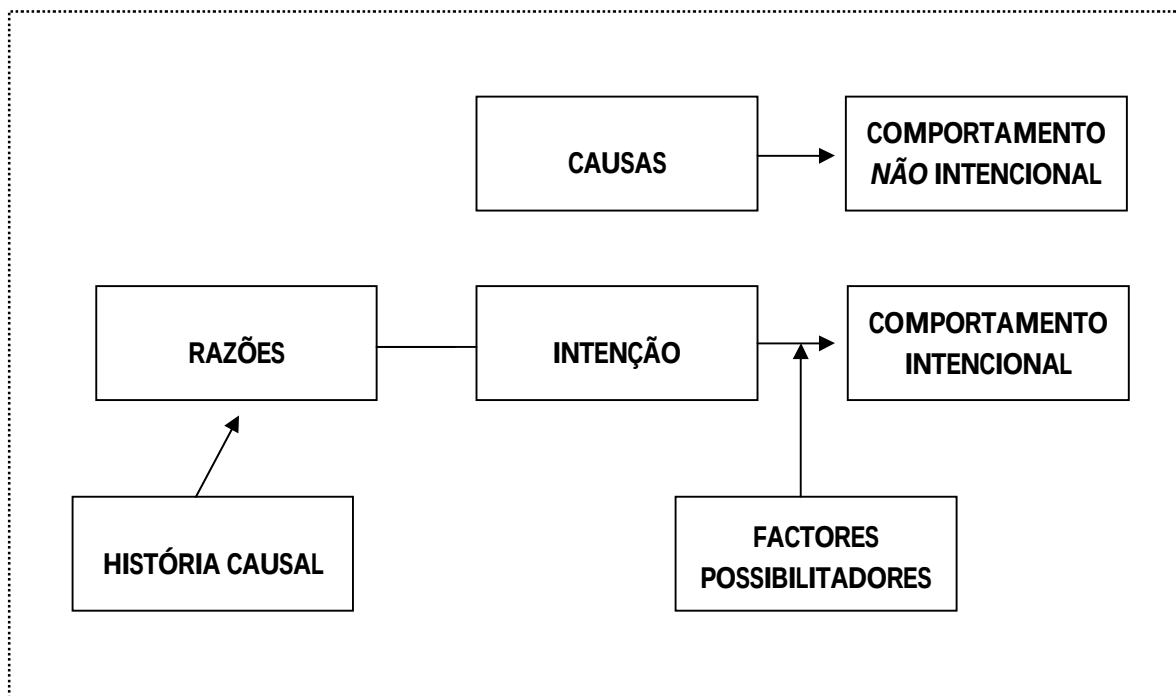
P5 refere-se às EC. Este género de explicações utiliza factores possibilitadores como elementos explicativos. Estes factores, de um ponto de vista psicológico popular, ajudam na transformação de uma intenção numa acção. Se eu digo “Maria trabalhou a noite toda...porque tomou muito café”, ou “José é tão pobre, como é que ele comprou aquele carro? Comprou-o porque ganhou no euromilhões”, estarei a explicar estas acções ressaltando os factores possibilitadores das razões em jogo (crenças, desejos, ou intenções).

P6 afirma que as pessoas comuns explicam os comportamentos não intencionais como se fossem eventos físicos típicos, como o cair de um vaso, ou o cair de uma pedra. As explicações apelam às funções fisiológicas do corpo, ou mesmo às leis da física, por exemplo. Dizemos coisas como: “Enquanto o João dormia, a sua perna foi parar ao chão, pois sem querer ele colocou a perna para fora da cama e a gravidade fez o resto”. Este postulado é também interessante, pois traz alguns problemas específicos. Por exemplo, parece razoável dizer que quando João bate com o joelho no pé da mesa isso é um comportamento involuntário de João. No entanto, parece também plausível dizer que as pessoas comuns poderão imputar a João algum tipo de intenção por “incapacidade”, isto é, João bateu com o joelho, mas tinha todas as condições para não o fazer. Não chega a ser uma explicação intencional, mas não parece ser uma explicação de um comportamento completamente não intencional. Este não é um assunto desta tese, no entanto, parecem-me que podem ser realizados experimentos para testar esta intuição.

É importante ressaltar que tanto em P4 como em P5 as explicações têm no seu núcleo o apelo a razões. No entanto, aquilo que é colocado em primeiro plano é ou a história causal das próprias razões, ou os factores que possibilitam que a intenção, causada por razões, se transforme em acção. Podemos pensar que de uma forma ou de outra, em quase todas as ocasiões, mesmo ao explicar um comportamento directamente por razões, temos em mente a história causal das razões e os factores possibilitadores da acção. Portanto, a essência da explicação psicológica popular continua a ser a explicação por razões. No entanto Malle sugere que ao explicar por razões estamos também comprometidos, algumas vezes, com explicações acessórias.

O Figura 3 (Malle, 1999) abaixo é uma ilustração do que acima foi descrito.

Figura 3: Quatro formas de explicação popular para os comportamentos intencional e não intencional



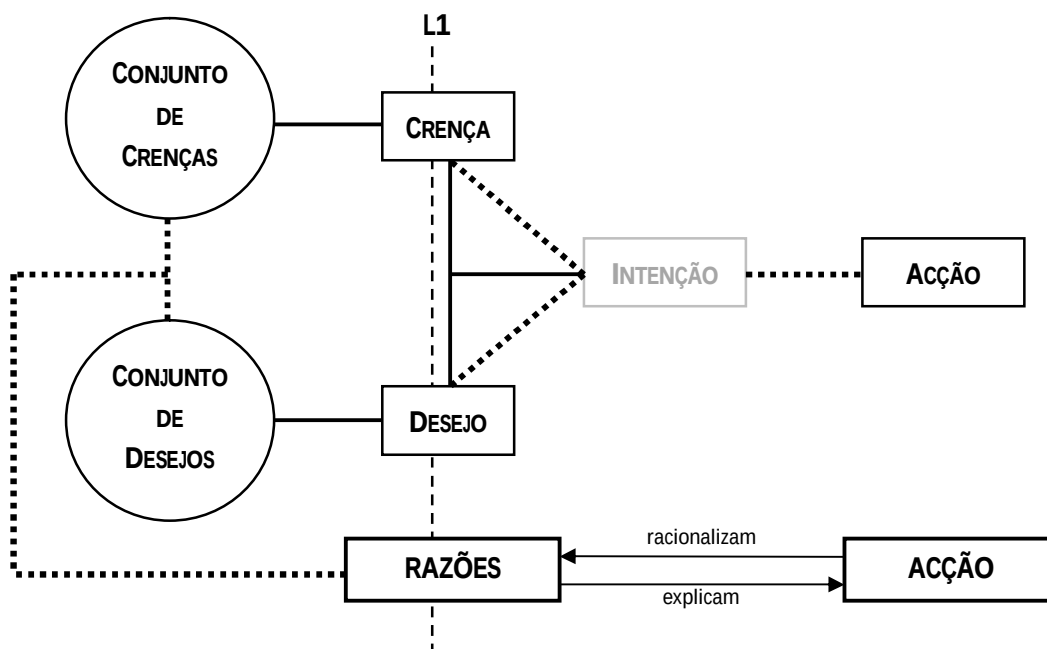
1.3.2.2 FILOSOFIA DA ACÇÃO: COMO É QUE D. DAVIDSON EXPLICA AS ACÇÕES?

De acordo com Davidson, uma acção é explicada se for encontrada a razão pela qual o agente levou a cabo essa acção. Davidson designa uma tal explicação '*racionalização*' (veja-se Figura 4). Uma acção explica-se, então, em termos de razões, mas apenas quando a razão apontada é, em princípio, a razão pela qual a acção foi executada. Isto porque alguém pode ter uma razão para realizar uma acção, levar a cabo essa acção, e ainda assim essa pode não ter sido a razão pela qual fez o que fez. Para ilustrar um tal caso, Davidson (1980, "Freedom to act") apresenta o exemplo de um montanhista que segue à frente de um outro sentindo-se incomodado pelo peso e pelo cansaço que essa situação lhe provoca. A certa altura, a corda solta-se e o segundo montanhista cai. Haveria algumas crenças e desejos que racionalizariam esta acção como acção. No entanto, apenas no caso de essas crenças e desejos terem causado a acção de forma apropriada, i.e. apenas no caso de serem as causas da acção, terá sido essa realmente a razão pela qual a acção foi realizada. Neste exemplo, tal poderia não ter sucedido, e o ocorrido pode nem ser considerado uma acção se não foi intencional. A racionalização davidsoniana pode ser entendida, numa interpretação alternativa, como um recurso metodológico que visa ajudar-nos a distinguir entre um acontecimento qualquer e um acontecimento intencional, ou seja, uma acção.

Uma tal definição ajuda-nos também a compreender a resposta de Davidson ao problema da natureza da acção. Davidson defende que: a) uma acção é um

acontecimento intencional porque tem uma descrição intencional (o que distingue uma acção é o facto de poder ser descrita recorrendo a crenças e desejos); b) uma acção é explicada por, pelo menos, um desejo e uma crença relevante (veja-se Figura 4). Davidson tentará mostrar ainda que a racionalização é uma explicação causal, ou seja, que as razões são as causas da acção – foi isso que fez a originalidade da sua proposta, quando pela primeira vez a avançou.

Figura 4: Explicação da Acção segundo D. Davidson



Concentremo-nos, agora, na questão da explicação da acção através do processo de racionalização. Quando é que, mais especificamente, uma razão racionaliza uma acção, na acepção davidsoniana? Isto acontece quando ela conduz o agente a perceber porque é que essa acção é apelativa – ou porque satisfaz algum desejo seu, ou porque se lhe afigura necessária, ou obrigatória, entre muitas outras hipóteses. Devido ao facto de para uma dada acção poderem, em princípio, ser

encontradas muitas explicações possíveis, é necessário saber qual é a que é de facto o caso. É neste sentido que Davidson afirma que racionalizar a acção é dar conta das pró-attitudes¹⁵ que um agente mostra relativamente a essa acção (desejar, valorizar, considerar imperativo algum dos seus aspectos, etc), assim como da crença que o agente tem de que essa acção vai cumprir aquela que é a sua intenção ao realizá-la. Assim, no conhecido exemplo de Davidson, a acção que eu levo a cabo ao carregar num interruptor é explicada pelo meu desejo de ligar a luz e pela minha crença de que carregando no interruptor eu ligarei a luz.

Racionalizar uma acção é, portanto, trazer à luz o par crença-desejo que torna essa acção inteligível. Dizer que alguém faz alguma coisa por alguma razão é dizer que estão aí implícitos pelo menos algum desejo e alguma crença (veja-se Figura 4). A esse par crença-desejo Davidson chama “razão primária” (veja-se Figura 4, especificamente a linha tracejada *L1*). Na Figura 4, as linhas tracejadas menores indicam a actuação da racionalidade. A partir do quadro que acabei de traçar, Davidson é considerado um dos mais conhecidos defensores do modelo crença-desejo na teoria da acção.

Aprofundando a questão e ainda a propósito da exposição da teoria davidsoniana, alguns aspectos devem ser esclarecidos, para que a minha proposta possa ser entendida na plenitude. Interessa-me particularmente sublinhar o seguinte: o que é que o processo de racionalização, de acordo com Davidson, não é? Racionalizar uma acção não é mostrar que essa acção é ‘racional’ no sentido de estar de alguma forma de acordo com certos padrões ou normas mais ou menos aceites por todos (por exemplo um imperativo moral ‘Não matarás’). Racionalizar

¹⁵ Neste conceito incluem-se coisas tão diversas como desejos, objectivos, valores, convenções sociais, princípios estéticos, etc.

uma acção é exclusivamente torná-la inteligível apresentando os motivos pelos quais ela se afigura desejável aos olhos do agente. Ou seja, uma acção que nos pode parecer totalmente descabida pode, ainda assim, ser racionalizada, no sentido em que pode ser encontrado o par crença-desejo que constitui a razão dessa acção.

Mas, então, em que sentido fala Davidson de racionalidade? Ele acredita ser possível imputar racionalidade aos agentes com base na observação das suas acções. De que forma? Como o próprio atesta, os nossos desejos e crenças não precisam de ser racionais em qualquer sentido normativo-prescritivo para servirem de razões para a minha acção (veja-se Figura 4). O que acontece é que a acção deve ser racional, dados os meus particulares desejos e crenças (Davidson, 1980, “Intending”). Que quer isto dizer? Que se eu tenho um determinado desejo e uma determinada crença com ele relacionada, a minha acção será racional se estiver em conformidade com esse desejo e com essa crença, se for desejável, dados os meus desejos e crenças (o que *não* é equivalente a dizer que constitui o melhor meio para atingir um fim). É por esta via que a conexão entre uma acção e o par desejo-crença que a racionaliza pode ser descrita como uma conexão racional – nos casos em que há uma relação lógica entre o conteúdo das crenças e desejos e a acção de tal forma que esta é considerada desejável pelo agente (veja-se Figura 4 – tracejado menor)

A racionalidade seria, portanto, uma questão de inteligibilidade do comportamento humano como acção ou, mesmo, uma condição para um agente ser tomado como mental (e não tanto uma questão de adequar meios a fins, como nos diz a teoria instrumental da racionalidade). Por exemplo, de acordo com a teoria instrumental, se eu quero matar alguém, realizarei uma acção racional se mobilizar

os meios adequados para atingir esse fim; se, por exemplo, disparar contra essa pessoa. Mas a minha acção será irracional se, em vez disso, fizer algo como oferecer um animal inofensivo a essa pessoa, digamos, um rato. No entanto, se de acordo com o modelo instrumental de racionalidade, esta pode ser considerada uma acção irracional, de acordo com a teoria da acção de Davidson, ela pode ainda ser considerada racional. A razão primária que, nesse caso, explicaria a minha acção seria constituída pelo meu desejo de matar essa pessoa e pela minha crença de que oferecendo um rato a essa pessoa, ela morreria, na medida em que disponho de uma outra crença relacionada com essa – a de que essa pessoa tem um medo terrível de ratos e sofrerá um ataque cardíaco se estiver em contacto directo com um. A minha acção seria racional apenas no sentido em que está em conformidade com as minhas particulares crenças e pró-attitudes: elas tornam aquele extracto de comportamento de um agente inteligível como acção (mesmo que se conte com crenças mal formadas).

É neste sentido que Davidson afirma que a explicação das acções humanas (por oposição ao acto de explicar nas ciências físicas) envolve a noção de racionalidade. Explicar uma acção é dar as razões dessa acção, e isso envolve inseri-la numa rede mais vasta de atitudes e crenças atribuíveis ao agente. A diferença proposta entre nós, animais racionais, e os animais não racionais consiste, então, precisamente na existência em nós de uma moldura coerente de crenças, uma vida mental holista. Nós só conseguimos identificar pensamentos particulares e saber o que eles significam porque conseguimos inseri-los no interior de uma rede de crenças que lhes dão sentido (Davidson, 2004, “Rational Animals”). É por isso que Davidson apresenta uma noção de racionalidade que é

necessariamente holista – isto é, nenhuma atitude proposicional é racional por si só, isso dependerá do conjunto das outras atitudes proposicionais onde ela se encontra incluída.

Dado este enquadramento geral, quais são as ocasiões em que podemos atribuir irracionalidade a um agente, de acordo com o modelo davidsoniano? Neste ponto do trabalho o tema será tratado de maneira mais resumida, já que o capítulo 3 será na sua totalidade, dedicado à acrasia, que é, segundo Davidson, um dos casos paradigmáticos de acção irracional.

Uma acção irracional – caso da acrasia, por exemplo – seria aquela que manifestasse inconsistência relativamente a certas crenças do agente. Mas trata-se de uma inconsistência interna (*inner inconsistency*) relativamente às crenças do agente e não de uma inconsistência que se estabelece a partir de fora, relativamente a padrões externos de qualquer natureza. Mais concretamente, seria um caso em que o agente age contrariamente às razões que ele próprio considera serem as mais relevantes no que toca à acção em questão.

O que se passa é que, num tal caso, o agente vai contra o princípio de continência, segundo o qual o agente deve agir de acordo com o seu melhor juízo, considerando todos os aspectos em jogo, e vai, por isso, contra ele próprio, já que Davidson considera ser esse um princípio racional aplicável a todas as criaturas que têm atitudes proposicionais. Porquê? Porque é apenas interpretando alguém ou alguma coisa como estando de acordo com certos princípios que nós lhe podemos atribuir atitudes proposicionais. A racionalidade é, neste sentido, uma condição para se ter pensamentos. Isto é, é porque consideramos que um agente se comporta a maior parte do tempo de acordo com princípios básicos de

racionalidade tais como o princípio de continência que nós o consideramos como um ser capaz de pensar e, por isso, de agir; de outro modo, não teríamos razões para tal. A racionalidade é, pois, o fundo sobre o qual se pode considerar a irracionalidade – se a maior parte das pessoas não agisse de forma racional a maior parte do tempo, não teríamos nunca razões para lhes imputar irracionalidade de nenhum tipo (Davidson, 2004, “Incoherence and Irrationality”)

No caso da acrasia, o agente age, de forma consciente e intencional, contrariamente ao seu melhor juízo, transgredindo o princípio em causa. Ou seja, ele tem razões para fazer o que faz, por isso age intencionalmente, mas ao mesmo tempo está a par do facto de que, considerando todos os aspectos em questão, há um melhor curso de acção disponível. A irracionalidade surge quando a razão para fazer o que fez suplanta o próprio princípio da continência. É irracional porque essa razão não é uma razão contra o princípio em si (não o põe em causa), mas é usada como tal, e assim o agente contraria-se a ele próprio (Davidson, 2004, “Paradoxes of Irrationality”).

Davidson apresenta os seguintes dois princípios como sendo evidentes: o princípio (P1) de acordo com o qual se um agente é livre de fazer X ou Y, e quer fazer X mais do que quer fazer Y, então ele intencionalmente fará X; e o princípio (P2) de acordo com o qual se um agente julga que seria melhor fazer X do que fazer Y, então ele quer fazer X mais do que quer fazer Y (Davidson, 2004, “How is weakness of the will possible”). Em conjunto, estes dois princípios formam o princípio de continência que seria partilhado por todos os seres racionais. Isto quer dizer que Davidson sustenta como absolutamente óbvio que se um agente livremente efectua qualquer espécie de juízo avaliador que vai no sentido de

escolher um determinado curso da acção, então ele necessariamente agirá dessa forma, se nada o impedir. Caso contrário, pode ser acusado de irracionalidade.

Feitos estes esclarecimentos essenciais relativos à teoria davidsoniana da acção e à forma como a racionalidade e a irracionalidade nela aparecem, passarei em seguida, no capítulo 3, a descrever a minha própria proposta acerca da impossibilidade da acrasia como uma acção (intencional) e irracional. Tentarei mostrar que a acrasia, que parece ser um fenómeno percebido pelas pessoas comuns, é na verdade um fenómeno moral. Noutras palavras, tentarei mostrar que a percepção de uma acção como acrática depende na verdade do valor moral atribuído à acção realizada e, de certa forma, à acção alternativa.

CAPÍTULO 2

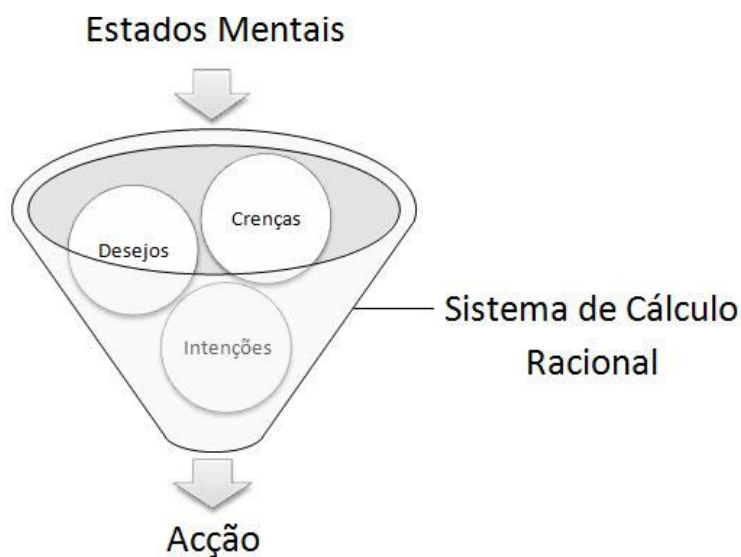
MODELO FUNCIONALISTA DA RACIONALIDADE NA ACÇÃO

Neste capítulo será descrito o modelo funcionalista da racionalidade na acção proposto neste trabalho, levando em conta aquilo que foi dito no Capítulo 1 quanto ao nível explicativo visado. Relativamente às ambições do modelo, afirmei já que se trata de um modelo afastado de quaisquer aspectos normativo-prescritivos, mesmo porque não parece adequado consagrar tais aspectos num modelo. Este é um capítulo simples mas importante do presente trabalho – pretende-se antes de mais objectividade e clareza na descrição do modelo proposto. Tratando-se de um modelo geral da racionalidade na acção (RA) o seu grau de especificidade deve ser o menor possível.

A RA está presente durante o processo de deliberação e durante a realização da própria acção. É a RA que coordena os esforços necessários para que se realize o cálculo deliberativo a partir de estados mentais (como crenças, desejos, ou intenções). Ora, a actuação da RA no sistema de cálculo não é uma actuação num sentido específico qualquer que possa/deva ser generalizado para todos os seres humanos, ou que derive de algum conteúdo moral último/universal. Isto é, quando a RA está em cena, dada a sua função no sistema de cálculo, o resultado esperado será a determinação do curso de acção do agente; os conteúdos morais são inputs como quaisquer outros no sistema de cálculo – como é o caso dos conteúdos morais de crenças e desejos.

Neste sentido, e retomando a Figura 1, parece possível reformulá-la para o formato da Figura 5 (abaixo), onde os inputs são os estados mentais, os outputs são as acções e a deliberação ocorre dentro de um sistema de cálculo alimentado pela racionalidade.

Figura 5



Alguém poderá perguntar “Mas onde estão os juízos morais neste sistema?”, outros poderão dizer que “os estados mentais não podem ser apenas dados, pois produzimos continuamente juízos avaliativos acerca das nossas crenças e dos nossos desejos; sendo assim, o cálculo é também realizado a partir destes juízos”.

A ideia é a seguinte: os juízos morais e outros juízos avaliativos estão contidos no sistema, no entanto não é este quem os produz (veja-se a Figura 6). Isto é, os desejos e as crenças, ou intenções são constantemente avaliados, reavaliados, modificados e produzidos por nós. Parece bastante razoável dizer que os avaliamos moralmente também. No entanto, esta avaliação não ocorre dentro do sistema de cálculo (SC), pois este tem uma função apenas operacional (veja-se Figura 6). Este sistema tem, por natureza, que ser o mais eficiente possível, isto é, tem que permitir a entrada máxima possível de inputs e produzir o(s) output(s) no instante em que o agente necessita – tendo como condicionante, por exemplo, a restrição de tempo deste agente. É um sistema que procura, naturalmente, a maior produtividade, a maior eficiência e maior eficácia possíveis – fazer mais em menos

tempo, fazer mais com menos custos para o sistema específico (SC) e para o sistema geral (agente) e fazer tudo isto de forma a concretizar a intenção do agente (independente do grau de consciência que o agente tem acerca da sua própria intenção eficaz).

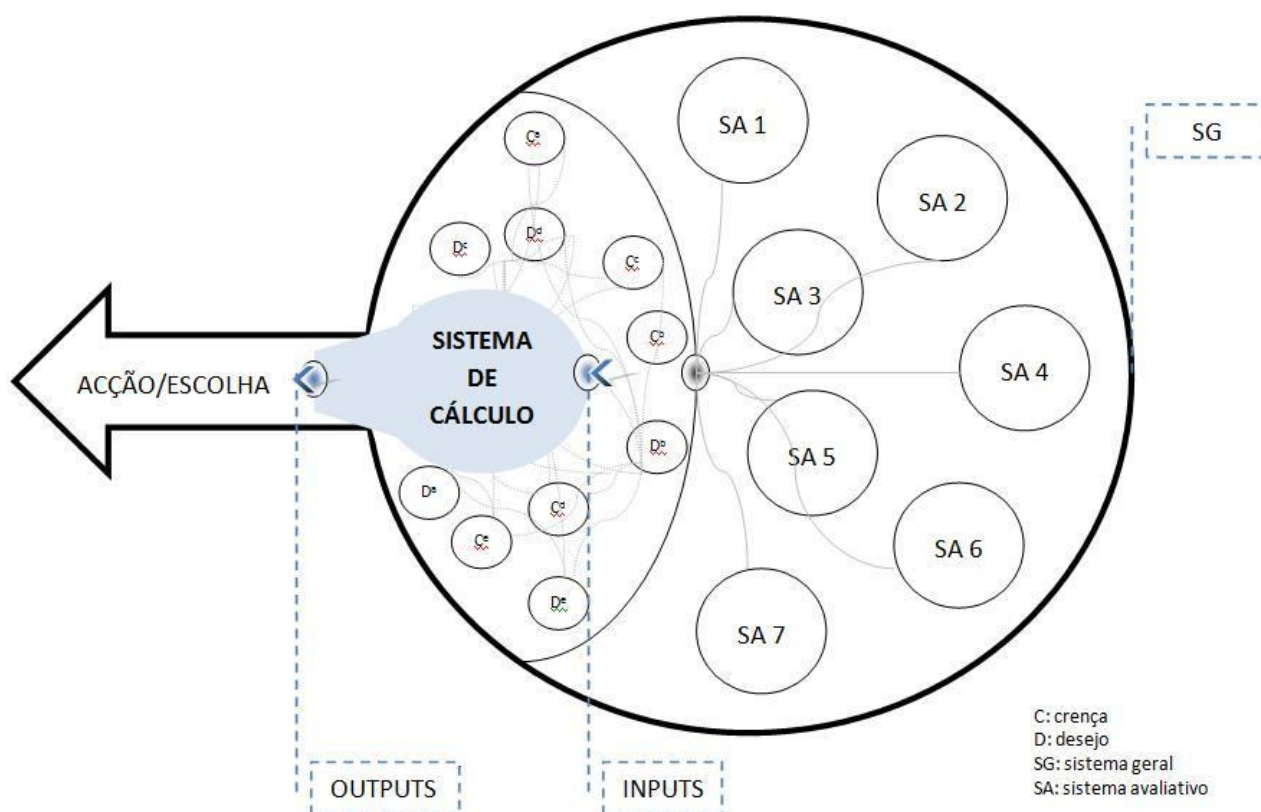
É razoável pensar que um sistema deste género opera de maneira complexa mas com dados simples. Por exemplo, o sistema tem que operar com crenças, com desejos, ou intenções que sejam similares a dados simples, isto é, uma “crença de que X” e um “desejo W”, e não uma “crença de que X, se W em T_1 ”, ou “um desejo W, se a crença de que X em T_3 ligar-se ao desejo V”. Estes são exemplos um bocado lineares, mas parecem servir como ilustração daquilo a que chamei dados simples. Estas conexões que, por exemplo, incluem o tempo e mesmo as relações entre crenças e/ou desejos são realizadas pelo próprio sistema. Operar de maneira complexa mas com dados simples, parece ser uma condição necessária, na medida em que parece plausível afirmar que o SC *não* pode depender de processos reflexivos complexos. Para que a acção ocorra, será necessário (e suficiente) que o sistema produza um output que conduza eficazmente o agente a agir. Para isto o SC tem que ser capaz de, dadas restrições externas tais como o tempo e outros recursos, e sem se fechar, produzir o output (acção). Se o próprio sistema produzisse elementos complexos e processos reflexivos, a deliberação acerca de algo simples como “que vinho beber?” poderia levar o agente à exaustão sem o levar à acção.

Por um lado, o sistema não pode fechar-se: parece intuitivo e plausível afirmar que estamos sempre a avaliar as nossas crenças e os nossos desejos, mesmo que não o façamos de modo consciente, ou completamente consciente. Isto parece

acontecer tanto numa acção relativamente longa tal como “pintar uma porta”, como também numa acção mais curta, tal como “tomar um copo de vinho”. Se o sistema se fechasse, seria o mesmo que afirmar que, num dado momento, não haveria *inputs* e nem a racionalidade operaria no sistema. Isto é, evocando de novo a analogia com o sangue, seria como se todos os estímulos parassem e o sangue parasse de circular de repente, por causa de algum processo do corpo.

Portanto, o sistema geral (agente) tem que agir com o SC aberto. Para isto acontecer, o próprio SC tem que ser capaz de levar o agente a agir num instante específico, dadas restrições internas e externas, considerando o estado actual dos processos avaliativos e reflexivos mais completos do sistema geral menos o sistema de cálculo (veja-se a figura 6). Estes processos avaliativos formam as crenças, os desejos, ou as intenções que estão presentes, ou que entrarão no SC. Assim, garante-se a permanente actualização do SC.

Figura 6



A figura 6 tenta mostrar, numa perspectiva sistémica, que a racionalidade na acção, que opera no interior do sistema de cálculo, é responsável pela escolha de crenças, desejos, ou intenções. Assim, nenhuma acção que seja intencional deixa de passar por este sistema de cálculo (deliberativo), seja uma acção intencional com uma deliberação aparente, seja uma acção intencional sem deliberação aparente. É, portanto, este o sistema que confere intencionalidade ao comportamento.

Pode dizer-se alternativamente que os sistemas avaliativo e de cálculo constituem um único sistema com duas funções, caracterizado pelo constante feedback/retroalimentação. No entanto, parece-me que o sistema de cálculo é aquele que, de facto, delibera. Assim sendo, parece útil identificar dois sistemas diferentes, mesmo que estejam intrínseca e directamente ligados.

É importante sublinhar que as crenças, os desejos, ou as intenções que estão no sistema de cálculo não precisam de ser conscientes para serem eficazes (levar o agente a agir). Isto é, este sistema fornece ao agente outputs derivados de desejos, crenças, ou intenções mais ou menos conscientes. A maneira como este processo pode intervir na explicação da acção será melhor descrita no capítulo 3.

Como vimos no capítulo 1, a TER explica a acção a partir dos conceitos de preferência, restrição orçamental e de escolha, mas o essencial é que o agente possui um ranking de preferências transitivas, que está de certa forma disponível para o agente a qualquer hora e em qualquer lugar. Isto é, é como se tivéssemos uma colecção de deliberações cujos resultados não são acções mas sim preferências gerais (desejos hierarquizados e estáveis). Num caso concreto de escolha/acção, o agente realiza um cálculo que leva em consideração o seu ranking

de preferências e a probabilidade (subjectiva, na maioria dos casos) de conseguir o resultado esperado com as suas escolhas em jogo. Assim, o agente escolheria a melhor combinação possível entre os desejos listados (ranking) e as probabilidades associadas à realização dos resultados esperados.

O problema é que olhando para um caso individual qualquer, parece muito pouco plausível que tenhamos à nossa disposição um tal ranking estável. Teríamos que aceitar que os nossos desejos não só são estáveis, para configurar um ranking ao longo do tempo, como, uma vez produzido o ranking, teríamos disponível um conjunto de valores associados aos nossos desejos que nos dão a possibilidade de associá-los às probabilidades de ocorrência destes mesmos desejos. Este mecanismo parece pouco dinâmico, pouco concordante com a intuição de que somos seres com uma psicologia dinâmica que nos faz alterar gostos e crenças com certa constância e rapidez.

A proposta que procuro avançar não afasta por completo a ideia de que possuímos algo parecido a um ranking de preferências (desejos). O que parece acontecer é que dada uma certa estabilidade das fontes de formação e avaliação de crenças e de desejos, temos a tendência a repetir escolhas anteriores em situações semelhantes. No entanto, qualquer modificação nos factores de escolha (crença, desejo, ou intenção) é o suficiente para mudar racionalmente o output (acção).

Frequentemente, quando dizemos algo como “que estúpido eu sou...podia/devia ter feito X e fiz Y” identificamos este agir como sendo o produto de uma falha de racionalidade. No entanto, de acordo com o que tem aqui sido defendido, a acção de fazer Y é racional, pois foi produzida pelos inputs disponíveis e pelo sistema de cálculo. No entanto, fazer Y e não X pode ter sido fruto da

ausência do input adequado na deliberação, como também dizer que “podia/devia ter feito X e fiz Y” pode ser simplesmente uma forma de racionalizar a acção no sentido psicanalítico, isto é, dizer que queria fazer X e não Y, mas que “algo” o fez agir de outra maneira (esta possibilidade será minimamente discutida no capítulo 3)

Vale a pena lembrar que a racionalidade na acção não leva em consideração a formação mais ou menos racional de crenças e, no limite, de desejos. Assim, dado um sistema racional individual, os outputs dependem dos inputs. Isso é diferente de dizer que os outputs dependem da racionalidade que deveria seleccionar os melhores inputs.

Não vou avançar por aqui, mas penso que é razoável afirmar que tendencialmente as pessoas que têm alterações mais constantes no modo de vida, seja ao nível fisiológico, seja ao nível cultural, têm maior inclinação para alterar crenças e desejos, não porque os alterem directamente mas sim porque as condições de avaliação e formação das crenças e desejos são outras. Ora, o que é que isto nos conduz a pensar no caso de uma intervenção, em termos de políticas públicas, num indivíduo? No caso limite de uma intervenção pública num indivíduo, o policy maker deverá levar em consideração não só os desejos e o suposto ranking de preferências revelado pelo indivíduo ao longo do tempo, mas também, principalmente, o sistema aqui apresentado em que a questão central não é a racionalidade em si, mas sim como produzir melhores resultados (outputs) a partir da modificação dos inputs, tendo em mente que a racionalidade na acção do indivíduo está dada. Talvez seja possível através da psicologia encontrar formas para modelar este sistema individualmente, mas este é um trabalho posterior.

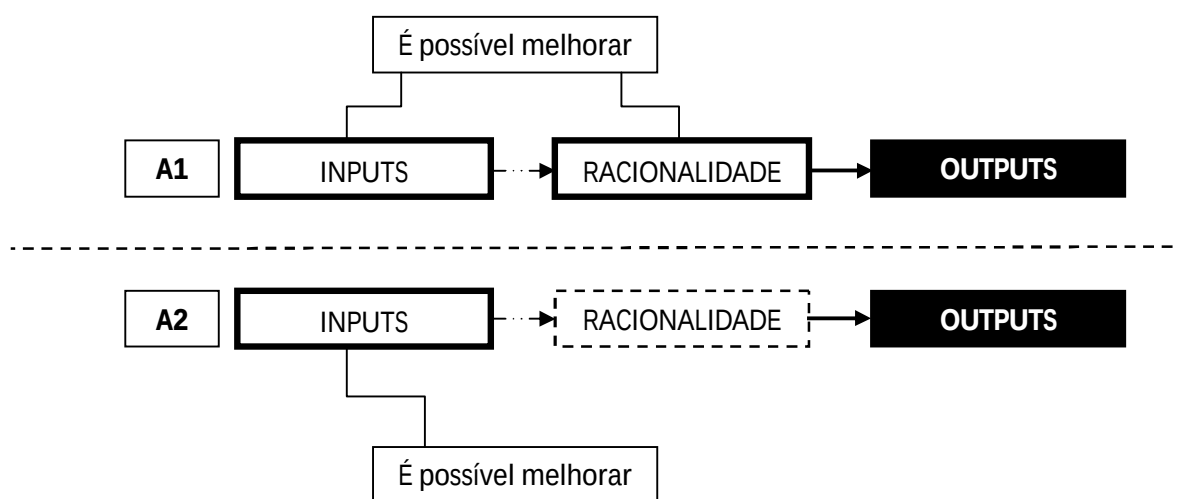
Alguém poderá dizer algo como: “se ensinarmos lógica às pessoas, elas vão raciocinar melhor e portanto vão produzir melhores resultados a partir dos inputs que lhes dermos”. Esse parece ser de facto o caso, no entanto o que aí teria acontecido seria não uma modificação da racionalidade na acção, mas sim uma modificação nas condições para a produção e avaliação de crenças e de desejos. Se ensinarmos mais música, mais história, mais filosofia – ou se, ao contrário, decidirmos acabar com as escolas, ou com a televisão, ou com a internet; enfim, qualquer coisa que façamos neste sentido – mudaremos as preferências, a partir da mudança das condições de formação dos desejos e das crenças.

Na figura 7, abaixo, procuro mostrar as duas abordagens aplicadas básicas do que acabei de dizer. A abordagem A1 é aquela na qual tentamos actuar sobre os inputs e também sobre a racionalidade do agente. Segundo o modelo aqui defendido, a actuação sobre a racionalidade não surtirá efeitos directos, isto é, tentar modificar a racionalidade na acção de um indivíduo não terá o efeito desejado sobre a racionalidade, mas deverá produzir algum efeito, ou alguma modificação nos inputs.

A abordagem A2 é a ilustração da aplicação aqui defendida. Isto é, a actuação da intervenção é apenas no input e não na racionalidade. Utilizando novamente a analogia com o sangue, pode ser que um dia, se a psicologia chegar lá, possamos dizer que a racionalidade no indivíduo S é R^{3+} , em que $3+$ significa algum tipo de classificação aproximada (nunca exacta, pois estamos falando de mentes, de estados mentais e não de cérebro e processos neuronais) do “tipo” de racionalidade na acção do indivíduo S. Isto dar-nos-ia maior eficácia na tentativa de variação dos inputs de modo a alterar os outputs no sentido desejado.

Esta é a estrutura esquemática do modelo funcionalista da racionalidade na acção que eu proponho. Como pano de fundo da minha proposta está a defesa do desenvolvimento de uma psicologia de intervenção, seja individual seja social, que leve em consideração o modelo. A partir daqui é necessário refinar filosoficamente o meu argumento.

Figura 7



CAPÍTULO 3:

ACRASIA – UM CASO DE IRRACIONALIDADE NA ACÇÃO?

3.1 ACRASIA: INTUIÇÕES COMUNS ACERCA DO CONCEITO

De que estamos a falar quando falamos de acrasia? Quando falamos de acrasia, estamos aparentemente a falar daquilo que normalmente as pessoas identificam como “pouca força de vontade”, ou como “falta de força de vontade”. Mas estaremos a falar de um problema moral, ou de um problema estritamente ligado à acção, comportamento e auto-controle, envolvendo nomeadamente a manutenção de uma decisão até ao final da acção?

Parece-me importante aqui colocar explicitamente uma questão acerca de método. Será que o filósofo deve investigar: (i) aquilo que ele considera acrasia levando em conta a sua observação, (ii) aquilo que os filósofos, em grupo, consideram ser a acrasia, (iii) aquilo que o agente real identifica como acrasia, ou (iv) uma combinação qualquer das três opções anteriores? A resposta foi em parte avançada no item 1.1. No entanto, quero tratar aqui especificamente da acrasia, procurando justificar preliminarmente a relevância do conceito popular para a análise levada a cabo no item 3.4.

Parece plausível, respondendo às perguntas acima, dizer que nos deparamos quotidianamente com acções nas quais os agentes parecem falhar de algum modo o cumprimento das suas intenções, dos seus desejos e até parecem loucos por contrariar as suas crenças acerca do mundo. Seja nas acções de outros ou nas nossas próprias acções, temos sempre a sensação de que “isto” está por todos os lados, que esta “doença”, este mal, não nos dá trégua – está sempre lá para nos sabotar.

Quantas vezes não dizemos para nós mesmos: “Hoje comerei esse bolo, mas, a partir de amanhã, começarei uma dieta ... tenho que emagrecer, estou gordo e preciso cuidar da minha saúde”? Acreditamos genuinamente que o melhor a fazer, para ter satisfeito o desejo de uma saúde melhor, é comer uma dose de rúcula e cenoura, um bom prato de legumes cozidos no vapor e, no máximo, um filete de peixe, ou de peito de frango. Na sobremesa, se houver, uma fruta. Porém, sentamos à mesa do restaurante e, entre um rissole de leitão e outro, pedimos os tradicionais rojões à moda do Minho e tomamos uma, ou duas, latas de Coca-Cola Light. Na sobremesa comemos uma taça de baba de camelo. Saindo do restaurante, com aquela sensação de ter comido uma bola de futebol (cheia), dizemos para nós mesmos: “Epá, essa foi a última vez... não faço mais isso...”. Dizemos essa frase, aceitamos o compromisso e passamos a ficar mais tranquilos. Parece que tudo volta ao normal. Parece que no momento em que nos comprometemos a não fazer mais isso, mesmo já tendo professado este compromisso 8.737 vezes nos últimos anos, voltamos ao normal e passamos, naquele instante, a ter a boa sensação de que estamos a coordenar melhor os desejos de curto e longo prazos com as nossas crenças actuais – pelo menos no que concerne aos desejos e crenças conscientes. As coisas parecem fazer sentido, parece que “recobramos a nossa racionalidade”, o nosso bom raciocínio e sentimo-nos, por isso, moralmente mais adequados, pois pretendemos não mais sucumbir à tentação do mal. Alguém poderá dizer que o glutão tem algum tipo de compulsão, ou algum tipo de vício que o impede de manter a palavra, sendo assim, ele não agiria livremente na próxima vez que comesse em excesso. Poder-se-ia, então, para tentar evitar o caso de uma possível compulsão, considerar o exemplo do desorganizado – aquele que adia a arrumação

do escritório de casa por toda a vida, mas tem sempre o compromisso de, no próximo fim-de-semana, ou no próximo feriado, arrumar e manter tudo bem arrumado. A questão a responder aqui é a seguinte: como podemos designar (adjectivar) estas duas acções, comer em excesso e adiar a arrumação? Serão acções acráticas?

Por outro lado, parece-me também comum um tipo de acção que se caracteriza pela dúvida que sente o agente acerca do sentido da acção. Por exemplo, podemos decidir, depois de uma longa deliberação, que vamos à praia com os amigos, mas, no momento em que descemos à garagem para pegar o carro, desistimos e afinal ficamos em casa a estudar para o exame de segunda-feira. Esse é um exemplo caricato, mas podemos pensar em algo mais comum e plausível: José está num outro país, numa fila para comprar um bolo e uma chávena de café e decide, depois de uma rápida deliberação, comprar um bolo de *rerda*, fruta típica da *vostalândia*. Tinha visto um bolo de chocolate, muito parecido com o da sua avó, mas, mesmo assim, decidiu-se pela *rerda*. Com muita firmeza e altivez permaneceu na fila até a sua vez chegar, quando, então, disse hesitante: “Por favor, quero um café e um bolo de...(pausa ligeira)...chocolate”. Nesse caso, a gula não é um elemento importante, pois, nessa história, José tem 1 metro e 90 centímetros, pesa 80 quilos e vai ao ginásio cinco vezes por semana...não precisa de dieta, tampouco pensou nisso ao escolher. Novamente, como podemos designar (adjectivar) estas duas acções, não ir à praia e comer bolo de chocolate? Serão acções acráticas?

Poderia dar muitos mais exemplos, e ao longo deste capítulo introduzirei ainda outros. Estes primeiros exemplos são no entanto suficientes para mostrar intuitivamente uma parte das minhas inquietações. Estas podem ser resumidas

assim: quando dizemos que uma acção é acrática estamos a referir-nos a qual tipo de acção? O que caracteriza uma acção desse género? A acção do glutão, ou do desorganizado é do mesmo tipo que a acção do estudante, ou do viajante? É o filósofo quem define, ou deve definir, o que é a acrasia, ou podemos dizer que o fenómeno sempre esteve lá e nós apenas lhe demos um nome? Devemos (os filósofos) procurar junto das pessoas comuns a resposta à pergunta ‘o que é a acrasia’? (refiro-me à procura de uma definição mais realista da acrasia). Será que o que, para já, chamamos acrasia tem alguma relação com algum tipo de acção psicologicamente possível?

Proponho que a melhor forma de definir, delinear ou descrever um objecto de investigação com natureza psicológica como este é observar o agente real e procurar compreender o que está em jogo quando este identifica um caso de *acrasia* ou *incontinência*, ou, mais vulgarmente, *fraqueza da vontade*¹⁶. A minha proposta tem as seguintes razões: como poderíamos justificar filosoficamente que um determinado conceito filosófico é contrário ao que indicam as nossas primeiras evidências? Neste caso, as principais evidências parecem estar do lado das pessoas comuns. São elas que nos mostram o que é a acrasia, como e quando a sensação de agir acriticamente acontece. Se queremos que a investigação tenha relevância para alguém além de um pequeno grupo de filósofos profissionais, temos que, nesse caso, perguntar (observar) às pessoas comuns e testar as nossas intuições. É claro que se confiamos na introspecção como a nossa melhor fonte e talvez a única disponível ao filósofo, ou se achamos que a filosofia não precisa de ter relevância

¹⁶ Nesse trabalho, assumo que a acrasia, a fraqueza da vontade e a incontinência são termos intercambiáveis. Alguns autores procuram, por exemplo, distinguir a acrasia da fraqueza da vontade, mas, nesse trabalho, essa distinção/alternativa não parece especialmente útil.

social, não teremos aqui problema algum. Poderemos dizer que a acrasia é qualquer coisa, desde que eu e mais algumas pessoas aceitemos que a acrasia é isto ou aquilo e aceitemos que o autor deu boas razões, ou justificou bem as suas opiniões. Em alguns casos, esse é o trabalho adequado ao filósofo, mas neste caso não creio que o seja. Antes de ser um objecto de estudo da filosofia, a acrasia é um fenómeno psicológico e como tal parece obrigar a um tratamento híbrido, através do qual o filósofo atravessa algumas fronteiras em direcção da psicologia.

Poder-se-ia objectar que a possibilidade psicológica da acrasia não é um problema do filósofo. Há três anos, num colóquio, ao apresentar um artigo sobre a acrasia, acusaram-me de estar preocupado com a possibilidade psicológica da acrasia e da racionalização davidsoniana. Afinal, isso não é trabalho, nem uma questão relevante, para o filósofo. Será que isso é verdade? Será que posso por exemplo afirmar, sem consequências graves para a plausibilidade da investigação filosófica, que uma acção intencional pressupõe que o agente possua razões conscientes que racionalizam a acção? Será que não é plausível defender que uma acção intencional se pode dar mesmo que as razões para agir sejam em parte não conscientes? Quando falamos da acrasia como um tipo de contradição entre a acção e o par crença-desejo consciente que racionaliza a acção, não estaremos a simplificar exageradamente o fenómeno? Não estaremos a definir a acrasia em termos puramente operacionais, isto é, não estaremos a dizer que a acrasia é possível independentemente da natureza psicológica dos agentes?

Parece haver algum desconforto na filosofia da acção quando temos que dizer algo sobre se as razões para agir têm que estar, ou não, presentes à consciência no momento da acção. É muito comum admitir-se que de facto existem razões para

agir que são inconscientes, mas os filósofos tendem a circunscrever sua investigação a decisões/acções que, como disse Broome (2007a: 219, 220), são deliberadas e conscientes. A questão, nesse caso, seria: As acções deliberadas são necessariamente realizadas através de razões conscientes? Isto é, não é possível haver algum tipo de deliberação eficaz, que nos leve à acção, que seja realizada com crenças e desejos em parte não conscientes? Vou tratar, de uma forma não exaustiva, essa questão no item 3.3.

Sobre esta questão, aproveitando material de um trabalho em curso, posso citar¹⁷ o que disse António Zilhão (Mauro & Cadilha, no prelo):

(...) as acções têm que estar conectadas de determinado modo com razões, e as razões têm que ser conteúdos de pensamentos. Agora, a questão da presença à consciência ou não presença à consciência já depende do modo como se define pensamento. Se tivermos uma definição cartesiana de pensamento, não pode haver acções que não estejam associadas a razões conscientes porque, deste ponto de vista, qualquer razão tem necessariamente que ser consciente. Mas, evidentemente, podemos ter uma concepção não cartesiana de pensamento. Ora, em contextos teóricos não cartesianos, o facto de um conjunto de itens de vida mental não ser consciente não quer automaticamente dizer que ele não possa constituir pensamento e, constituindo pensamento, que não possa estar naquela relação particular de consequência com o comportamento desempenhado pelo agente que nos permita classificá-lo como um conjunto de razões. Portanto, dentro de algumas destas concepções não cartesianas, pode ter-se pensamentos não conscientes que são razões. Nestas condições, o comportamento deles decorrente é uma acção.

Outra questão importante aqui é a seguinte: para mim, a acrasia faz mais sentido como um fenómeno moral e não meramente como uma falha operacional

¹⁷ O interesse dessa citação é o de mostrar que a questão das razões não conscientes é algo que devemos considerar quando falamos, por exemplo, da acrasia. O texto citado faz parte de um projecto no qual estou envolvido, juntamente com Susana Cadilha, chamado *Diálogos sobre a Acção Humana* que terminará com a publicação de um livro com o mesmo nome. Essa publicação está programada para 2010.

na acção (irracionalidade). Assim como Platão e Aristóteles, as pessoas comuns identificam a fraqueza da vontade com uma acção moral negativa (isso será discutido no item 3.2). Essa é uma intuição forte e que não deve ser descartada. Parece que as pessoas comuns ligam a acrasia a aquilo que podemos chamar de apetites, mais especificamente aos apetites que “devemos” evitar por razões morais.

Para responder às questões levantadas até aqui, começarei por tentar mostrar que existem definições relevantemente diferentes entre si do conceito de acrasia, que podem levar-nos por caminhos bastante diferentes e a resultados mutuamente exclusivos.

Parece-me que o fenómeno da acrasia nos agentes reais foi melhor captado por Platão e por Aristóteles e que, de alguma forma, o conceito actualmente utilizado não capta o fenómeno que atinge o tal agente real, transformando-se num mero exercício teórico pouco fértil e importante.

3.2 [BREVEMENTE] PLATÃO, ARISTÓTELES, DAVIDSON ACERCA DA POSSIBILIDADE DA ACRASIA

Quase todos os textos descritivos sobre a acrasia mencionam, pelo menos parcialmente, aquilo que vou apresentar na primeira parte deste capítulo. Mesmo considerando duvidosa a utilidade, em termos argumentativos estritos de trazer à tona alguns dos usos históricos do conceito de acrasia, penso que é útil tornar explícito aquilo que, para os meus objectivos filosóficos, parece ser o eixo central

do debate sobre a acrasia. Este é formado, essencialmente, pelos seguintes filósofos: Platão, Aristóteles e D. Davidson. O meu primeiro objectivo aqui é mostrar a mudança de área, de abordagem e, consequentemente, de definição do conceito de acrasia – passando da filosofia moral para a filosofia da acção. Uma vez que vou introduzir as concepções de Platão e de Aristóteles, devo dizer que isso não significa que tenha realizado uma extensa e profunda investigação acerca dos sistemas filosóficos propostos por esses filósofos. O meu foco foi apenas o tema da acrasia, sem tratar sistemática e historicamente os autores eles próprios.

Pretendo, apesar disso e precisamente por não ser um especialista em filosofia antiga, dar ao leitor a oportunidade de refutar as minhas conclusões acerca do que penso ser o problema da acrasia para Aristóteles e para Platão, remetendo para fontes directas e explícitas. Para evitar o excesso, circunscrevi as citações às passagens que mais claramente dão origem às minhas afirmações mais relevantes.

É inútil tentar escrever algo breve que seja, ao mesmo tempo, importante e original acerca dos autores e temas que se seguem. No entanto, é possível escrever um breve resumo das principais ideias deles sobre a acrasia (com algumas raras análises). O que se segue terá a função de enquadrar minimamente o debate desenvolvido nesta tese em alguns parâmetros teóricos e históricos.

3.2.1 PLATÃO

A “*akrasia*” é um termo grego clássico formado a partir de um prefixo de negação e *kratos*, que significa “força” ou “poder”.

No Protágoras, Platão não utiliza a palavra *akrasia*, mas trata do fenómeno que, posteriormente, será tratado criticamente por Aristóteles – já com a utilização desse termo.

Sócrates diz a Protágoras (Platão, Protágoras, 352d):

...sabes que a maioria dos homens não se deixa convencer nem por mim nem por ti; dizem antes que muitos, sabendo o que é melhor, não querem agir desse modo, quando podiam fazê-lo, mas agem de outra maneira. E que aqueles a quem eu tenho perguntado qual a razão desse comportamento respondem que são obrigados a fazer o que fazem, dominados pelo prazer, pela dor, ou por qualquer desses outros sentimentos de que eu falava ainda agora.

Nessa passagem, Platão lança aquela que será uma das mais duradouras e, a meu ver, mais interessantes discussões filosóficas acerca da *akrasia*: será a *akrasia* realmente possível?

Platão continua o texto com Sócrates e Protágoras a simularem um diálogo no qual tentam convencer aqueles que têm a sensação de serem dominados pelo prazer de que estão enganados. Os “dominados pelo prazer” perguntariam: “Protágoras e Sócrates, se essa tal sensação não é ser dominado pelo prazer, é então o quê?” (Platão, Protágoras, 352d). É comum, em muitos artigos acerca da *akrasia*, dizer-se simplesmente que para Platão a *akrasia* era impossível e que tal acção/erro era provocado pela ignorância. Para tentar evitar essa simplificação,

muitas vezes excessiva, encontra-se abaixo um esquema no qual procuro, de maneira resumida, expor a resposta de Sócrates a essa pergunta na sequência em que aparece no Protágoras.

- (1) As pessoas muitas vezes dizem que são dominadas pelos prazeres (ex. “da comida, da bebida, do sexo”) e agem de acordo com eles, mesmo sabendo que eles são prejudiciais.
- (2) Essas acções são prejudiciais porque no futuro produzem resultados negativos tais como “doenças e pobreza e muitas outras desgraças semelhantes”, e não porque dão prazer momentâneo. Isto é, é o resultado futuro que é prejudicial, não o prazer presente.
- (3) Essas acções prejudiciais são más porque “culminam em desgraça e privam o homem dos outros prazeres”.
- (4) Existem coisas “boas penosas”. Por exemplo: os resultados provenientes do “exercício físico, das campanhas militares, do tratamento feito pelos médicos, através de cautérios, amputações, medicamentos e privação de alimentos”.
- (5) Essas coisas são boas porque culminam no futuro em prazer “libertando e prevenindo de dores”.
- (6) Em geral as pessoas “procuram o prazer porque é bom e fogem da dor porque é má”.
- (7) O prazer é mau quando “priva de prazeres maiores que os que ele próprio proporciona, ou quando culmina em dores maiores que esse mesmo prazer”.
- (8) A dor é boa quando “liberta de dores maiores que essa ou quando culmina em prazeres maiores que as dores”.
- (9) O bom é o prazer e o mal é a dor.
- (10) “(...) digamos que o homem que sabe que as más acções são más as pratica na mesma. E se alguém, então, nos perguntar: ‘Por que razão?’, responderemos: ‘Porque está dominado?’, Por quem?’, perguntar-nos-á ele. Nós, aí, já não vamos dizer que pelo prazer – é

que o outro nome que substitui esse prazer é o bom. Assim, se respondermos dizendo 'Dominado', ele perguntará: 'Por quem?', e nós, por Zeus, responder-lhe-emos que pelo que é bom" (Platão, Protágoras, 355c)

- (11) O agente, de posse de todas as informações, escolherá sempre o que for bom, isto é, aquilo que lhe der mais prazer, ou, mesmo que opte pelo que é mal, o fará por ser esse mal menor do que o mal futuro. Também é curioso perceber como Platão inclui o tempo neste cálculo.

Este esquema mostra-nos, até aqui, uma clara semelhança entre o pensamento de Platão e a análise custo-benefício na TE18. Permito-me aqui, para ilustrar a referida semelhança, uma citação mais longa.

"(...) se alguém disser: 'Mas, Sócrates, há uma grande diferença entre prazer momentâneo e o prazer e a dor que vêm com o tempo', eu, pela minha parte, responder-lhe-ia que não, decerto, por outra razão senão por serem prazer ou dor. Não há mesmo outro motivo! Mais, é como se um homem bom em pesagens, somando prazeres com prazeres e somando dores com dores, depois de ajustar na balança a proximidade e a distância, disser quais são maiores; porque se pesares prazeres com prazeres terá de aceitá-los sempre como de maior tamanho e em maior número, se forem dores com dores em menor número e em menor tamanho. Agora, se forem prazeres com dores, se os prazeres as excederem, seja a proximidade menor que a distância ou a distância menor que a proximidade, terás de agir segundo o que estes ditarem. Se forem as dores a exceder os prazeres, não terás de o fazer. 'Não é assim, meus amigos?', perguntarei eu. Sei bem que não poderiam responder outra coisa." (Platão, Protágoras, 356bc)

- (12) O tratamento (cálculo) de todas essas variáveis depende do conhecimento.¹⁹

¹⁸ Sobre a relação custo-benefício, veja o item 1.3.1.

¹⁹ Platão fala também do comedimento, mas escreve que "(...) já que se trata de comedimento, suponho que forçosamente será também uma arte de conhecimento" (Platão, Protágoras, 357b)

(13) Errar por falta de conhecimento é agir por ignorância.

(14) Para concluir, Sócrates diz:

“Ora, diziam vocês que o prazer, muitas vezes, supera até o homem que possui conhecimento, e, quando nós discordamos, nessa altura, perguntaram-nos: ‘Protágoras, e Sócrates, se essa situação, o ser dominado pelo prazer, não existe, então é o quê, que lhe chamam vocês? Digam lá.’ Se vos tivéssemos respondido, logo de imediato: ‘Ignorância’, ter-se-iam rido de nós. Mas agora se rirem de nós, estarão a rir-se também de vocês próprios, uma vez que concordaram que quando se erra na escolha em matéria de prazeres e dores – ou, o mesmo será dizer, de coisas boas e más – se erra por falta de conhecimento (...)” (Platão, Protágoras, 357c,d)

Portanto, para Platão, a acrasia é impossível e quando as pessoas dizem que “o prazer, muitas vezes, supera até o homem que possui conhecimento” aquilo a que se referem é resultado da ignorância, da falta de conhecimento. Essa não é uma ideia tão invulgar contemporaneamente, como se pode imaginar. O que Platão defende é que uma pessoa de posse de todas as informações e com conhecimento suficiente para utilizá-las escolherá sempre o que dá mais prazer, seja, por vezes, o prazer momentâneo, ou um prazer a longo prazo, em outros momentos. Esse prazer é o resultado de um cálculo complexo, que aparentemente Platão achava possível realizar, e que envolvia basicamente o tempo, o prazer e a dor. Seria algo como trazer todos os valores futuros para valores presentes, somá-los e a acção a ser realizada seria aquela que envolvesse a maior quantidade de prazer, ou a menor quantidade de dor, para o agente. Fazer isso, por exemplo, em termos financeiros é muito fácil já que os dados são claros, objectivos e partilhados pelo mercado – qualquer calculadora financeira faz isso com apenas uma dúzia de toques. No entanto, a capacidade cognitiva humana parece não nos permitir

realizar tais cálculos tão complexos em tempo útil e, além disso, quase nunca possuímos os dados ou informações requeridos para tal.

Platão parece assim defender que a ignorância é a condição essencial para se agir contrariamente à acção que produz o maior prazer. A ideia de que este tipo de acção é fruto da ignorância é sedutora, mas se a levarmos às últimas consequências teremos que assumir, contemporaneamente, que todas as acções humanas são fruto da ignorância e que seriam, nesse sentido, quase sempre, contrárias à acção que produz o maior prazer, na medida em que não há, em princípio, uma escolha/acção que seja realizada a partir de cálculos perfeitos, com informação completa.

Seguindo Taylor (1980, p. 505), podemos concluir a partir da leitura do Protágoras que Platão (358b-e) diz algo próximo a:

- (i) Se o prazer é o bom, então, se alguém julga que x é melhor do que y, e ele é capaz de fazer ou x ou y, ele faz x;
- (ii) Se A julga que x é bom e que y é mau, ele não está disposto a fazer y preferencialmente a x;
- (iii) Se A julga que (a) x é mau, (b) y é mau e (c) y é pior do que x, e se A tem que escolher entre x ou y, ele escolherá x.

De facto, parece-me correcta essa interpretação. No entanto, uma questão relevante é saber se é possível dizer que os conceitos de *prazer*, *bom* e *melhor*, por um lado, e *dor*, *mau* e *pior*, por outro, são intercambiáveis no contexto teórico actual. Por ora, essa pergunta deve ser apenas registada. Quando passarmos à descrição das ideias de D. Davidson vou retomá-la para, de algum modo, lhe responder negativamente.

3.2.2 ARISTÓTELES

Para Aristóteles, retomando a origem da palavra *akrasia*, o poder em causa era o poder de alguém se controlar a si mesmo diante das tentações. Resumidamente, na versão aristotélica a *akrasia* era tratada como uma deficiência de auto-controlo. Nesse caso, o autocontrolo deve ser entendido como a capacidade de se controlar a si mesmo diante das tentações. Este era um problema moral, isto é, o problema era decidir-se pelo bem e fazer o mal.

A acção acrática caracteriza-se pela presença de dois desejos conflitantes durante e após a realização da acção, levando o agente a realizar uma acção contrária àquela que é fruto da sua própria deliberação. Há a ideia de que o desejo da satisfação, de um prazer corporal excessivo, é mais forte do que aquele de ser comedido. O agente será, então, conduzido ao excesso (Muños, 2002, 222).

Seria algo como o que estava já presente num exemplo anterior: alguém que tinha decidido fazer uma determinada dieta comendo salada e frango grelhado, come, de facto, metade de um leitão da Mealhada e toma duas garrafas de vinho verde e, obviamente, depois, arrepende-se do excesso. O elemento negativo, do ponto de vista moral, não é comer o leitão e sim comer em excesso, ou fazer aquilo que é mau – é algo similar ao pecado da gula.

Aristóteles parece ter-se interessado pela *akrasia* por causa da aparente contradição entre o resultado da deliberação, que produziu esta acção, e a acção. O problema tratado por Aristóteles tinha as seguintes características: a pessoa S sabe, no momento em que age, que a melhor acção a ser realizada é X (fruto da

deliberação), mas faz Y e depois arrepende-se. Ele percebeu que esse comportamento, longe de ser raro, era muito frequente. Aristóteles percebeu também que quando a acrasia acontece parece haver dois níveis relevantes de desejos: um mais forte (dY – seguindo o exemplo acima) que o leva a agir no sentido de Y diferente daquele (dX) que o levou a decidir-se por X.

Aristóteles define ainda duas formas de falta de domínio, a que ocorre por precipitação e a que ocorre por fraqueza:

“Há a considerar ainda duas formas de falta de domínio, uma é por precipitação, a outra por fraqueza. Uns são os que, mesmo tendo chegado a uma decisão através de um processo de deliberação, não permaneceram fiéis ao resultado das suas deliberações, quando sucumbem à paixão; os outros são levados pela onda da paixão, sem sequer terem chegado a iniciar qualquer processo de deliberação.”
(EN 1150b15,20)

Para mim, esta passagem é significativa porque pode trazer algumas dificuldades para os filósofos da acção. Aristóteles admite um tipo de falta de domínio em que há deliberação e outro em que esta não existe. Hoje em dia, como mais à frente será notado, a acrasia define-se como uma acção que contraria o melhor juízo do agente, depois de tudo considerado. Nesse sentido contemporâneo, é necessária uma decisão fruto da deliberação para que ocorra a acrasia – o problema fica circunscrito, então, à irracionalidade prática. Se considerarmos a falta de domínio por precipitação como sendo também acrasia, além da falta de domínio por fraqueza, o que poderemos dizer acerca deste tipo de acrasia no qual a deliberação não ocorre? Essa não tem sido uma possibilidade desenvolvida pelos filósofos contemporâneos da acção.

Nos dois casos (precipitação e fraqueza) o agente arrepende-se do que faz – “quem perdeu o domínio de si tende a arrepender-se sempre do mal que faz” (EN 1150b30). Aristóteles compara, nessa passagem, o incontinente com o devasso, que, segundo ele, “não se arrepende nunca do mal que faz, ou seja, é fiel à decisão que tomou” (EN 1150b29). Em termos morais, o devasso estaria num nível mais abaixo, isto é, o devasso delibera fazer o mal, faz e não se arrepende, enquanto o incontinente delibera fazer o bem, faz o mal, mas se arrepende do que faz – “o devasso é incurável, enquanto quem perde o domínio de si é curável” (EN 1150b30-35).

Quando Aristóteles escreve que o devasso é fiel à decisão que tomou, diz, por contraste, que o incontinente não é fiel à decisão. Portanto, há claramente a ideia de que há uma decisão que será contrariada pelo acrata, e que, então, há dois elementos essenciais que definem uma acção acrática: a decisão (resultado da deliberação) e a acção. O cálculo, isto é, a deliberação, é também um elemento importante, mas está já incluído na decisão. Aristóteles diz que “quem se domina a si permanece fiel às determinações do seu cálculo, enquanto quem não se domina está inclinado a traí-lo” (EN 1145b10).

É relevante perceber e ressaltar que Aristóteles compara moralmente o incontinente e o devasso, pois, para ele, a acrasia e a devassidão são problemas morais e não de raciocínio simplesmente. Aristóteles (EN 1145a35-b1) contrasta ainda os seguintes dois conjuntos de disposições: (i) falta de autodomínio e lassidão e (ii) autodomínio e perseverança. Aristóteles escreve claramente que “o autodomínio e a perseverança são disposições sérias e louváveis, enquanto, da outra parte, a falta de autodomínio e a lassidão são vis e repreensíveis” (EN

1145b5,10). O acrata decide e percebe o que é bom, mas sucumbe. Noutra passagem relevante para essa caracterização, Aristóteles diz:

“Há, assim, um certo tipo humano que fica fora de si sob efeito da paixão e age contra o sentido orientador, mas de tal sorte que se, por um lado, a paixão o domina ao ponto de não o deixar agir em conformidade com o sentido orientado, por outro, a paixão não domina nele ao ponto de o fazer convencer-se de que deve perseguir negligentemente prazeres desse género. Este é o que não domina, e é melhor do que o devasso. Não é, de resto, completamente mau, porque o melhor nele, o princípio da acção, fica salvaguardado. (...) É evidente a partir do que foi visto que uma disposição é séria e outra é má.” (EN 1151a20-29)

Sobre a decisão e a deliberação (EN 1111b25,30 – 1113b5), Aristóteles, no Livro III, escreve que toda decisão conduz a uma acção voluntária, mas nem toda acção voluntária resulta de uma decisão e que só se pode decidir acerca daquilo que se acredita vir a acontecer através do próprio agente. Podemos, por exemplo, fazer uma analogia entre esta ideia e o conceito de conjunto de oportunidades tratado no item 1.3.

Segundo ele, desejamos/ansiamos pelos fins, mas o objecto da nossa decisão são os meios. Aristóteles dá o seguinte exemplo: “nós ansiamos por restabelecer a saúde, mas apenas decidimos aquilo através do qual viremos a obter saúde; também dizemos que ansiamos por ser felizes, contudo, dizer que decidimos ser felizes não é adequado”.

Diferencia também a decisão da opinião. Segundo ele, formamos opinião acerca de tudo, inclusive de coisas impossíveis, de coisas que não dependem de nós e, claro, de coisas que dependem de nós. Acerca de uma opinião podemos dizer se é falsa ou verdadeira, mas acerca de uma decisão podemos dizer se é boa ou má.

Para ele a decisão implica sempre um processo de pensamento e a deliberação é sempre uma investigação. Ele escreve:

O que está na situação de deliberar parece investigar e analisar do modo referido como se estivesse a investigar e a analisar um diagrama (parece que nem toda a investigação é uma deliberação, como no caso das matemáticas, mas toda deliberação é uma investigação), e o passo final na ordem da análise é o primeiro na ordem da execução.

Podemos pensar num processo que tem início num desejo (fim) e que termina numa acção. Podemos, simplificando o processo, dizer que a partir de um fim o agente delibera sobre o melhor meio para se chegar ao fim desejado e o final da deliberação é uma decisão/acção.

O que move alguém à acção é o desejo (*orexis*). O desejo, por exemplo o apetite, pode mover alguém à acção, mas o intelecto não move ninguém sozinho. Isto é, o desejo sempre está presente e, em alguns casos, pode levar sozinho o agente à acção.

Aristóteles escreve:

“Na verdade, mostra-se que o intelecto não faz mover sem o desejo (pois a vontade é desejo, e quando se é movido de acordo com o raciocínio, também se é movido de acordo com a vontade), mas o desejo move deixando de lado o raciocínio, pois o apetite é um tipo de desejo” (DA, 433a17)

A forma mais primitiva de acção para Aristóteles é aquela em que um animal (humano, ou não) age simplesmente porque percebe um objecto como prazeroso, ou doloroso. Dahl (1984, p. 25) dá o exemplo do animal que bebe porque tem sede, e estar sedento é, parcialmente, estar disposto a perceber o “beber” como

prazeroso. Nesse sentido, uma acção primitiva parece ser aquela em que quando perguntado “Porque é que você bebeu isto?” o agente responde citando o desejo (apetite) presente na acção – nesse caso responderia: “Porque estava com sede – eu queria algo para beber”.

Porém, o desejo (*orexis*) em Aristóteles é um conceito que tem três formas: o apetite (*epithumia*), paixão (*thumos*) e o desejo racional (*boulesis*).

Quando ocorre a acrasia, é o desejo sozinho que move o agente, pois “mesmo que o intelecto ordene e o raciocínio diga que se evite ou busque algo, o indivíduo não se move, mas age de acordo com o apetite, como no caso dos incontinentes” (DA, 433a1).

Tenho feito referências à decisão e à deliberação, mas é preciso dizer algo mais acerca deste processo, relativamente ao raciocínio prático / silogismo prático. Aristóteles considerava que a conclusão de um silogismo prático é uma acção. Esse silogismo é constituído por duas premissas de dois tipos diferentes e uma conclusão. Uma premissa é denominada *premissa maior* e prende-se com a definição do que é bom. A outra premissa, denominada *premissa menor* prende-se à definição do que é efectivamente o caso. A premissa maior, muitas vezes, prende-se também, nos exemplos de Aristóteles, ao que é desejado e ao que é necessário (Dahl, 1984, p. 26 e 27). Diz Aristóteles: “(...) quando se forma uma opinião a partir da geminação das duas primeiras, a alma é obrigada no horizonte teórico a afirmar a conclusão derivada. No caso de se tratar do horizonte prático, a alma é obrigada a agir imediatamente” (EN 1147a25). Desse modo, obtemos a conclusão do silogismo. Aristóteles assume evidentemente que a conclusão de um silogismo prático é uma acção e não uma decisão.

Para exemplificar um tal argumento, segue-se imediatamente um dos seus exemplos específicos de um silogismo prático²⁰:

Todas as coisas doces devem ser provadas	(premissa maior)
Isto é doce	(premissa menor)
... [acção]	(conclusão)

A conclusão não é enunciada. Em vez disso, Aristóteles diz:

“Por exemplo, se se diz ‘tem de se provar todas as coisas doces’, e se ‘esta coisa aqui é doce’, enquanto uma das coisas particulares que são doces, aquele que tem essa possibilidade, e nada o impede, tem necessariamente que pôr logo esse conhecimento na prática” (EN 1147a30).²¹

Um raciocínio teórico, por oposição ao prático, termina numa crença, como é o caso do seguinte exemplo:

Todas as coisas doces devem ser provadas
Isto aqui é doce
Logo, isto deve ser provado

Se para Aristóteles o raciocínio prático termina numa acção, como é possível ocorrer a acrasia? Utilizando o exemplo anterior (silogismo prático - acção), podemos exemplificar e explicar como é possível a acrasia – assim como fez Aristóteles (EN 1147a30 – 1147b15).

²⁰ Aqui sigo a formulação de Broome (2007), por estar mais intuitiva. No original, veja em EN 1147a30,35.

²¹ Broome (2007) utiliza uma tradução diferente: “Uma vez relacionadas estas premissas, estamos obrigados, se nada nos impedir e tivermos capacidade para tal, a provar a coisa em questão”.

Acção não acrática

Todas as coisas doces devem ser evitadas
 Isto aqui é doce
 ... [Não como o doce]

(*mas*)
 (e)

Acção acrática

[Tudo o que é doce é agradável]
 [Tenho apetite de comer doce]
 ... [Como o doce]

No caso acima, formou-se uma opinião universal na mente do agente (premissa maior) que indicava o impedimento de provar coisas doces e, simultaneamente, forma-se uma segunda opinião de que se sente que “tudo o que é doce é agradável”. Além disso, formou-se uma premissa menor que indicava que “isto aqui é doce”, e, ao mesmo tempo, ocorre o apetite para comer doce. Com isso, forma-se um conflito entre a premissa maior “todas as coisas doces devem ser evitadas” e o meu desejo (apetite) de comer doce. Esse desejo é capaz, como vimos anteriormente, de levar o agente a agir contrariamente ao que seria a acção resultante do silogismo prático. Gera-se assim a perda do autodomínio e, conseqüentemente, a acção incontinente tem lugar.

3.2.3 DAVIDSON

É já lugar-comum afirmar que o ponto inicial do debate contemporâneo acerca da acrasia é o artigo “How is Weakness of Will Possible?” de Donald Davidson. No entanto, pode ser útil, como indicam Stroud e Tappolet (2007: 2, 3), olhar para a discussão sobre a acrasia promovida por R. M. Hare no livro *The*

Language of Morals (1952), pois o artigo de Davidson é em parte uma reacção à negação da possibilidade da acrasia estrita²² defendida por Hare.

De forma simples, pode dizer-se que Hare nega a possibilidade da acrasia estrita porque defende o prescritivismo. Hare, então, defende que juízos morais como “Eu devo fazer X” são imperativos e isso significa que se eu for livre para fazer X, farei X. Essa abordagem é uma versão forte da ideia internalista de que há uma relação entre os juízos práticos, a motivação e a acção.

Hare afirma que o que realmente ocorre em casos de suposta acrasia é o seguinte: (a) *o agente realmente não julga que deve fazer X*, e/ou (b) *o agente não está livre física, ou psicologicamente, para agir*.

A citação abaixo é bastante reveladora do pensamento de Hare. O que está abaixo citado é parcialmente o que defendo nesta tese. Hare pergunta acerca dos princípios morais, já a minha pergunta refere-se às crenças e desejos (intenções)²³ dos agentes. Caso eu aceitasse o prescritivismo de Hare, parece-me que seriam perguntas idênticas. Hare diz:

“Se perguntássemos a uma pessoa ‘Quais são os seus princípios morais’ a maneira pela qual poderíamos ter mais certeza de uma resposta verdadeira seria estudando o que ela faz. Ela pode, logicamente, professar em seu discurso toda sorte de princípios que desconsidera completamente em suas acções; mas, quando estivesse frente a escolhas ou decisões entre cursos de acção alternativos, entre respostas alternativas à questão ‘Que devo fazer’, conhecendo todos os factos relevantes de uma situação, ela revelaria em quais princípios de conduta realmente acredita. A razão pela qual as acções, de uma maneira peculiar, são reveladoras

²² Segundo Alfred Mele a acrasia estrita é uma acção intencional e livre que é contrária a uma crença consciente que o agente tem, no instante em que age, de que é melhor fazer X e não Y, mas faz Y. Aqui “melhor” define-se “a partir da perspectiva dos seus próprios valores, desejos, crenças, entre outros, em oposição, por exemplo, a uma perspectiva avaliativa comum que ele não endossa” (Mele, 2006: 729).

²³ A pergunta seria: Quais são/foram as suas crenças e desejos na acção?

de princípios morais é que a função dos princípios morais é orientar a conduta. (Hare, 1996: 1)”

Se quisermos saber quais foram as crenças e os desejos que explicam a acção X de um agente temos que primeiro olhar para X e depois procurar o desejo e a crença que explicam X. No caso da acrasia, quando é considerado um fenómeno ligado exclusivamente à irracionalidade prática, temos que aceitar que existe um princípio qualquer que deveria levar-nos a agir de acordo com aquilo que chamamos o “melhor juízo”, seja a acção moral ou imoral. A esse princípio Davidson dá o nome de princípio de continência. Sendo assim, se agimos através das nossas próprias crenças e desejos, mas contrariamos o princípio de continência, fomos acráticos. Isto é, se tínhamos consciência de que havia uma outra acção a qual julgávamos melhor do que aquela que realizámos e estávamos em posição de realizá-la e não a levamos a cabo, houve uma falha da racionalidade. Aristóteles, nesse caso, certamente, diria que foi um desejo (apetite) que desviou o caminho do agente. Porém, no caso de Aristóteles e mesmo de Platão, o que estaria em jogo era o bom e o mau, o maior, ou o menor prazer/dor. No caso do princípio de continência, temos apenas que seguir aquilo que decidimos e que consideramos o melhor juízo, seja ele o mais, ou o menos prazeroso (bom), ou o mais ou menos doloroso (mau).

Assim, contemporaneamente, a acrasia tem sido tratada de maneira muito diferente, relativamente a Platão e a Aristóteles, como espero que esteja a ficar um pouco mais claro. Basicamente, a acrasia deixou de ser um problema da filosofia moral e passou a ser um problema da filosofia da acção. O filósofo Ricardo Santos, que é um dos poucos autores de língua portuguesa que escreveu directamente

sobre o tema, afirma que: “(...) a possibilidade da acrasia é um problema de filosofia da acção, e não de filosofia moral ou de ética normativa” (no prelo: 1). Essa afirmação não só está adequada sociologicamente, em termos científicos, como simboliza um dos problemas acerca da acrasia que pretendo aqui enfrentar: Será que de facto a acrasia existe como uma falha no processo decisão/acção, ou será que é uma acção caracterizada pelas pessoas comuns através de elementos morais? Qual das duas definições é mais relevante em termos filosóficos e em termos práticos psicológicos?²⁴

A partir de Davidson, o problema da possibilidade da acrasia transfere-se para a filosofia da acção – com novos pressupostos e sem o comprometimento original com a filosofia moral.

O artigo “How is Weakness of the Will Possible?”, de Donald Davidson (1970), é considerado o ponto de partida do debate contemporâneo acerca da acrasia, como já foi dito. Davidson tenta desmontar a ideia de Hare acerca da impossibilidade da acrasia. Não faz isso somente e especificamente acerca dos juízos morais, mas acerca dos juízos práticos, de maneira mais geral. Tenta mostrar que a acrasia não é uma questão apenas de agir contra um juízo moral, mas sim uma questão de agir livre e intencionalmente contra um juízo prático.

Davidson, então, propõe que as acções intencionais seguem os seguintes princípios, como vimos no item 1.3.2:

(P1) Se um agente quer fazer x mais do que quer fazer y e acredita que pode fazer ou x ou y, então fará intencionalmente x, se fizer ou x ou y intencionalmente.

²⁴ Ver item 3.4 para uma tentativa de resposta.

(P2) Se um agente julga que seria melhor fazer x do que fazer y, então quer fazer x mais do que quer fazer y.

Segundo Davidson, esses princípios são intuitivos e ele os considera auto-evidentes. No entanto, P1 e P2 parecem levar à falsidade de um outro princípio, segundo ele também auto-evidente:

(P3) Existem acções incontinentes

Para defender a possibilidade da acrasia, Davidson distingue, então, dois tipos de juízos avaliativos: (a) “juízos avaliativos incondicionais” (ou juízos avaliativos *sans phrase*), do tipo “o que é melhor”; (b) juízos avaliativos *prima facie*, do tipo “o que é melhor, todas as coisas consideradas (*tcc*)”

A diferença entre eles é que os juízos avaliativos *prima facie* são relacionais e os incondicionais não. Um juízo do tipo *tcc* não envolve um comprometimento adicional do agente com a opinião expressa pelo próprio juízo. Ele não diz nada acerca do que é melhor de maneira incondicional, mas o que é melhor a partir de uma razão (*r*). Davidson propôs que a forma lógica desse juízo avaliativo relacional envolve um operador *prima facie* que comanda todo o juízo (Stroud & Tappolet, 2007, p. 4).

A forma geral desses juízos é a seguinte: *pf* (X é melhor do que Y, *r*) (Stroud & Tappolet, 2007, p. 4; Santos, 1997, p.96). Aqui, *r* refere-se às razões do porquê “X é melhor do que Y” (juízo). Isto é, X é melhor do que Y, *tcc* – onde *tcc* significa que o agente julga X melhor que Y de maneira relacional/condicional às suas próprias razões para isso, por exemplo, às suas próprias crenças e desejos.

Davidson define, a partir disso, a acção acrática da seguinte forma:

Ao fazer X, uma pessoa age de modo incontinente se e somente se: (a) faz X intencionalmente; (b) acredita que há uma acção alternativa Y que pode fazer; e (c) julga que, consideradas todas as coisas, seria melhor fazer Y do que fazer X.

Davidson também define um requisito racional para se fazer cumprir um juízo condicional que indica uma acção como sendo a melhor tendo em vista todas as razões relevantes disponíveis. Dá-lhe o nome de “princípio de continência”.

Princípio de Continência: Realiza a acção que julgas, com base em todas as razões relevantes disponíveis, ser a melhor.

Esta proposta de Davidson torna possível a acção acrática. Esta tem como característica básica a falha do princípio de continência quando o agente realiza uma acção que contraria o juízo avaliativo condicional, ou *prima facie*. Isto é, a acrasia existirá quando a acção X for considerada, pelo agente, melhor do que Y com base em todas as razões relevantes disponíveis (*r*) (crenças e desejos do próprio agente) e esse agente levar a cabo Y.

Assim, de acordo com a definição, uma acção acrática é aquela que manifesta inconsistência relativamente a certas crenças ou desejos do agente (uma inconsistência interna relativamente às crenças/desejos do agente e não uma inconsistência que se estabelece a partir de fora, relativamente a padrões externos).

Mais concretamente, seria um caso em que o agente age contrariamente às razões que ele próprio julga serem as mais relevantes no que toca à acção em

questão. Ele opta por um determinado curso de acção, por lhe parecer o melhor, mas acaba por agir num sentido diferente, que até pode ser o oposto.

Nesse caso, como foi dito, estaria a ser violado o chamado princípio de continência (Davidson, 2004, “Paradoxes of Irrationality”). Há irracionalidade porque a razão que leva o agente a agir em sentido oposto suplanta o próprio princípio de continência – essa razão não é uma razão contra o princípio em si (não o põe em causa), mas é usada como tal; há irracionalidade porque o agente não atribui às considerações relevantes que estão em jogo o peso que elas de facto teriam (ou esse peso não é suficiente para o incitar a agir).

É importante destacar que a concepção davidsoniana da acrasia parece levar à uma concepção pouco realista da racionalidade. Isto acontece porque para agir de modo não acrático, portanto agir de modo racional, o agente tem que levar em conta e pesar todos os elementos importantes. Isto leva-nos a concluir que, em grande parte dos casos, o esforço para conhecer estes elementos e reflectir acerca deles é quase infundável. Sobre isto, Zilhão (2009, p. 114 e 115), de maneira categórica, escreve:

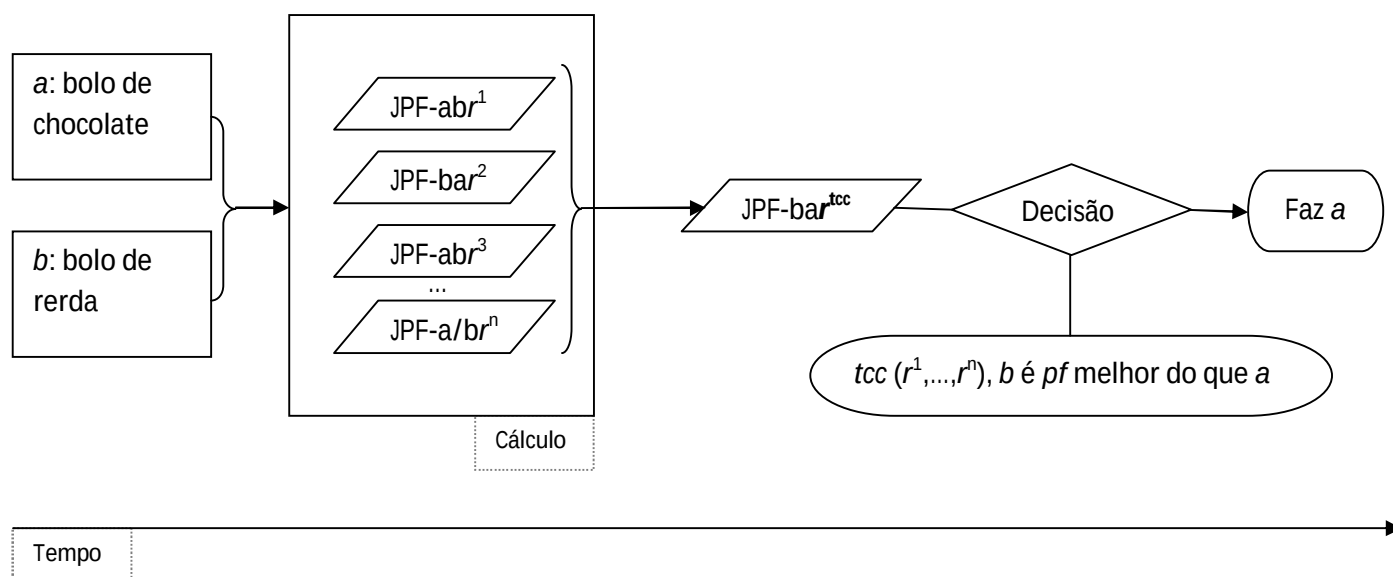
Davidson diz-nos que um agente é continente se e somente se ele faz uma busca exaustiva no seu stock de crenças com vista a assegurar-se que *todas* as razões relevantes foram tomadas em consideração e apropriadamente sopesadas na formação do seu próprio juízo acerca da melhor forma de agir e se ele age de acordo com esse juízo.²⁵ Enquanto forma de acção racional, a acção continente é assim concebida, simultaneamente, como uma instanciação de um modelo de pura racionalidade ilimitada e como tendo um sentido psicologicamente relevante. Por outro lado, um agente é considerado por Davidson como tendo sido incontinente se e somente se, na sua acção, ele *negligenciou* ou *não levou em conta* uma parte relevante das suas próprias crenças. E é isto que faz da sua acção incontinente uma acção irracional. Todavia, ele levou em conta algumas dessas crenças. E é isso que faz da sua acção uma acção intencional. Os argumentos cognitivos contra a

²⁵ Nota da citação [Cf. Davidson, 1980, “How is Weakness of Will Possible?”, pp. 40-41.]

validade descritiva de um modelo de pura racionalidade ilimitada são bem conhecidos. Não há assim necessidade de os passar em revista aqui. Pressupondo que eles são basicamente correctos, a análise de Davidson dos conceitos de continência e incontinência implica a consequência de que toda a acção humana real é incontinente. Se este é o caso, o par de conceitos continente/incontinente deixa de ter qualquer relevância explicativa para quaisquer fins psicológicos.

Abaixo encontra-se um fluxograma (Figura 8) produzido a partir do que foi acima exposto e que servirá de ilustração para alguns exemplos que vêm a seguir.

FIGURA 8 – ACRASIA



Voltando a um exemplo já introduzido, temos o sujeito S, que está num outro país, que tem o desejo de comer algo que acompanhe uma chávena de café e que tem que decidir entre dois bolos possíveis: (a) bolo de chocolate (parecido com o da sua avó) e o (b) bolo de rerda, fruta típica da vostalândia.

O processo de deliberação deverá terminar numa decisão que forma uma intenção e que, em princípio, o levará à acção. S começa a deliberação (cálculo) levando em consideração algumas razões para fazer *a* ou para fazer *b* [JPF-a,brⁿ] (veja-se Figura 8). Abaixo, estão alguns exemplos de possíveis *juízos prima facie* (JPF) que pendem para *a* (JPF-abr), ou para *b* (JPF-bar) dadas determinadas razões (*r*).

JPF-abr¹: *a* é *prima facie* melhor do que *b* porque S gosta de chocolate e não conhece rerda (razões *r*¹)

JPF-bar²: *a* é *prima facie* melhor do que *b* porque S gosta de experimentar sabores novos – o sabor do chocolate S já conhece e o da rerda não (razões *r*²)

JPF-a/b, b/arⁿ: *a/b* é *prima facie* melhor do que *b/a* porque ... (razões *r*ⁿ)

Depois do processo de deliberação ter chegado ao fim, S deverá ter consciência de um juízo condicional do tipo “todas as coisas consideradas” (JPF-bar^{tcc}, no exemplo acima). Nesse caso, o juízo *prima facie* condicional (decisão) produzido por S foi “o bolo de rerda é melhor do que o bolo de chocolate, todas as coisas consideradas”. Desse modo, S decidiu-se por comer bolo de rerda.

JPF-bar^{tcc}: *b* é *prima facie* melhor do que *a* porque ... (razões *r*^{tcc})

[JPF-bar^{tcc} = melhor juízo de S, todas as coisas consideradas]

Se S comer (*b*) bolo de rerda estará a agir de maneira continente, no entanto, se comer (*a*) bolo de chocolate, estará a ser incontinente (acrático). Porque não

seguir o tal requisito racional, o princípio de continência (realiza a acção que julgas, com base em todas as razões relevantes disponíveis, ser a melhor).

Do ponto de vista lógico informal, a explicação davidsoniana da possibilidade da acrasia parece adequada tendo em vista os longos anos de debate acerca desse fenómeno sem que alguém tenha (que se saiba) proposto uma solução desse género. No entanto, o ponto que pretendo retomar para negar a possibilidade da acrasia não se prende com a “viabilidade lógica” da proposta de Davidson. Prende-se sim com a presunção de viabilidade cognitiva ou psicológica da expressão “o melhor juízo, todas as coisas consideradas” (JPF- r^{icc}).

A ideia de que a acção acrática se define por um desejo (apetite, ou mesmo paixão) que aparece de maneira mais ou menos fantasmagórica e corrompe a decisão, sem estar envolvido na deliberação, parece-me muito pouco plausível dada a psicologia humana. Acho que não é possível dar conta da possibilidade da acrasia apenas nos termos de Davidson. Se a acrasia continuasse a ser tratada como um problema moral, não se colocava essa questão, pois, como problema moral seria discutido como um fenómeno dependente dos juízos morais, da capacidade motivadora e da formação desses juízos, etc. Não estaria em jogo a racionalidade do sujeito na acção, como está em Davidson.

Na próxima secção tentarei tratar dessas questões a partir de alguns exemplos.

3.3 OBSERVAÇÕES ACERCA DA FRAGILIDADE PSICOLÓGICA E CONCEPTUAL DA ACRASIA COMO UM FENÓMENO DA RACIONALIDADE NA ACÇÃO

Nesta parte do capítulo, tentarei argumentar em favor de duas estratégias de negação da possibilidade da acrasia como Davidson a define. Com algumas variações, algumas mais importantes do que outras, mas nenhuma “definitiva”, o conceito contemporâneo de acrasia gira em torno da definição proposta por Davidson.

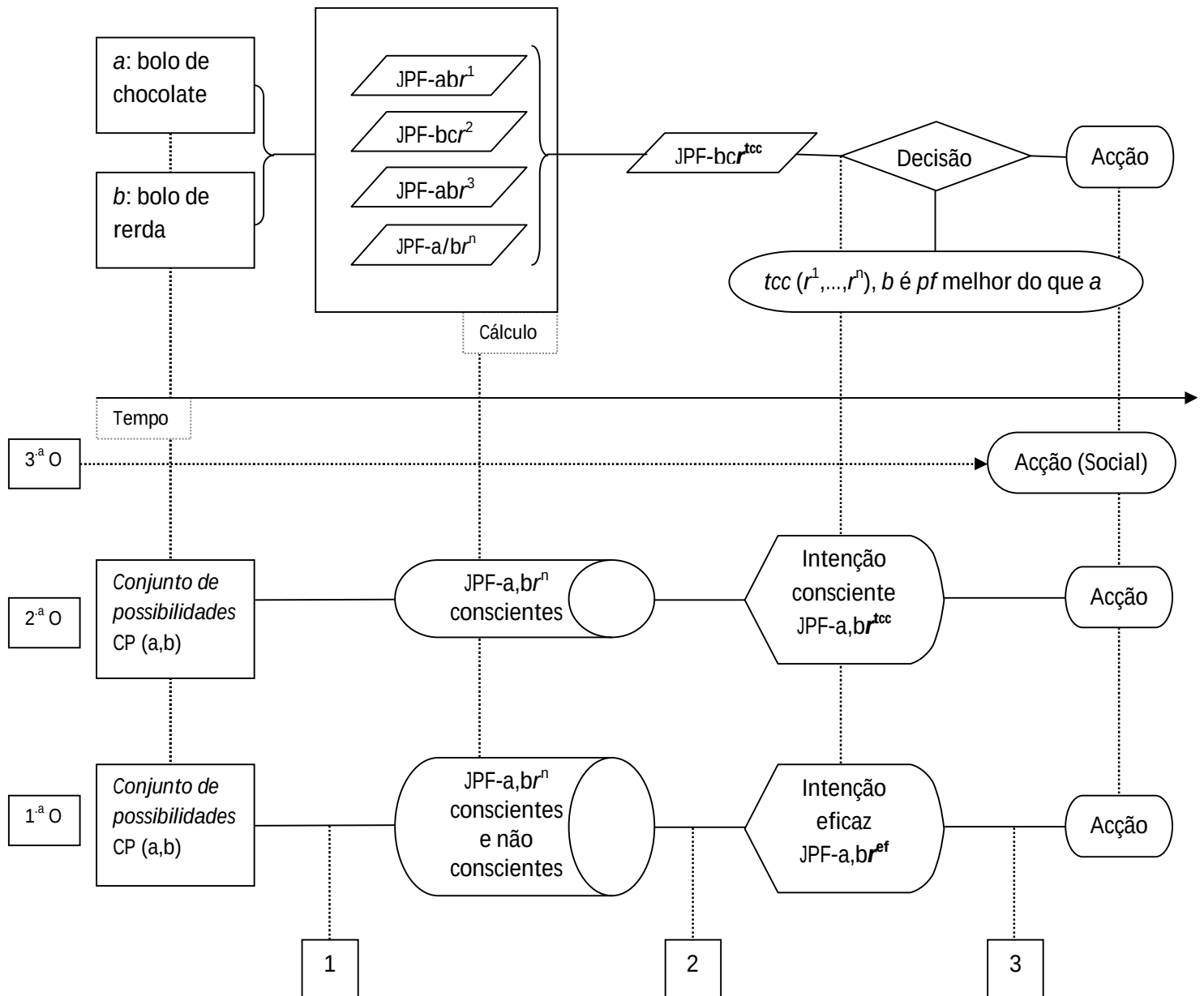
Na primeira parte, tento defender a ideia de que o processo definido por Davidson (Figura 9) é um processo de segunda ordem e que há um processo de primeira ordem que é aquele que, de facto, leva o agente à acção, onde a acrasia não é possível.

Na segunda parte, tento defender uma outra ideia, a de que a acrasia depende do conceito de decisão e que se compreendermos o processo de deliberação como um processo que tem como conclusão uma acção e não uma intenção, que num tempo diferente levará o agente à acção, poderemos defender que a acrasia não é possível. Não é possível porque o desejo provocador da acrasia, como no caso de Aristóteles, já está incluído no processo de deliberação.

3.3.1 JUÍZOS DE 1^A. E DE 2^A. ORDENS

A probabilidade de ser viável cognitiva/psicologicamente “o melhor juízo, todas as coisas consideradas” é, para mim, menor do que a probabilidade de ser viável “o melhor juízo consciente, consideradas as razões relevantes e disponíveis à consciência – dado um determinado contexto” (*crdc*). Abaixo, segue um fluxograma (Figura 9) que servirá de base para a descrição dessa alternativa.

Figura 9



(2ª O) CP (a,b) provoca o cálculo JPF-a,brⁿC, produz IC (JPF-a,br^{tcc}) que motiva a acção [cálculo parcial que racionaliza a acção] = JPF tcc

(1ª O) CP (a,b) provoca o cálculo JPF-a,brⁿC,ñC produz I.Ef (JPF-a,br^{ef}) motiva a acção [cálculo completo] = JPF ef

----- 1ª O, 2ª O, 3ª = primeira, segunda e terceira ordens -----

[1] CP (a,b) provoca o cálculo JPF-a,brⁿC // CP (a,b) provoca o cálculo JPF-a,brⁿC,ñC

[2] O cálculo JPF-a,brⁿC, produz I.Cons (JPF-a,br^{tcc}) // O cálculo JPF-a,brⁿC,ñC produz I.Ef (JPF-a,br^{ef})

[3] I.Cons (JPF-a,br^{tcc}) motiva a acção // I.Ef (JPF-a,br^{ef}) motiva a acção

De maneira directa e simples, quero dizer que quando S produz um JPF- $\bar{b}ar^{tcc}$ está, na verdade, a “dizer” que dadas as razões conscientes, relevantes e disponíveis, num determinado contexto, b é melhor do que a . No entanto, se dissermos que esse juízo é já um juízo de segunda ordem no qual intervêm basicamente razões conscientes, que passam ao mesmo tempo por um processo de racionalização anterior à acção, podemos dizer que esse “melhor juízo” é, na verdade, um juízo parcial e que dá ao agente um curso de acção recomendável, possível, racionalizadamente aceitável, mas, não necessariamente motivador da acção. Esse juízo é de segunda ordem porque o juízo de primeira ordem é, de facto, o juízo que sempre motiva as acções do agente. Muitas vezes os dois juízos coincidem (1ª e 2ª ordens), mas muitas vezes não. Nas vezes em que coincidem, dizemos que fomos racionais e nas vezes em que não coincidem dizemos que fomos irracionais, por exemplo acráticos. A partir dessa submissão avaliativa das acções passadas aos juízos de segunda ordem, muitas vezes, procuramos adequar as nossas futuras acções a esses juízos de segunda ordem que foram descartados, por exemplo, numa suposta acção acrática. Isso acontece porque a necessidade que tem o agente de corresponder aos juízos de segunda ordem é intensa, dada a natureza social do agente humano. Quanto mais o agente tem consciência dos seus juízos de primeira ordem, menor será o conjunto de acções que ele considerará acráticas, dado o maior equilíbrio entre as suas expectativas de primeira e de segunda ordem.

A entrada dos juízos de segunda ordem no cálculo (sistema) de primeira ordem acontece por via dos juízos intersubjectivos (aqueles cujas razões r dependem fortemente da intersubjectividade) que são levados em consideração

pelo agente, mas nada garante que a ponderação desses juízos na equação (cálculo) tenha valores altos ou baixos. Essa ponderação (multiplicador) dependerá de factores que ultrapassam o escopo desse trabalho, mas parece razoável pensar que quando falamos de liberdade (social) podemos estar a falar disso, ou seja, uma pessoa mais livre (socialmente) é aquela que tem no seu sistema de cálculo uma “equação” cujos multiplicadores dos juízos intersubjectivos são baixos. Enfim, aqui há nuances que podem ser exploradas posteriormente. Porém, para uma breve clarificação, podemos pensar, por exemplo, que existem vários tipos de juízos intersubjectivos, que possuem multiplicadores diferentes entre si. Talvez possamos dizer que existem juízos intersubjectivos que se prendem primordialmente com as relações sociais de carácter colectivo, como a religião, cultura nacional, profissão, etc., e que, ao mesmo tempo, existem juízos intersubjectivos que se prendem primordialmente com as relações sociais de carácter privado, como a família do agente, amigos relevantes, etc. Podemos imaginar, por exemplo, que existem pessoas genuinamente afastadas de parte dos valores sociais, sem que isso caracterize algum tipo de patologia que retire a liberdade dessas pessoas. “Genuinamente” porque as suas acções parecem não ter uma relação importante com o contexto social (e só as acções podem dizer algo acerca da referida equação e ponderação). Podemos dizer, em princípio, que essa pessoa tem um sistema de cálculo que pondera com valores baixos os juízos intersubjectivos de carácter colectivo. É possível dizer que esse tipo de pessoa dificilmente se sentirá incontinente por causa de acções que ferem normas sociais, desde que essa ponderação seja consciente o suficiente para que os juízos tcc sejam uma boa expressão dos seus juízos *crdc*.

Por outro lado, no movimento contínuo que conduz à acção, sustento ser fundamental a distinção entre o processo de deliberação (seja a decisão declarada ou não – isto é, quer o agente declare verbal e explicitamente qual será a sua acção, quer seja uma tomada de decisão “interior”, “silenciosa”, não formulada) e o momento em que a acção é realizada – ou seja, a acção propriamente dita. Se no *processo de deliberação* que conduz à decisão estão em jogo muitas crenças e desejos, até mesmo desejos e crenças contraditórias, no momento em que a acção é levada a cabo apenas um juízo particular composto por, pelo menos, uma determinada crença e um determinado desejo é eficaz.

Concretizando, por meio de um exemplo banal – um atleta julga que, dadas as suas circunstâncias, o melhor que tem a fazer no feriado que se avizinha é treinar para a próxima prova. Contudo, acaba por passar o dia com os amigos a passear. Agiu, portanto, contrariamente ao seu melhor juízo tcc, e a sua acção, por definição, é tida como irracional, ou acrática. Isto porque parece existir uma incoerência entre a acção realizada e as crenças ou desejos do agente (nomeadamente a crença de que tem melhores razões para treinar do que para sair com os amigos, e a própria crença de que deve seguir aquela que lhe parece a melhor alternativa de acção possível, conjuntamente com o desejo de ganhar a prova). O que sustento é que não é possível dizer que existe tal incoerência, porque no exacto momento em que ele age, só um juízo *ef* é determinante – possivelmente produzido pelo desejo de se divertir²⁶ e a crença de que tal será possível escolhendo esse curso de acção. Não há qualquer contradição no próprio momento da acção, senão ele não agiria

²⁶ Desejo que poderia ser realizado por meio de outras acções como, possivelmente: beber, ler, contemplar, conversar, beijar, ter relações sexuais. É importante entender que são outras acções e que cada uma terá o seu par crença-desejo.

de todo. Nesse preciso momento, foram esse desejo e essa crença específicos que se sobrepuseram aos outros e provocaram a acção – o agente não pode ser acusado de irracionalidade, nem de “fraqueza de vontade”, só porque esse par crença-desejo não coincide com o que anteriormente havia julgado ser o mais razoável. Como vimos, esse é um exemplo da relação entre os juízos *tcc* e *ef*.

O que pretendi demonstrar com esse exemplo é que não é sustentável acusar um agente real de irracionalidade, apelando para uma possível inconsistência/contradição entre as suas crenças e desejos e a sua acção. O que se passa é que, a um dado ponto, outras crenças (mesmo que mal formadas) e outros desejos (mesmo que imorais ou inadequados) adquiriram maior peso e tornaram-se os propulsores da acção (juízo *ef*). Quando um agente real (por oposição a uma entidade abstracta, que se regesse por princípios formais) vai contra o seu melhor juízo *tcc* não se pode falar de contradição – o que pode haver é um revelar daquelas que são as verdadeiras intenções do agente, ou uma mera mudança de planos. No exacto momento da acção, foi um conjunto particular de desejos e de crenças que se revelou decisivo para o agente (intenção *ef*). Com efeito, se dois desejos contraditórios estivessem presentes no momento da acção, sem que o agente pendesse para um dos lados, então este não agiria²⁷.

²⁷ Esse seria um caso como o do “Burro de Buridan”. História criada por Jean Buridan (1298) que foi reitor da Universidade de Paris e que ilustra uma espécie de impossibilidade de decisão, referindo-se à paralisia de um ser que estivesse dividido entre motivações estritamente equivalentes, que se equilibrariam, como duas cordas que o puxassem de dois lados opostos, com duas forças idênticas, e que fariam com que ele ficasse absolutamente imóvel no ponto onde se encontrasse. O burro tinha tanta sede como fome, e colocaram o animal a igual distância de um recipiente de aveia e de um balde de água. Não sabendo determinar se tinha mais fome do que sede, o animal não foi capaz de decidir se devia comer ou beber em primeiro lugar. Ficou indiferente ao aprovisionamento proposto e morreu – de fome e de sede. Parece improvável que um burro não agisse numa situação dessas, assim como é muito improvável que isso seja possível num agente humano, a não ser em casos patológicos (neurológicos) onde há, por falha do sistema, uma incapacidade de decidir relevante e constante (Pena-Ruiz, 2009)

No instante em que eu opto por um curso de acção que não coincide com aquele que foi determinado pelo meu melhor juízo *tcc*, eu posso estar a agir em sentido contrário ao que é socialmente aceite, posso estar a ser irresponsável ou a revelar um comportamento imoral, mas não estou a ser irracional. O facto é que, para que a nossa acção seja considerada racional, não temos que nos pautar por aqueles desejos que, avaliados todos os dados em jogo numa determinada situação, concluimos serem os mais acertados mediante qualquer referência externa, seja ela moral ou não. Até porque podemos estar simplesmente a seguir aquilo que social e contextualmente é tido como o melhor, o que não tem nada a ver com racionalidade. Ser racional não é seguir normas previamente instituídas – apelando para uma definição mais lata, tudo pode ser racional se tivermos razões para tal, e se agimos intencionalmente é porque tivemos essas razões (*intenção ef*).

Ser racional no sentido em que tenho vindo a referir é inerente à própria acção (ver capítulo 2). Dessa forma, o indivíduo não age sempre nos sentidos considerados mais acertados em termos instrumentais, ou nos sentidos que representam o melhor resultado para ele próprio ou para outras pessoas. Uma acção racional não pode ser definida à partida, com base em princípios ou normas abstractas. A racionalidade deve ser considerada a um nível distinto, não como uma característica adjectiva de certas acções, mas como uma propriedade/qualidade substantiva da realidade humana. Por analogia, podemos dizer que a racionalidade é o “sangue” da acção intencional; isto é, da mesma maneira que se não houver sangue não haverá vida, se não houver racionalidade não haverá sequer acção intencional (ver capítulo 2).

3.3.2 DELIBERAÇÃO, DECISÃO E ACRASIA

Pretende-se em seguida explorar a ideia de que o “agir segundo o seu melhor juízo, *tcc*” está associado intrinsecamente ao conceito de decisão. Se isso se mostrar defensável, será também plausível dizer que se não houver, de facto, “a decisão” – que compromete naturalmente o agente a levar a cabo uma determinada acção – não haverá também algo como “o agir segundo o seu melhor juízo *tcc*”. Isso levar-nos-á imediatamente à conclusão de que se não há o tal melhor juízo, não haverá também acrasia. Isto acontecerá, pois a acção acrática é definida como sendo aquela que é levada a cabo em detrimento do melhor juízo *tcc* do próprio agente, como vimos anteriormente.

Pensemos num exemplo em que Marta recebe uma carta na qual Rita declara o seu amor e pede que, caso seja um amor correspondido, ela (Marta) a encontre no bar da faculdade, na manhã seguinte, às 11h00 da manhã. Marta pode agir de muitas maneiras diferentes, isto é, tem um conjunto de oportunidades razoavelmente vasto. Marta pode, por exemplo, (a) ir ao encontro e declarar o seu amor por Rita, já que a ama, como (b) pode não ir e reprimir o seu amor por ela, (c) pode aparecer e dizer que tem dificuldade em assumir o romance e pedir um tempo para conseguir fazer isso, pode (d) ir ao bar e fingir que não a viu, enfim, há uma série de acções possíveis, desde as mais óbvias como, também, possibilidades invulgares.

Neste caso específico, Marta teve 12 horas para deliberar acerca do que fazer. Recebeu a carta às 23h00 do dia anterior ao encontro e só conseguiu dormir às

3h00 da manhã do dia seguinte após ter “decidido” não ir ao encontro e negar a atracção que sentia por Rita (JPF *tcc*). No entanto, acordou às 9h00, mas passados 15 minutos estava novamente em dúvida e recomeçou o processo de deliberação. A tortura parecia não ter fim quando às 10h00 algo aconteceu...sentiu que de facto havia decidido e que aquela decisão representava o seu melhor juízo (JPF *tcc* confirmado). Marta havia decidido ir ao encontro, mas dizer à Rita que não queria nada com ela ($[D_c+C_c]$ - veja diagrama 3ª à frente). Marta tinha chegado a essa conclusão porque pensou: “os meus pais são muito conservadores, os meus amigos também, corro o risco de ser marginalizada por todas essas pessoas que amo...e mais, o meu pai é cardíaco e já tem 85 anos e se algo lhe acontecer por conta disso, não vou me perdoar... será mesmo que consigo passar os últimos anos de vida do meu pai zangada com ele?” Enfim, Marta havia invocado e reunido um conjunto de crenças que justificava a sua “decisão” (JPF *tcc*). No entanto, Marta chega ao encontro às 11h00 em ponto e vê o lindo corpo de Rita contra a luz. Rita estava especialmente sensual naquele dia. Aquela imagem deixou Marta pouco confortável, pois percebeu mais uma vez a atracção que sentia por Rita. Quando se aproximou, Marta sentiu o delicioso perfume de Rita que tinha acabado de sair do banho e ainda tinha os cabelos molhados. Marta sucumbiu... beijou Rita com a força mais visceral possível (D_k+C_k). Como diria o mais ferrenho defensor do compromisso humano com a decisão, Marta foi realmente fraca, não teve força de vontade para seguir a sua própria decisão. Será que isso é verdade?

A nossa tendência é achar que Marta primeiro deliberou, isto é, realizou algum tipo de cálculo, tendo como recursos as suas próprias crenças e desejos, decidiu e agiu. Parece razoável dizer que Marta tem crenças e desejos e que realizou algum

tipo de raciocínio prático. No entanto, não parece óbvio que houve um momento, um instante específico em que “decidiu” o que fazer. Algo como, “calculei, decidi e agora vou agir” (tenho-que-agir-no-sentido-da-decisão). O que provavelmente aconteceu foi que, até agir, Marta formou algumas intenções, algumas mais fortes e outras mais fracas, e pode até ter pensado que havia decidido negar o amor que sentia por Rita, segundo o seu melhor juízo. Nesse sentido, a intenção que explica a acção de Marta é o único elemento que pode, de facto, ser levado em consideração para definir se Marta agiu ou não de acordo com o seu melhor juízo, pois a única intenção que foi eficaz e que, portanto, deve representar o seu melhor juízo é aquela formada pelo desejo DK e pela crença CK (Figura 10). Só poderíamos aceitar a ideia de que o melhor juízo era aquele formado pela crença C e pelo desejo C se aceitássemos a intuição que de facto a acrasia não passa de um juízo moral e nada tem a ver com a explicação da acção a partir da deliberação do próprio agente.

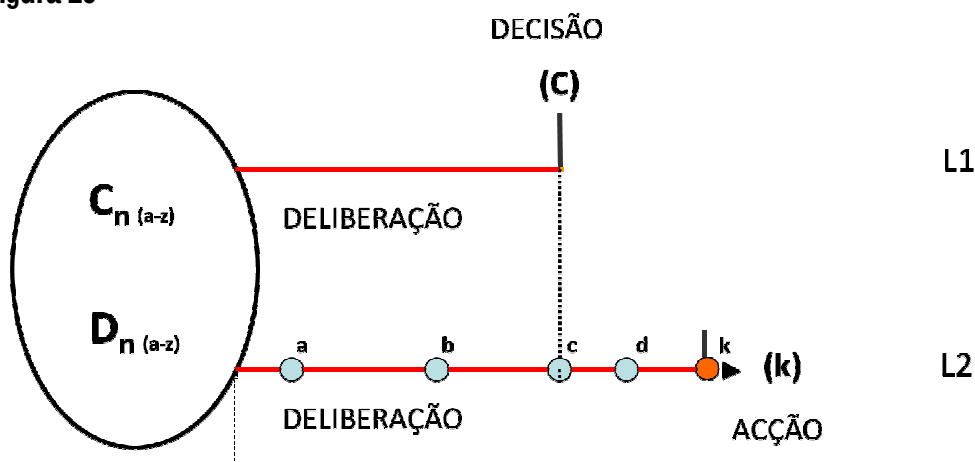
A minha proposta é que não há esse momento de decisão e que, portanto, não podemos identificar uma acção como acrática. A proposta é muito simples: posso dizer que quando Marta se decidiu por C, a intenção em questão era de tal maneira forte que para um observador privilegiado, ou para ela própria, aquele seria, inevitavelmente, o sentido da sua acção segundos depois da formação dessa intenção. No entanto, como tudo mais não se manteve constante, como é plausível defender, Marta formou uma intenção (I_k) e só esta se mostrou, de facto, eficaz no instante em que Marta começou a agir e manteve-se eficaz durante todo o tempo em que a acção ocorreu. Às diversas intenções formadas na deliberação, mas que se revelaram ociosas, podemos chamar “intenções na deliberação” (I_β). A intenção que corresponde à acção pode denominar-se “intenção na acção” (I_α).

Se aceitarmos a ideia de que para explicar uma acção podemos prescindir da decisão, pode também dizer-se que o raciocínio prático termina não numa decisão que reflecte o melhor juízo, mas sim na acção e que esta sim reflecte o melhor juízo. Assim sendo, não haveria condições para identificar uma acção como acrática.

Sei que isto parece muito simples e que parece qualquer coisa como apenas um jogo em que eu faço desaparecer a decisão e conseqüentemente a acrasia também desaparece. Mas, de facto, não é nada disso: é minha convicção que a acção é a melhor expressão do melhor juízo, mesmo que o agente não consiga explicar a sua acção à luz das suas crenças e desejos conscientes, no momento em que é impelido a explicar a própria acção. Evidentemente, isso significa que admito que também razões não conscientes explicam a acção.

Há uma outra questão importante que é: “De onde deve partir uma explicação da acção?” Será que deve partir do conjunto de razões do agente reconstruído ou criado a partir das observações de outra pessoa ou das declarações do próprio agente? Ou será que deve partir da própria acção, que é, na verdade, a única informação razoavelmente segura? Isto é, não seria mais interessante explicar uma acção a partir dela própria, procurando descobrir quais crenças e quais desejos de facto estiveram envolvidos na formação da intenção na acção (I_α) e por que esta acção corresponde ao melhor juízo do agente? Creio que é este o caso.

Figura 10



● $C_{\alpha k} + D_{\alpha k}$ (intenção na acção – eficaz)

● $C_{\beta a} + D_{\beta a}$; $C_{\beta b} + D_{\beta b}$ (intenções na deliberação – ociosas)

Quando alguém age contra o seu melhor juízo, parece não agir por uma razão que racionaliza (no sentido de Davidson) a acção, uma outra forma comum de dizer é que o agente não tem controlo sobre a acção. Essas descrições, tanto no caso do agir sem razões racionalizadoras, como do agir sem controle, traz-nos à questão da identidade do agente. Parece intuitivo dizermos que o “verdadeiro agente” é aquele que visivelmente controla a acção. No entanto, tanto o agente como o observador identificam a acção como tendo sido, de algum modo, efeito da autoridade do agente sobre a acção. Dizemos coisas como “as razões do agente não estavam sobre o seu próprio controle”, ou “o agente não está no controle das suas próprias acções” (Pears, 1984).

Será que uma acção causada por uma razão do agente, mas que vai contra o seu próprio juízo é uma acção, ou um simples acontecimento físico? Isto é, será que esse salto fantasmagórico entre o melhor juízo e a acção é plausível? Será que de facto o que chamamos “melhor juízo” é o melhor juízo do próprio agente? Qual é a plausibilidade psicológica de um agente levar a cabo uma acção intencional contra aquilo que o move à acção? Se respondermos que o melhor juízo não move o agente à acção, teremos que dizer que o melhor juízo não cumpre papel algum além de ser um juízo aparentemente mais forte e certamente intermediário entre o início do processo de deliberação e a própria acção. O conceito de melhor juízo só tem importância porque, aparentemente, é o que motiva eficazmente o agente à acção. Se isso é verdade e aceitarmos que existe a acrasia, teremos que estar dispostos a aceitar que o agente age contra o juízo motivador directamente. Curiosamente, temos todos a sensação de que somos vítimas ou autores de acções produzidas pela fraqueza da vontade. Não é por acaso que esse tema interessa aos seres humanos há tanto tempo. Como é possível termos a sensação de viver a acrasia e mesmo assim ela ser impossível? O primeiro pensamento seria, “bom, ok, a ideia da impossibilidade parece interessante, mas não é plausível, pois sentimos isso ‘na pele’, quase todos os dias”. A resposta a dar aqui, e aquilo que eu defendo, é que aquilo a que chamamos acrasia nada mais é do que um mecanismo psicológico simples de penalização social ou auto-imposta. Noutras palavras, a declaração de acrasia é ela própria uma forma de racionalização, e, curiosamente, nos dois sentidos, no de Davidson e no da Psicanálise²⁸. No primeiro caso, dizer que uma

²⁸ Quando uma sugestão feita durante a hipnose tem efeito só depois, e provoca um ato surpreendente num sujeito, ele frequentemente alega um motivo plausível para sua conduta, a fim de lhe dar uma aparente coerência. O mesmo fenómeno se produz cada vez que uma explicação aparentemente

acção é acrática é uma forma subtil de dar uma explicação, uma razão para aquela acção ter sido realizada, mesmo que não saibamos qual é – joga-se tudo sob o véu útil da fantástica acrasia. No segundo caso, a acrasia também funciona como uma simulação, isto é, como uma máscara que o agente coloca para fugir à declaração frontal acerca das suas próprias crenças e desejos.

No entanto, se tivermos como a informação primeira a levar em consideração a própria acção, parecer-nos-á razoável dizer que uma acção X foi levada a cabo por um desejo W e uma crença Y, mesmo que esses não sejam aqueles declarados como os formadores do melhor juízo *tcc*. Isto porque o melhor juízo, a partir da acção, só poderá ser o juízo eficaz, isto é, aquele que levou o agente a agir de facto.

São muitos os motivos que nos levam a identificar um juízo como sendo o melhor juízo *tcc*, mas não acontece necessariamente que o juízo que identificamos como sendo o melhor juízo o seja eficazmente (JPF *ef*).

racional é utilizada para justificar uma intenção ou um ato cuja determinação inconsciente permanece desconhecida. (...) Trata-se, portanto, de um procedimento de camuflagem, que, tal como a elaboração secundária do sonho, resultaria de uma pressão para a unificação do eu: de fato, com ele o sujeito tenta estabelecer um controle sobre seus pensamentos e seus sintomas, o que mascara a causa inconsciente destes e institui uma lógica do semblante. Essa razão imposta se opõe, portanto, ao reconhecimento da racionalidade dos fenómenos ligados ao saber inconsciente que ultrapassa o sujeito ou lhe escapa. E, evidentemente, a construção de uma teoria é propícia à racionalização: assim, só é possível desmontá-la na elaboração analítica na medida em que esta se mostra capaz de pôr em xeque aquilo que, em si mesmo, é indício dos efeitos do inconsciente (Kaufmann, 1998, p.443).

3.4 UM EXPERIMENTO ACERCA DO CONCEITO POPULAR DE ACRASIA E AS SUAS IMPLICAÇÕES FILOSÓFICAS

3.4.1 FILOSOFIA EXPERIMENTAL [X-PHI]: QUESTÕES BÁSICAS

O objectivo central desta secção é simples: no item seguinte deste capítulo, apresentarei os resultados de um experimento realizado por mim²⁹ acerca da percepção popular da acrasia. Isto é, pretendo, como foi já dito no capítulo 1 e no início do capítulo 3, reconhecendo a importância da concepção popular de acrasia, redefinir este conceito no âmbito filosófico, de forma a aumentar a aderência deste ao que, de facto, parece acontecer no agente real. Essa redefinição passará, como já foi dito, pela eliminação da acrasia como um problema de racionalidade/irracionalidade na acção (3.3), ao mesmo tempo que se propõe que a acrasia pode ser vista e definida como um problema moral.

Para apresentar os referidos resultados é necessário introduzir ao leitor a chamada *filosofia experimental* (x-phi). Foi no âmbito desta nova estratégia metodológica que o experimento foi concebido e realizado. Consciente da estranheza que pode causar a inclusão de um experimento com análise estatística num trabalho de filosofia, julgo ser importante mostrar o âmbito metodológico específico no qual este está inserido.

²⁹Com a participação/ auxílio da Rita Gaio e do Pedro Cardona. Os resultados deste experimento foram apresentados na 17ª Conferência da *European Society for Philosophy and Psychology* (ESPP), em Budapeste, em Agosto de 2009.

Dado esse intuito geral, esta secção tem os seguintes objectivos específicos: (i) introduzir a Filosofia Experimental (x-phi) aos leitores que pela primeira vez têm contacto com o tema; (b) sistematizar um conjunto básico e mínimo de conhecimentos acerca da x-phi de forma a eliminar possíveis dúvidas acerca de questões introdutórias centrais; (c) abrir e incentivar o debate tanto com aqueles que conhecem a x-phi e têm objecções importantes a fazer como com aqueles que já são praticantes mas que podem ter algo a objectar em relação a algumas das ideias acerca da x-phi aqui expostas.

Não tenho conhecimento da existência de outros textos introdutórios no âmbito da filosofia experimental, escritos em Português. Daí que tenha escrito esta secção sobretudo com o propósito de suprir essa falta e de cumprir esse propósito. Por essa razão, o leitor não deve esperar uma secção estritamente técnica acerca das implicações profundas da filosofia experimental, nem uma descrição exaustiva de todos os casos estudados. Aquilo que se pretende é fornecer elementos suficientes ao leitor para que este possa objectar o âmbito no qual o referido experimento foi concebido e realizado.

Para minimizar os efeitos de alguma parcialidade derivada do facto de eu ser um investigador ligado à x-phi, sugiro a leitura de alguns críticos importantes da x-phi, tais como Sosa (2007a) ou Kauppinen (2007), assim como as respostas dos representantes da x-phi a algumas dessas críticas (Nadelhoffer e Nahmias, 2007).

DESMITIFICAÇÕES NECESSÁRIAS

A primeira parte desta secção é necessária para acalmar os ânimos dos leitores que têm crenças elementares equivocadas acerca da x-phi. Esta parte é importante para que a leitura do resto do texto seja produtiva – para evitar que concepções distorcidas do que a filosofia experimental é e pretende estejam latentes durante a leitura do que se segue. Para aqueles que nada sabem acerca da x-phi, a sugestão é que essa secção seja lida para que se evite que tais más concepções elementares apareçam.

Antes de avançar uma explicação mais directa acerca da filosofia experimental, é preciso dizer, para uma imediata “desmitificação”, que a x-phi *não* é um movimento que prescreve a solução de problemas filosóficos através da estatística, do cálculo matemático ou das pesquisas/experiências científico-sociais, assim como *não* defende que o único e exclusivo caminho da filosofia é perguntar às pessoas comuns o que elas pensam acerca deste ou daquele conceito, ou desta ou daquela intuição.

Há muitos filósofos que pensam que a x-phi consiste em perguntar às pessoas o que elas pensam, ou que os filósofos experimentais prescrevem a utilização dos métodos experimentais e quantitativos para resolver ou solucionar problemas filosóficos. Curiosamente, tais filósofos acreditam em mitos como estes acerca da X-Phi sem aparente justificação – uma vez que os principais autores e os artigos centrais da x-phi³⁰ não defendem tais ideias. Por conta dessa atitude, tem havido

³⁰ Veja-se, por exemplo, os artigos contidos em Knobe e Nichols, 2008.

alguma precipitação e confusão acerca da x-phi. Basta o interessado ler dois ou três destes artigos centrais para perceber do que trata a x-phi.

É preciso, porém, distinguir, entre os praticantes da filosofia experimental, como fez S. Matthew Liao (2008), entre *experimentalistas radicais* e *experimentalistas moderados*. Os experimentalistas radicais pensam que a filosofia experimental *mina* outros métodos-padrão de fazer filosofia, enquanto os experimentalistas moderados pensam que ela os *suplementa*. (May, forthcoming)

Segundo Liao (2008),

“Nos últimos anos, uma série de filósofos vem realizando estudos empíricos que investigam as intuições das pessoas sobre diversos temas da filosofia. Outros filósofos, como Gilbert Harman e Frank Jackson indicaram também o apoio a este tipo de abordagem empírica à filosofia, mesmo que não tenham (ainda) realizado tais estudos empíricos. Podemos chamar esta abordagem empírica à filosofia de “experimentalismo” e aqueles que são sensíveis a esta abordagem de “experimentalistas”. (...) No conjunto, o experimentalismo parece ser um bom desenvolvimento para a filosofia. Quando filósofos colocam os argumentos, muitas vezes, apelam para intuições. Se algo é intuitivo, tende a contar a favor de uma determinada posição, e se algo é contraintuitivo, tende a contar contra esta posição. Além disso, se um filósofo descobre que outros têm intuições diferentes das dele acerca de um caso específico, tal revelação pode ser bastante informativa, porque ele deverá, então, investigar se existe um verdadeiro conflito e, em caso afirmativo, tentar verificar quem está errado, se ele ou os outros. Podemos chamar a isto de abordagem filosófica da Intuição como Evidência (IAE). Na medida em que o experimentalismo tem o potencial para descobrir intuições duvidosas, parece ser algo que os filósofos devem saudar. No entanto, entre aqueles que são simpáticos ao experimentalismo, é possível distinguir entre os “experimentalistas moderados”, que acreditam que o experimentalismo pode complementar a IAE, e os “experimentalistas radicais”, que parecem sustentar que o experimentalismo deveria substituir a IAE.”

Liao dá o nome de experimentalistas aos comumente chamados filósofos experimentais e distingue entre dois tipos: os moderados e os radicais. Aqui não

cabe promover uma discussão metafilosófica de fundo, porém é importante clarificar o meu posicionamento acerca desta divisão proposta por Liao.

A versão da x-phi que me seduz é claramente a versão moderada. O experimento realizado (ver item 3.4.2) segue a ideia moderada de que é possível, através da realização de experimentos específicos, o filósofo sofisticar e refinar as intuições que tem, sem a necessidade de abandoná-las. Na verdade, nesta tese o apelo às intuições e à psicologia popular constitui a principal base metodológica. O experimento, acima referido, parece-me relevante porque é uma forma de refinar, ou avaliar, uma intuição específica acerca de como as pessoas comuns percebem e definem a acrasia. Além disso, qualquer das explicações deste fenómeno, seja uma explicação de tipo davidsoniana, ou a minha tentativa de explicação, baseia-se na psicologia popular e certamente em intuições do proponente acerca do que se passa no agente real.

Jonathan Weinberg e Joshua Alexander (2007), apontados por Liao como sendo experimentalistas radicais – aqueles que sustentam que o experimentalismo deveria substituir a filosofia baseada em intuições como evidência (IAE), criaram uma tipologia das formas de apelo às intuições, constituída pelos seguintes tipos: o solipsista, o elitista e o populista. Abaixo, segue uma descrição (Alexander & Weinberg, 2007) mínima destes tipos:

Em primeiro lugar, poderia supor-se que quando um filósofo invoca as intuições como evidência está confiando apenas nas suas próprias intuições como prova. Vamos chamar este ponto de vista solipsismo das intuições (intuition solipsism). Em segundo lugar, ele pode estar a confiar nas suas próprias intuições porque considera que estas intuições são representativas das intuições da classe profissional dos filósofos. Vamos chamar este ponto de vista elitismo das intuições (intuition elitism). Em terceiro lugar, ele pode estar a confiar nas suas próprias intuições porque as toma como

representativas das intuições de uma classe mais ampla que inclui os não-filósofos - vulgarmente designada por " povo". Vamos chamar este ponto de vista populismo das intuições (intuition populism).

Sem querer alongar-me num tema que é paralelo a esta tese, gostaria de fazer algumas pequenas observações acerca dos experimentalistas radicais e desta tipologia.

Parece-me razoável dizer que qualquer filósofo é capaz de reconhecer a existência destas diferentes formas de tomar as intuições. No entanto, é preciso saber se o filósofo tem como fazer filosofia sem o apelo às intuições, principalmente nas áreas que esta tese abrange.

A questão é saber se os filósofos têm condições metodológicas para fazer filosofia sem o apelo às intuições, seja na formulação do problema, seja na argumentação. O que me interessa, assim como deve interessar aos filósofos experimentais moderados, é poder utilizar métodos experimentais e métodos quantitativos para aumentar a quantidade de dados disponíveis ao trabalho filosófico.

Repare-se que a utilização de resultados científicos pelos filósofos não é de todo nova. Alguns podem dizer que isso já bastaria, pois o filósofo pode utilizar tais resultados para fazer filosofia. Isto é, não seria necessária a ida do filósofo ao laboratório, ou ao campo, à procura de dados. O problema é que quase sempre os experimentos não são feitos à medida dos interesses filosóficos. Isto é, as ciências seguem a sua própria agenda e não se desviam (nem me parece que deveriam desviar-se), com raras exceções, para atender às necessidades da filosofia. Assim sendo, os filósofos vão, de certa forma, a reboque da agenda científica. É claro que

há exemplos de questões levantadas por filósofos e que se tornaram relevantes no meio científico, principalmente na filosofia da mente. Porém, em outras áreas, como a filosofia da acção, não parece ser este o caso. A aproximação do filósofo das áreas experimentais parece ser de grande valia para a filosofia e também para as ciências.

Algumas das reacções mais inflamadas contra a x-phi, além daquelas derivadas do experimentalismo radical, parecem ser o resultado de uma indistinção entre as questões filosóficas inerentes à x-phi e elementos extra-filosóficos ou não-filosóficos, envolvendo todo um conjunto de acessórios como logótipos, vídeo clipes, t-shirts, entre outras coisas, que têm surgido de maneira espontânea e supostamente a favor da filosofia experimental – “supostamente” porque em alguns casos o seu efeito parece ser nocivo. O logótipo mais conhecido da x-phi é um cadeirão incendiando, como uma alusão ao suposto declínio da chamada “filosofia de cadeirão” – como veremos adiante, identificada com uma filosofia que supostamente prescreve o conhecimento *a priori* como o único a ser produzido pela “verdadeira filosofia”.

A x-phi tem conquistado adeptos em todos os grupos científicos relevantes nos Estados Unidos, especialmente entre os jovens alunos de licenciatura, de mestrado e de doutoramento. Alguns parecem mesmo adeptos, no sentido futebolístico da palavra. Como ilustração desse tipo de “folclore”, sugiro ao leitor que entre na página da filosofia experimental no myspace³¹, que assista ao clipe do hino da x-phi³² e veja as t-shirts à venda³³, assim como a selecção de músicas x-phi³⁴. Esta é

³¹ <http://www.myspace.com/experimentalphilosophy>

³² <http://www.youtube.com/watch?v=tt5Kxv8eCTA>

³³ <http://www.cafepress.co.uk/xphi>

uma imagem de grupo bastante forte, que tem associado algum fundamentalismo. Trata-se de um radicalismo algo ingénuo que vê a filosofia experimental como a nova panaceia e que encara o debate como uma guerra entre “novos” e “velhos” métodos, da qual só sairá um vencedor.

Pensamos que as reacções teóricas acesas, inflamadas e polarizadas parecem ocorrer precisamente por causa do ataque que os filósofos experimentais fazem ao método que foi predominante na filosofia analítica durante o século XX. Isto é, a X-Phi ataca frontalmente a ideia de que cabe ao filósofo realizar exclusivamente trabalhos metodologicamente apriorísticos – a esse modo de fazer filosofia, chamam “filosofia de cadeirão”. Como foi dito, partindo deste ataque, alguns adeptos da x-phi passaram a actuar fora do campo estrito da filosofia, talvez numa estratégia inocente de marketing, e isso tem provocado algum desconforto a alguns filósofos experimentais. O logótipo mais conhecido da x-phi, acima referido (a poltrona incendiando, em alusão à ideia de que essa filosofia de poltrona está, em muitos casos, com os dias contados), tem provocado reacções polarizadas e completamente afastadas do debate filosófico mais importante. Isso tem, para alguns, desviado o debate para questões absolutamente inúteis e que provocam muita incompreensão acerca da x-phi. Porém, o debate central permanece, em princípio, imune a essa distorção. Têm sido publicados artigos, respostas e críticas e têm sido realizados experimentos que não estão contaminados por essa discussão não-filosófica, e será este o meu foco. Isso tudo acontece principalmente nos Estados Unidos, onde se baseia grande parte do conjunto de filósofos

³⁴ <http://www.mixtapecollective.org/mixtape/?mix=2740>

experimentais que tem realizado experiências e publicado artigos sob a denominação de filosofia experimental.

É importante ressaltar que em nenhum dos elementos centrais actuais da x-phi – o x-phi-blog³⁵, a XPS (Experimental Philosophy Society)³⁶, e o livro x-phi – há qualquer referência relevante, à excepção do cadeirão incendiado, a algum tipo de marketing agressivo e desproporcionado. Não é uma marca através da qual os filósofos experimentais querem conquistar terreno dentro da filosofia. Essas coisas surgiram espontaneamente e não fazem parte, ao que parece, de uma política deliberada de combate filosófico. Mesmo porque esse seria um enorme erro estratégico.

O x-phi blog, por exemplo, é uma excelente fonte de informação e é, principalmente, um grande fórum de debate acerca da x-phi. Há um blog em português chamado “filosofia experimental” que tenta reunir interessados para formar um fórum através do qual possam ser discutidos experimentos, questões metodológicas e questões de fundo acerca da x-phi.

Em 2009 foi criada a “Experimental Philosophy Society”, que tem em seu comité executivo investigadores, como por exemplo, Stephen Stich (filosofia e ciência cognitiva: Rutgers), Shaun Nichols (filosofia: Arizona), Joshua Knobe (filosofia e ciência cognitiva: Yale), Joshua Greene (psicologia: Harvard) e Jonathan Weinberg (ciência cognitiva: Indiana).

Em 2008 foi publicado pela Oxford Press o primeiro livro dedicado à X-Phi, com o título “Experimental Philosophy”, editado pelos filósofos Joshua Knobe e Shaun Nichols. Muitos artigos têm sido publicados sob a denominação X-Phi nos

³⁵ <http://experimentalphilosophy.typepad.com/> e <http://filosofiaexperimental.wordpress.com/>

³⁶ <http://x-phi.org/>

últimos 10 anos, mas só em 2008 foi publicado esse livro dedicado exclusivamente à explicação do que é a X-Phi, através de um “Manifesto” e de exemplos de investigações/ experimentos levados a cabo por filósofos experimentais, como Stephen Stich, Edouard Machery, Shaun Nichols, Joshua Knobe, Jesse Prinz, entre outros.

A ligação entre a filosofia e a ciência cognitiva está na base deste movimento. Um dos representantes/fundadores do movimento é o filósofo Stephen Stich que tem uma história vasta de investigações na filosofia e na ciência cognitiva. Muitos filósofos experimentais têm ligações teóricas e experimentais directas com outras áreas como, por exemplo, a psicologia, a antropologia, a neurociência, etc. É o caso de Joshua Knobe e Bertram Malle, um filósofo e um psicólogo experimental, respectivamente, que uniram esforços no sentido de tentar perceber qual é a concepção comum do conceito de “intencionalidade”, mais concretamente, qual é o critério que as pessoas normalmente usam para decidir se um dado efeito colateral de uma acção pode ser considerado intencional. Uma outra ligação acontece entre a filosofia experimental e a psicologia do desenvolvimento. Esta área do conhecimento preocupa-se sobretudo com a forma como as crianças entendem determinados conceitos, e como essa concepção evolui ao longo do tempo. Shaun Nichols (Nichols & Folds-Bennett, 2003) é um filósofo que, com a psicóloga T. Folds-Bennett, delineou experiências em psicologia do desenvolvimento com o intuito de perceber se a noção de “moralmente errado” é ou não, por defeito (*by default*), uma noção objectivista entre as crianças, e como essa percepção se altera na idade adulta. Este é de resto um bom exemplo para se perceber o que está e o que não está em causa: o facto de essa ser a intuição das crianças não significa que

se possa defender, sem mais, que os juízos morais são objectivos; no entanto, saber como é que a noção de objectividade é adquirida pode lançar luz sobre a legitimidade da sua aplicação.

O QUE É A X-PHI: UMA ABORDAGEM GERAL

Tendo sido feitas essas observações iniciais, posso passar a tentar explicar o que é a filosofia experimental.

Parece ser uma boa estratégia, no caso de uma primeira abordagem geral, utilizar extractos de dois artigos extremamente básicos de divulgação da x-phi. Assim, talvez possa desenvolver uma boa e intuitiva aproximação ao assunto, atendendo ao requisito básico de olhar para o que de facto dizem os filósofos experimentais cujos trabalhos são reconhecidamente centrais.

Abaixo, cito dois parágrafos de dois textos diferentes, um do Joshua Knobe e outro do Shaun Nichols e vou identificar e numerar secções desses parágrafos que serão, posteriormente, tratadas em conjunto ou separadamente, de modo a tentar aproveitar o máximo possível o conteúdo dessas citações.

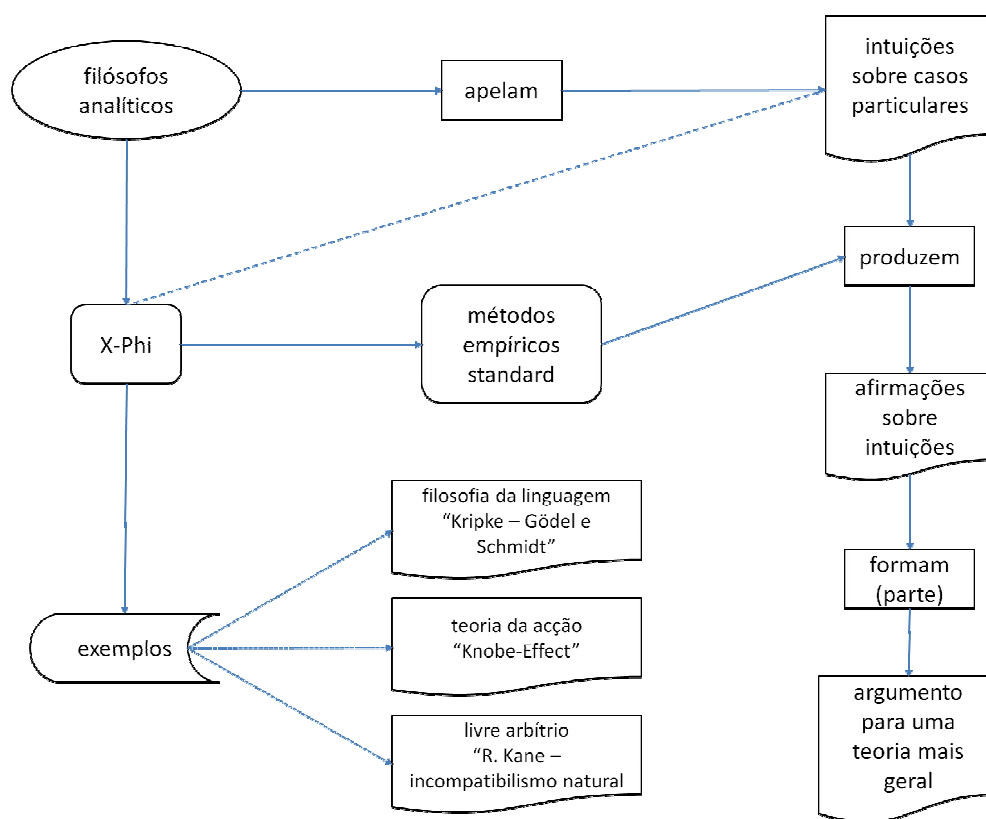
Antes disso, no entanto, vou tentar desenhar, resumida e basicamente, diagramas que sejam intuitivos e que representem o que está escrito nos parágrafos seleccionados. É importante ressaltar que os parágrafos seleccionados não representam a totalidade do pensamento dos dois autores acerca da x-phi. Esses parágrafos foram seleccionados, pois representam, talvez, o que de mais básico se pode dizer acerca da x-phi. É importante ter-se em mente que há outras formas de abordar e definir o que é e qual é a abrangência da filosofia

experimental³⁷. A abordagem do J. Knobe é aquela que serve como referência essencial nesta tese.

Sobre a x-phi, Joshua Knobe (forthcoming) escreve: (os números foram introduzidos por mim, com o objectivo acima referido)

“[1] Desde os primeiros dias da filosofia analítica, tem sido prática comum apelar às intuições sobre casos particulares. Tipicamente, o filósofo apresenta uma situação hipotética e depois produz uma afirmação na seguinte forma: “nesse caso, podemos certamente dizer...”. Essa afirmação sobre as intuições das pessoas forma, depois, parte de um argumento para uma teoria mais geral sobre a natureza dos nossos conceitos ou sobre o nosso uso da linguagem. [2] Um aspecto complicado ou intrigante dessa prática é que raramente se faz uso de métodos empíricos standard. Apesar de os filósofos muito frequentemente produzirem afirmações sobre “o que as pessoas podem ordinariamente dizer”, eles raramente se asseguram acerca dessas afirmações perguntando realmente às pessoas e procurando padrões nas respostas. Em anos recentes, no entanto, alguns filósofos têm colocado essas afirmações sobre intuições em teste, utilizando métodos experimentais para revelar o que as pessoas realmente pensam acerca de casos hipotéticos particulares. Algumas vezes, os resultados têm sido extremamente surpreendentes.”

³⁷ refs

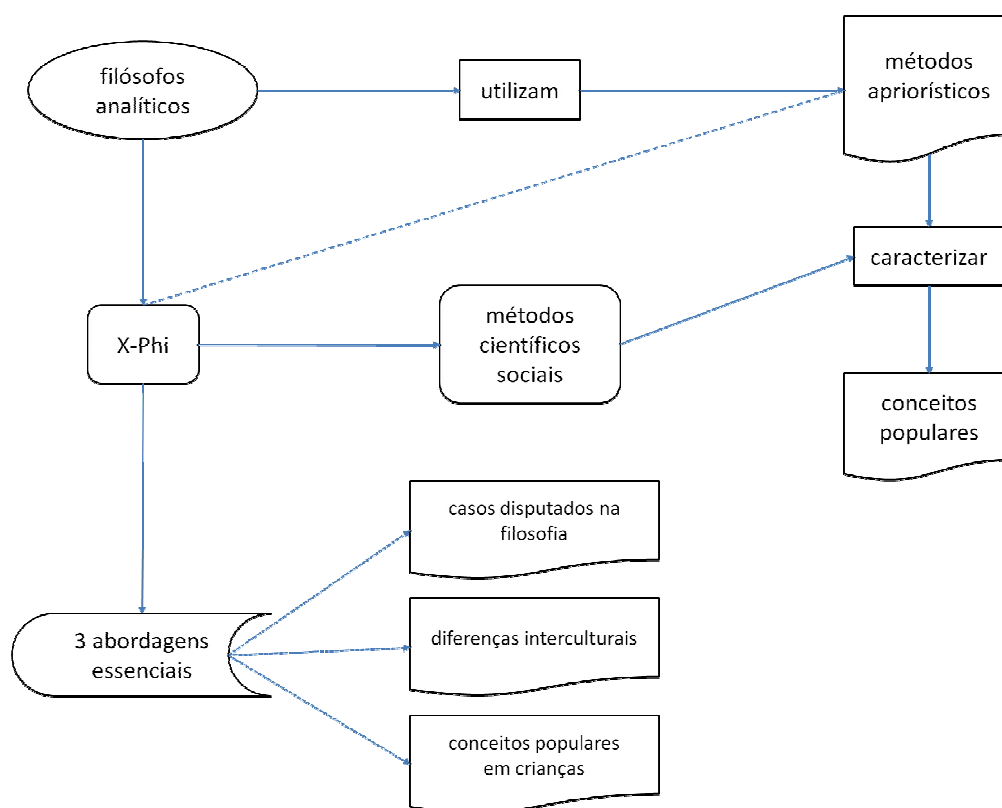


O diagrama mostra resumidamente que a ideia central desse parágrafo é mostrar que os filósofos experimentais utilizam métodos empíricos para produzir afirmações acerca das intuições das pessoas de maneira mais confiável. Isto é, nos casos em que é importante para a filosofia saber o que as pessoas pensam, ou quais as intuições das pessoas acerca desse ou daquele conceito, a filosofia experimental fornece um caminho mais seguro e confiável. Ou seja, essas crenças serão formadas a partir de uma outra fonte. Os três exemplos de Knobe serão explicados mais à frente.

Sobre a x-phi, Shaun Nichols (2004b) escreve:

[1] Os filósofos analíticos têm desde sempre utilizado métodos apriorísticos para caracterizar conceitos populares como o conhecimento, a crença, ou 'ser moralmente errado'. Recentemente,

alguns investigadores têm vindo a explorar metodologias das ciências sociais para caracterizar esses conceitos populares. Uma linha de trabalho tem explorado as intuições populares acerca de casos que são disputados dentro da filosofia. Uma segunda abordagem, com implicações potencialmente mais radicais, aplica métodos da psicologia intercultural às intuições filosóficas. Trabalhos recentes nessa área sugerem que as pessoas em diferentes culturas têm sistematicamente diferentes intuições acerca de conceitos populares. Uma terceira linha de investigação explora a emergência e a natureza de conceitos populares nas crianças. Essas abordagens para caracterizar conceitos populares fornecem recursos importantes que vão complementar, e talvez em alguns casos dispensar, abordagens apriorísticas.



O diagrama mostra resumidamente que a ideia central deste parágrafo é mostrar que os filósofos experimentais utilizam métodos científicos sociais ao invés de métodos apriorísticos para caracterizar conceitos populares que são importantes para a filosofia. Nos dois autores, o sentido é basicamente o mesmo e a forma de caracterizar a x-phi é também a mesma. Os exemplos que Nichols e

Knobe nos oferecem são praticamente os mesmos e, por uma questão de espaço, serão aqui descritos apenas alguns desses exemplos.

Tentando, agora, clarificar e aprofundar um pouco o que dizem esses dois autores, passo à análise das secções identificadas e numeradas nestes textos.

Nas secções [1] e [1'] desses parágrafos, Knobe e Nichols chamam-nos a atenção para o facto de que, na filosofia, apelamos constantemente às nossas intuições como pensadores e às intuições de um conjunto vasto de pessoas – generalizando muitas vezes determinadas intuições acerca das intuições dos outros. Os leitores da filosofia analítica estão bastante acostumados a ler justificações parciais de premissas que compõem um argumento expressas da seguinte maneira: “as pessoas intuitivamente pensam A, portanto, nesse caso, podemos afirmar B”, ou “no caso em que C, as pessoas intuitivamente pensam D, portanto, nesse caso, podemos afirmar F”. Parece que temos estado dispostos a aceitar determinadas intuições e a fazer isso de forma sistemática e aparentemente passiva. A questão é: será que temos sido suficientemente prudentes ao aceitar pacificamente determinadas intuições acerca de tópicos filosoficamente complexos, intuições que parecem evidentes demais para serem questionadas?

Um dos exemplos emblemáticos da X-Phi é aquele no qual se discute o livre arbítrio através da intuição de que as pessoas são “naturalmente” incompatibilistas. Robert Kane, afirma (veja-se Knobe, forthcoming):

“Na minha experiência, as pessoas comuns, na sua maioria, começam por ser incompatibilistas naturais. Elas acreditam que há alguma espécie de conflito entre liberdade e determinismo; e a ideia de que liberdade e responsabilidade podem ser compatíveis com o determinismo parece-lhes à primeira vista como um ‘quagmire of evasion’ (William James) ou um “subterfúgio

miserável” [wretched subterfuge] (Immanuel Kant). As pessoas comuns têm de ser convencidas a abandonar este incompatibilismo natural pelos argumentos inteligentes dos filósofos.”

Kane parte do pressuposto, portanto, de que o compatibilismo é contra-intuitivo. Através da filosofia experimental, os filósofos Eddy Nahmias, Thomas Nadelhoffer, Jason Turner e Steve Morris quiseram testar isso. Aos participantes na experiência, apresentaram a seguinte situação (veja-se Knobe, forthcoming):

“Imagine que no próximo século descobrimos todas as leis da natureza, e construímos um super computador que pode deduzir a partir destas leis da natureza e do actual estado de coisas no mundo, exactamente tudo o que acontecerá no mundo em qualquer altura do futuro. Pode verificar tudo acerca de como o mundo é, e prever tudo acerca de como será com 100% de exactidão. Suponha-se que tal super computador existe, e olha para o estado do universo, num determinado ponto, a 25 de Maio de 2150 d.C., vinte anos antes de Jeremy Hall nascer. O computador deduz então, a partir desta informação e das leis da natureza, que Jeremy vai, com toda a certeza, assaltar o Fidelity Bank às 18:00h do dia 26 de Janeiro de 2195. Como sempre, o super computador está correcto; Jeremy assalta o Fidelity Bank às 18:00h do dia 26 de Janeiro de 2195.”

Em seguida, questionam-se os participantes. A pergunta colocada foi: sera que Jeremy é moralmente responsável pelo assalto ao banco? A maioria diz que sim, indicando que a intuição inicial de Kane não estaria suficientemente fundamentada, ou, no mínimo, é problemática e precisa ser reformulada.

Na secção [2] do parágrafo, Knobe chama-nos a atenção para o facto de que se fazemos afirmações acerca do que as pessoas pensam, ou do que elas podem dizer, mas não utilizamos métodos standard já consagrados em áreas científicas como, por exemplo, a Psicologia Social, poderemos cometer erros filosóficos importantes.

Alguns filósofos e alguns simpatizantes da filosofia têm considerado abusivo, para não dizer inadequado, o uso desses métodos empíricos standard. Dizem que existem áreas de especialização razoavelmente bem definidas e que os filósofos, por norma, não seriam bons experimentadores, não realizariam bons cálculos e que, portanto, deveríamos deixar a realização de experiências e de cálculos para os psicólogos, físicos, químicos, neurocientistas etc.. Caberia, então, ao filósofo a utilização dos dados produzidos pelos cientistas como fonte para os seus insights.

Estas afirmações espantam-me: como é possível dizer que um filósofo não é capaz de estruturar uma boa experiência ou que a estatística é um “bicho-de-sete-cabeças”? Um filósofo deve ser capaz de compreender coisas muito mais complexas e abstractas do que os métodos experimentais e estatísticos em questão. Pode aliás, em pouco tempo, ser capaz de aprender a lidar com determinados métodos estatísticos, tal como fazem os economistas, os psicólogos, os sociólogos etc. Além disso, na filosofia experimental, é comum a parceria entre investigadores de áreas diferentes, como filósofos e psicólogos, filósofos e antropólogos, economistas, etc.

Não vejo realmente nada que possa ser objectável no que diz respeito à capacidade do filósofo de realizar experiências utilizando métodos quantitativos e experimentais.

Além disso, parece extremamente imobilizador e frustrante querer saber algo acerca de uma intuição que está disponível “a um experimento de distância”, isto é, a um pequeno passo, e nós, como filósofos, não podermos querer saber (o que é pior do que não poder saber) porque “não é esse o nosso papel”.

Não me parece que os filósofos disponham de dados científicos suficientes que abarquem as necessidades específicas deles. Daí a importância de colocarem a mão na massa e produzirem os dados que para eles fazem mais sentido – dados que estejam directamente ligados às perguntas que fazem e que realmente querem tentar responder.

Não quis aqui dizer que esses dados vão responder às perguntas filosóficas, nem dizer que passaremos a validar sistematicamente as respostas filosóficas a partir de experiências. No entanto, uma filosofia ligada directamente ao laboratório e ao campo parece ser uma filosofia extremamente interessante e excitante. Esses dados poderão, muitas vezes, criar mais e maiores problemas filosóficos do que o simples apelo às intuições dos filósofos acerca do que as pessoas pensam.

Uma outra linha de investigação da filosofia experimental liga-se à amplitude das generalizações em que normalmente incorremos. Mesmo que se desse o caso de as intuições dos filósofos coincidirem com as intuições das pessoas comuns, a que pessoas estaríamos a referir-nos? Podemos por exemplo partir do princípio absoluto de que as pessoas que se inserem na cultura ocidental partilham concepções acerca de tópicos variados como por exemplo a causalidade ou a intersubjectividade com aquelas que se integram numa cultura diferente?

Os filósofos Edouard Machery, Ron Mallon, Shaun Nichols e Stephen Stich (Machery, Mallon, Nischols e Stich, 2004) conduziram um estudo empírico³⁸ cujo objectivo era testar essa ideia no que toca à forma como é entendida a noção de referência.

³⁸ ref

Para isso, usaram um exemplo célebre usado por Saul Kripke em filosofia da linguagem e metafísica, através do qual ele pretende mostrar como a teoria da referência à qual ele se opõe (a teoria descritivista) é contra-intuitiva. O caso que ele coloca é o seguinte (veja-se Knobe, forthcoming):

Suponha que Gödel não era de facto o autor do teorema [de Gödel]. Um homem chamado 'Schmidt'... na verdade fez o trabalho em questão. O seu amigo Gödel, de alguma forma, conseguiu a posse do manuscrito, e a partir daí foi atribuído a Gödel. Numa determinada perspectiva [a então popular teoria 'descritivista'], quando o homem comum usa o nome 'Gödel', ele pretende realmente referir-se a Schmidt, porque Schmidt é a única pessoa que satisfaz a descrição 'o homem que descobriu a incompletude da aritmética'.

De acordo com a tal teoria descritivista, o nome refere aquele objecto que melhor se adequa à descrição associada a esse nome. Portanto, se o nome “Gödel” significa “aquela pessoa que descobriu a incompletude da aritmética”, então, quando dizemos “Gödel” estamos na verdade a referir-nos a Schmidt.

A maioria dos filósofos analíticos foi de encontro à opinião de Kripke de que algo de errado se passava com esta história. E na experiência levada a cabo por Machery et. al. (2004) , os indivíduos originários dos EUA também concordaram com a ideia de que algo está mal numa teoria da referência na qual “Gödel” pode referir Schmidt. No entanto, surpreendentemente, o mesmo não se verificou nas respostas dos indivíduos que viviam em Hong Kong. Entre estes, a maioria disse que a palavra “Gödel” de facto se referia a Schmidt.

Portanto, utilizar como reforço num argumento filosófico a ideia de que uma intuição é óbvia e partilhada por todos é algo que deve ser tratado com mais cuidado.

O QUE É A X-PHI: UMA DISCUSSÃO DENTRO DA FILOSOFIA

A x-phi e a análise conceptual não estão em lados completamente opostos – elas partilham interesses em tópicos comuns e também o uso de intuições das pessoas sobre casos específicos. A diferença mais superficial entre os filósofos experimentais e os analistas-conceptuais é que o primeiro grupo utiliza métodos experimentais e o segundo não. O analista conceptual escreve algo como: “Nesse caso, alguém certamente diria.....”. O filósofo experimental diria: “Nesse caso, 77% dos participantes disseram...”. Essa comparação entre o que diria um e outro é já muito conhecida e aparece quase sempre como uma comparação superficial, mas importante, entre as duas.

A análise conceptual procura identificar o significado preciso de um conceito partindo-o em seus componentes essenciais, os quais envolvem outros conceitos. Tentando determinar o significado de um conceito, filosoficamente relevante, verificamos através da análise conceptual se o conceito se aplica em vários casos possíveis, tentando chegar a níveis cada vez mais precisos.

Sobre a análise conceptual, Knobe e Nichols (Knobe & Nichols, 2008, pp. 4 e 5) escrevem:

Este programa de investigação é, de qualquer perspectiva, extremamente difícil. O filósofo esforça-se por construir um conjunto de condições necessárias e suficientes para o conceito investigado, digamos lápis. Mas então, quando apresenta os seus resultados, acontece inevitavelmente que alguém no fundo da sala dá um exemplo de um objecto que cumpre todas as condições mas que não é um lápis. Isto leva o filósofo de volta para o seu estudo para fazer alguns ajustes à sua definição. O objectivo da análise conceptual é altamente controverso. Há quem considere que está a fazer progressos consideráveis e que eventualmente convergirá em análises correctas de certos conceitos filosóficos importantes; outros consideram que nunca fomos capazes de analisar nada em

termos de outra coisa e que este falhanço denuncia uma falha intrínseca nas suposições em que o próprio projecto se funda. Independentemente do que pensemos acerca desta controvérsia, é importante compreender o modo como o objectivo da filosofia experimental difere do da filosofia tradicional.

De facto, independentemente do que se pensa sobre a controvérsia acerca da análise conceptual, deve ficar claro que a x-phi constrói um caminho complementar à análise conceptual através dos métodos experimentais e também através de conjuntos de estratégias e objectivos diferentes. Enquanto a análise conceptual busca a identificação mais precisa possível de um conceito através da sua divisão em componentes essenciais que acabam por envolver outros conceitos, a x-phi procura explicar os factores que influenciam a aplicação de um conceito, principalmente os processos internos que estão por trás da sua aplicação.

De forma mais precisa e também para resguardar a análise original, cito novamente Knobe e Nichols (Knobe & Nichols, 2008, pp. 4 e 5):

Tanto quanto sabemos, nenhum filósofo experimental forneceu uma análise de um conceito em termos de outro. Em vez disso, o objectivo é usualmente o de fornecer uma explicação dos factores que influenciam a aplicação de um conceito, e em particular, os processos psicológicos internos que subjazem à sua aplicação. O progresso é medido não em termos de precisão com os quais se pode caracterizar os padrões actuais das intuições das pessoas mas em termos do grau de profundidade explicativa que podemos atingir.

POSSÍVEIS OBJECÇÕES

Tal como referi inicialmente, há algumas críticas válidas endereçadas à filosofia experimental, das quais alguns dos principais autores da filosofia experimental procuram defender-se de uma forma que considero segura. Os

exemplos que menciono em seguida foram retirados daquele que é o texto fundador do movimento – o “manifesto” da filosofia experimental (Knobe & Nichols, 2008).

1) Por toda a academia, confiamos nos especialistas para que garantam a pesquisa. Seria absurdo que biólogos ou físicos fizessem inquéritos sobre intuições comuns acerca de biologia ou de física. Em vez disso, os físicos e os biólogos especializam-se nas suas áreas e contribuem para a sua área usando o seu conhecimento especializado. O mesmo é verdade na filosofia. Da mesma forma que os físicos não consultam leigos em física, os filósofos não têm de consultar leigos de filosofia.

A resposta para uma objecção como esta pode ser a seguinte: os problemas com que biólogos ou físicos se deparam não dependem da concepção que as pessoas têm acerca da estrutura física do mundo, por exemplo, por isso essa concepção está naturalmente arredada da investigação científica. Ao contrário, alguns problemas em filosofia provêm, precisamente, de uma tensão entre o que pode ser considerado uma concepção comum e o modo como o mundo tem que ser de forma a acomodar essa concepção. O problema do realismo moral, por exemplo, emerge do choque entre a concepção que, implicitamente, temos quando elaboramos um juízo moral – consideramos que se trata de um juízo objectivamente verdadeiro – e a imagem que temos do mundo natural como um local improvável para aí encontrarmos coisas como valores morais.

2) É verdade que nos preocupamos com questões acerca de conceitos de senso-comum. A questão importante é que os filósofos podem usar esses mesmos conceitos – os vulgares conceitos de senso-comum que as pessoas usam todos os dias – com uma precisão e subtileza que as pessoas comuns não conseguem atingir. Pois os filósofos estão especialmente treinados para traçarem distinções precisas e para pensarem cuidadosamente; e os filósofos trazem estas capacidades para conseguirem descortinar a

verdadeira natureza das nossas intuições de senso-comum. Como resultado, os filósofos possuem uma capacidade muito mais desenvolvida para chegarem a definições descomprometidas sobre casos do que as pessoas comuns.

Este ponto, defendem os filósofos experimentais, é algo que precisa, ele próprio, de fundamentação empírica. Podemos afirmar, sem mais, que o filósofo tem um acesso privilegiado ao conhecimento, por ter tido uma formação especializada, com domínio de técnicas próprias da filosofia? Como podemos ter a certeza que as intuições dos filósofos não estão, também elas, sujeitas aos factores (constrangimentos cognitivos, culturais, de classe, etc.) que por vezes distorcem e influenciam as intuições comuns? E não está a visão do filósofo, por outro lado, contaminada por preconceitos e defeitos de formação, de uma forma que a visão das pessoas comuns não está?

3) Estão simplesmente a passar ao lado do objectivo da filosofia. A filosofia não é somente uma questão de olhar para as intuições das pessoas e de tentar entender como as pessoas pensam. Em vez disso, quando fazemos filosofia, necessitamos de sujeitar as intuições das pessoas a um escrutínio, olhando para os argumentos que podem mostrar que as intuições das pessoas estão erradas em certos casos.

Esta é uma objecção que não melindra a filosofia experimental, na medida em que é um ataque ao lado relativamente àquilo que a filosofia experimental se propõe fazer. Isto é, em nenhum momento, este novo movimento toma como objectivo a substituição da análise filosófica pela prospecção da opinião comum. O que pretende é lançar mais um dado a ser objecto dessa análise – mostrar em que é que as intuições das pessoas comuns podem ser diferentes das intuições dos filósofos, e porquê, e o que é que isso revela acerca dos próprios problemas em questão.

4) Nunca conseguirão chegar a lado nenhum se simplesmente fizerem uma grande quantidade de experiências. Assim, suponha que nos interrogamos acerca de alguma questão específica em filosofia moral. Podemos descobrir que um determinado processo psicológico tende a produzir um tipo particular de intuição sobre certas questões, mas que tal conhecimento não nos ajudará a não ser que já possuamos alguma informação sobre se o processo é fiável ou as intuições correctas. E como vamos descobrir isso? Certamente não vão fazer mais experiências!

Certamente que não. Mas o ponto é que esses resultados podem fazer-nos pensar melhor acerca das teorias de que partimos, e dos factores, nomeadamente psicológicos ou culturais, que devem ser levados em consideração na argumentação filosófica. Os padrões que podem ser descobertos uma vez analisadas as respostas das pessoas comuns podem dizer-nos coisas acerca do funcionamento da mente que são elas próprias relevantes até para a formulação dos problemas filosóficos subjacentes.

5) Uma outra crítica diz respeito ao tipo de intuições que a filosofia experimental pode captar: alguns críticos (Kauppinen 2007) defendem que a filosofia experimental só é capaz de chegar até às intuições superficiais e não às intuições robustas, que são as que interessam aos filósofos no seu trabalho conceptual, e que são produzidas em condições ideais – isto é, sem falhas de competência ou de performance, e que só são descobertas através do diálogo participativo (método socrático), ou então através da reflexão individual de sujeitos treinados (os filósofos), mas não através de métodos estatísticos.

Por um lado, o que pode estar em causa aqui é a intuição como fundamento do edifício filosófico. A filosofia lida com conceitos. Mas que é que contribui para a discussão e fixação desses conceitos? Pode a intuição desempenhar o mesmo papel na filosofia que a percepção na ciência? Este é um problema que não é apenas da filosofia da experimental, mas de qualquer elaboração que recorra a intuições. O

que a filosofia experimental faz é tornar essas intuições mais explícitas. Mas pode haver uma segunda leitura que afecte somente a filosofia experimental: são as intuições “populares”, das pessoas comuns, relevantes para o exercício da filosofia? Será que as pessoas percebem realmente os problemas que estão em causa, isto é, são, por assim dizer, “falantes competentes”? Os filósofos experimentais tentam usar métodos para assegurar que as pessoas percebem convenientemente as questões e não são influenciadas por factores irrelevantes. Por exemplo, usam métodos estatísticos, e amostras suficientes para permitir uma leitura o mais fidedigna possível dos resultados; por vezes testam previamente se os participantes compreenderam os cenários, ou fazem os mesmos testes com diferentes cenários ou a diferentes populações. Mas mesmo que haja algum tipo de distorção proporcionado por falhas cognitivas ou outras, esse é precisamente um aspecto que merece estudo e atenção por parte dos filósofos. Os filósofos experimentais estão interessados em perceber como funciona a mente humana, e essas experiências podem ser importantes indicadores. Se de facto existirem essas diferenças na base das nossas intuições, essa é mais uma das razões pelas quais não podemos dar como garantidas certas intuições comuns, que servem depois de apoio para as análises conceptuais.

3.4.2 EXPERIMENTO: CONCEPÇÃO VULGAR DE ACRASIA

3.4.2.1 RAZÕES PARA UM EXPERIMENTO SOBRE A ACRASIA

A acrasia é um conceito que parece ter uma natureza intrinsecamente prática, isto é, a sua definição mais precisa provavelmente depende de observações que possamos fazer aos agentes reais. Quando falamos, por exemplo, da racionalidade, não no sentido estrito aqui empregado, parece que falamos de um conceito que tem uma natureza mais abstracta e que pode prescindir dos casos concretos – é como se o próprio conceito fosse uma régua para medir todos os casos concretos. No caso da acrasia, o que se procura é uma definição mais próxima possível de um fenómeno que parece ocorrer no agente real. Isto é, não é possível, como é no caso da racionalidade, produzir uma boa definição do fenómeno se não tentarmos encontrar a resposta no agente que declara ter vivido/experimentado a fraqueza da vontade. Parece-me pouco plausível dizer que uma pessoa experimentou a racionalidade, mas parece bastante defensável dizer que alguém experimentou a acrasia. A acção acrática, como já vimos, depende da percepção do agente, ou do observador, de que, de alguma forma, agiu contra o melhor juízo, tendo a possibilidade de realizá-lo.

A hipótese defendida neste trabalho é que a acrasia deve ser considerada um conceito moral e não um conceito ligado à racionalidade ou irracionalidade da acção. Como vimos nas primeiras partes deste capítulo, tratar da acrasia como um problema moral não é, em princípio, nenhuma novidade. A novidade parece estar

na forma de explicar a acção e especialmente a acção acrática, dada uma concepção funcionalista da racionalidade na acção. Além disto, a utilização de um experimento específico para testar a intuição de que a percepção das pessoas comuns de uma acção acrática parece depender do valor moral desta acção, é uma mais valia produzida pela inclusão da filosofia experimental no tratamento deste tema.

Comummente, em Português, dizemos com muito mais frequência coisas como: “S não tem força de vontade para fazer X”. Quando dizemos isso, estamos dizendo: “S não tem força de vontade para fazer X, mesmo sabendo que X é o melhor a fazer, mas, apesar disso, faz Y”. Podemos dizer, intuitivamente, que X dificilmente será algo moralmente negativo. Isto é, é difícil imaginar alguém a dizer: “S não tem força de vontade para consumir heroína”, ou “S não tem força de vontade para matar K”. Nos dois casos, no limite, poderíamos pensar em algo como: “S não tem coragem de injectar heroína”, ou “S não tem coragem de matar K”. No entanto, como já vimos, por exemplo, segundo uma concepção davidsoniana, não matar K pode ser caracterizado como “pouca” força de vontade, ou como fraqueza de vontade, caso o melhor juízo de S, todas as coisas consideradas, seja matar K.

Em Português é mais comum a utilização da expressão “força de vontade”, ao invés de “fraqueza da vontade”. No entanto, parece plausível admitir que a força de vontade é o inverso da fraqueza de vontade e, nesse sentido, podemos medir o mesmo fenómeno utilizando uma expressão mais coloquial/comum. Assim, de maneira vulgar, acrasia = fraqueza da vontade = pouca/fraca/falta de força de vontade.

Como vimos anteriormente, a acrasia tem sido definida como uma falha de racionalidade e não como um fenómeno moral. A acrasia seria “apenas” o agir contra o melhor juízo do próprio agente, todas as coisas consideradas, independentemente do conteúdo moral desse juízo e da acção realizada.

O experimento que foi realizado, procura mostrar como as pessoas definem ou percebem a fraqueza da vontade (acrasia). A hipótese teórica aqui defendida é de que as pessoas comuns associam a fraqueza da vontade a algum tipo de fraqueza moral, sem relacionar directamente a fraqueza da vontade com a não realização da acção que corresponde ao melhor juízo do agente, todas as coisas consideradas. Isto é, para considerar uma acção acrática, as pessoas comuns têm que perceber que o juízo do agente não é moralmente positivo (em qualquer sentido relevante para quem está em posição de determinar se uma acção foi, ou não acrática – seja o próprio agente, ou seja um observador) e que, portanto, a acção é moralmente negativa. O agente, nesse sentido, age contra um determinado juízo moral que ele próprio ou o observador considera como sendo o juízo que deveria levá-lo à acção. Assim, a acrasia parece ser um conceito ligado ao valor moral da acção e não à dissonância entre a decisão (melhor juízo TCC) e a acção, como defende a filosofia da acção de origem davidsoniana (ver item 3.1).

Portanto, o desafio desse experimento é mostrar que as pessoas comuns apenas imputam a qualificação de acrática à uma acção quando consideram que o valor moral desta é negativo, se comparado com aquilo que poderia o agente ter feito, considerando que o próprio agente teria declarado a sua decisão, no sentido moral positivo, antes da acção.

Isto é, o desafio é mostrar que as pessoas comuns consideram a acrasia apenas quando S diz que fará X (acção moralmente positiva) levando todos os aspectos em consideração e, no fim das contas, faz Y (acção moralmente negativa). Por outro lado, o maior desafio é mostrar que quando, segundo essa definição filosófica, S leva a cabo uma acção acrática na qual decidiu fazer Y (acção moralmente negativa) levando todos os aspectos em consideração e, no final das contas, faz X (acção moralmente positiva), as pessoas comuns não considerarão esta uma acção acrática – muito provavelmente vão achar que a acção se qualifica como sendo fruto da força de vontade de S.

Para tentar demonstrar a presença desse efeito moral no conceito popular de acrasia, o melhor desenho/estratégia de investigação mostrou-se ser a chamada “experiência simples” (*simple experiment*). Esse método foi também utilizado por Joshua Knobe em Side Effects (2003b).

O experimento simples é um dos caminhos mais fáceis e seguros, para se estabelecer as causas de um efeito. O efeito neste caso é aquele que pode ser definido como o efeito do valor moral sobre a percepção da acrasia (definida pela teoria davidsoniana - TD). Espera-se, portanto, definir que, dada uma *teoria D*, o conteúdo moral dos casos apresentados aos indivíduos provocará um determinado efeito sobre a percepção desses indivíduos acerca da acrasia (como é definida por TD). Por um lado, esse efeito será medido a partir de TD, mas, por outro, tentar-se-á avançar que a própria definição de acrasia (TD) está errada e que, portanto, o próprio efeito, que neste caso revela-nos o tal erro, poderá encaminhar-nos para uma melhor definição do fenómeno, desviando a caracterização da acrasia como um fenómeno de irracionalidade para uma caracterização como fenómeno moral.

Para realizar um experimento simples, é preciso ter dois grupos que não devem diferir em nenhum aspecto relevante e sistemático. Estes dois grupos serão tratados de maneira diferentes, isto é, receberão, neste experimento, casos diferentes para análise. Um dos grupos chama-se *grupo de controlo* e receberá um caso em que espera-se que os indivíduos respondam em concordância com TD. O outro grupo chama-se grupo experimental e receberá o caso em que espera-se que os indivíduos respondam em discordância com TD, pois o sentido moral deste caso é diferente do anterior.

É importante ressaltar que os dois casos não diferem em nada a partir da definição de acrasia pela TD. Apenas serão trocadas as posições dos factores nos casos entregues ao grupo de controlo e ao grupo experimental. Supostamente, não deveria haver diferença alguma entre as respostas dos dois grupos, mas tentaremos mostrar o tal efeito acima mencionado, que produzirá resultados diferentes no grupo experimental.

Em termos mais formais, para realizar-se um experimento simples é preciso, além de uma escolha aleatória de participantes, uma hipótese experimental, como, neste caso, a predição de que o conteúdo moral (valor negativo) dos casos apresentados aos indivíduos do grupo experimental provocará um efeito tal sobre nestes indivíduos que os resultados comparados entre os dois grupos serão diferentes e, provavelmente, inversos.

A hipótese nula deste experimento, que é a inversa da hipótese experimental, é neste caso aquela que prevê que não haverá diferenças entre o grupo de controlo e o grupo experimental, isto é, a hipótese nula prediz que o conteúdo moral não terá efeito algum sobre os resultados do experimento.

3.4.2.2 EXPERIMENTO, RESULTADOS, MÉTODOS

DESCRIÇÃO DA AMOSTRA

Foram seleccionados aleatoriamente quatro conjuntos de cem participantes adultos em diferentes ruas da cidade do Porto, de Vila Nova de Gaia e de Póvoa do Varzim, entre as 10h00 e as 21h00, entre os dias 20 e 24 de Julho de 2009. Os indivíduos foram convidados a responder duas perguntas, bem como a dar a sua identificação em termos de sexo e faixa etária. O inquérito era anónimo e a expectativa era de que cada entrevistado demorasse entre um minuto e meio e 2 minutos. No entanto, mesmo sem ter cronometrado as entrevistas (pois não era uma variável a ter em conta), vale a pena ressaltar que os indivíduos demoraram, pelo menos, o dobro do tempo esperado, algo em torno de 4 minutos. O tamanho total da amostra foi de 397 indivíduos: 179 (45%) mulheres e 218 (55%) homens; com faixa etária determinada pela seguinte distribuição: 228 (57%) indivíduos com idade inferior ou igual a 40 anos, 122 (31%) tinham entre 41 e 65 anos, e 47 (12%) eram indivíduos com idade superior a 65 anos.

Os entrevistadores abordaram os indivíduos dando as seguintes informações: a) tratava-se de uma investigação da Universidade do Porto; b) não havia resposta certa, ou errada – intencionava-se que as pessoas expressassem sem constrangimentos o quanto concordavam ou discordavam das afirmações; c) o entrevistador poderia responder a qualquer dúvida do entrevistado.

Dizer que se tratava de uma investigação da UP cumpriu os seguintes objectivos: (i) facilitar a abordagem aos indivíduos, por uma questão de segurança

(indivíduos); (ii) deixar o entrevistado mais tranquilo acerca da utilização da informação (apesar de ser um inquérito anónimo); (iii) desvincular esta actividade das acções de promoção comercial; (iv) propiciar maior seriedade/compromisso do indivíduo naquela actividade.

Ao dizer que não havia resposta certa, ou errada, objectivou-se neutralizar qualquer desvio produzido pela “tentativa de acertar”. Essa atitude poderia levar os indivíduos a duvidarem de suas próprias respostas, levando-os a assinalar a resposta menos intuitiva – algo como: “aqui deve ter algum ‘rato’, portanto, a resposta certa deve ser a resposta contrária a que as pessoas acham que é”. É possível que esse efeito fosse neutralizado pela própria pulverização deste efeito pela amostra, fazendo com que as distorções “se compensassem”. No entanto, optou-se por tentar minimizar este possível problema.

Dizer ao indivíduo que o entrevistador poderia responder a qualquer pergunta objectivou, de facto, ajudar o indivíduo, quando fosse o caso, a compreender os casos e mesmo resolver dúvidas operacionais, como, por exemplo, se deveria “fazer um X”, ou “fazer um círculo” (essas regras estavam no inquérito). Muitas vezes, questões desse género, se não forem resolvidas, podem atrapalhar a colecta dos dados e respondê-las não compromete a qualidade destes dados.

A distribuição de frequências associada ao sexo e à faixa etária está apresentada na Tabela 3.

Tabela 3: Distribuição de frequências associada ao sexo e à faixa etária, para o total da amostra.

Grupo Etário	Sexo		Total
	<i>Feminino</i>	<i>Masculino</i>	
<i>Até 40 anos</i>	93 (0.234)	136 (0.343)	229 (0.577)
<i>De 41 a 65</i>	62 (0.156)	60 (0.151)	122 (0.307)
<i>66 ou mais</i>	24 (0.060)	22 (0.055)	46 (0.116)
Total	179 (0.451)	218 (0.549)	397

DESCRIÇÃO DOS CASOS

O inquérito aplicado para cada conjunto de 100 pessoas continha dois casos: um deles era comum a todos os indivíduos da amostra (397) e outro era diferente para cada conjunto de 100 indivíduos (casos específicos). Em cada caso específico, os indivíduos foram convidados a expressar o seu nível de concordância a uma afirmação, variando de -3 (discordo totalmente) a 3 (concordo totalmente) – conforme a figura 11.

Figura 11: Escala de Concordância aplicada

Discordo Completamente			Neutro		Concordo Completamente	
-3	-2	-1	0	+1	+2	+3

O caso unificador (*Uni*) foi considerado a fim de verificar se todos os participantes estavam na mesma condição de partida. Os quatro casos que eram diferentes para cada grupo foram designados por "casos específicos" - C1.1, C1.2,

C2.1, C2.2, sendo: C1.1 e C2.2 casos aplicados aos grupos de controlo e C1.2 e C2.1 casos aplicados aos grupos experimentais. Os indivíduos responderam primeiro ao caso específico e depois ao caso *Uni*.

O experimento completo foi dividido em dois experimentos, A e B. No experimento A, utilizou-se a expressão *fraqueza da vontade*, tradução do termo inglês *weakness of will* que, por sua vez, é a expressão inglesa para *Akasia*. No experimento B, utilizou-se a expressão *força de vontade*. Em Português, utiliza-se muito pouco, ou quase nunca, a expressão *fraqueza da vontade*. Por isso, optou-se pela utilização desta expressão e também da expressão *força de vontade* que é mais comum/coloquial em Português. A utilização desta última expressão não traz, em princípio, nenhum problema conceptual, pois a ideia/sentido por de trás da expressão *força de vontade* é a mesma, mas na direcção oposta, da *fraqueza da vontade*.

O outro elemento da definição standard da acrasia, o melhor juízo (todas as coisas consideradas), está representado pela frase: “S decidiu, considerando todos os aspectos”, onde a decisão representa o melhor juízo, pois S considerou todos os aspectos. Se fossem utilizados os termos acrasia e melhor juízo, os indivíduos poderiam não compreender o caso, assim como a afirmação. A ideia central da acrasia está resguardada nos casos que foram aplicados.

Os casos eram os seguintes:

Experimento A

Experimento A, Grupo 1.1

Caso: C.1.1 (controlo)

João é um assassino profissional.

João decidiu, considerando todos os aspectos, que, no dia seguinte, iria procurar um emprego diferente, que não envolve violência.

No dia seguinte, João não procura o tal emprego diferente e mata mais uma pessoa por dinheiro.

Afirmção: A *fraqueza da vontade* é a ideia que melhor define a acção de João.

Experimento A, Grupo 1.2

Caso: C.1.2 (experimental)

João é um assassino profissional.

João decidiu, considerando todos os aspectos, que, no dia seguinte, iria matar mais uma pessoa por dinheiro.

No dia seguinte, João não mata mais uma pessoa e procura um emprego diferente, que não envolve violência.

Afirmção: A *fraqueza da vontade* é a ideia que melhor define a acção de João.

Experimento B

Experimento A, Grupo 2.1 **Caso: C.2.1 (experimental)**

João é um assassino profissional.

João decidiu, considerando todos os aspectos, que, no dia seguinte, iria matar mais uma pessoa por dinheiro.

No dia seguinte, João não mata mais uma pessoa e procura um emprego diferente, que não envolve violência.

Afirmção: *A força de vontade* é a ideia que melhor define a acção de João.

Experimento A, Grupo 2.2 **Caso: C.2.2 (controlo)**

João é um assassino profissional.

João decidiu, considerando todos os aspectos, que, no dia seguinte, iria procurar um emprego diferente, que não envolve violência.

No dia seguinte, João não procura o tal emprego diferente e mata mais uma pessoa por dinheiro.

Afirmção: *A força de vontade* é a ideia que melhor define a acção de João.

UNI: para EXPERIMENTOS A e B

Experimentos A e B**Caso Uni** (para C.1.1, C.1.2, C.2.1, C.2.2)

João é um homem saudável e sem problemas financeiros.

João decidiu, considerando todos os aspectos, que iria comer Pescada grelhada no almoço.

No almoço, João não come Pescada grelhada, mesmo estando disponível, mas come Badejo grelhado.

Afirmação: *A fraqueza da vontade* é a ideia que melhor define a acção de João.

Para simplicidade da apresentação, cada grupo será nomeado a partir de agora, como: C1.1, C1.2, C2.1 ou C2.2. Desse modo, o nome do grupo corresponde ao nome do caso.

Sexo e Faixa Etária

Os perfis de sexo e faixa etária obtidos para cada um dos quatro grupos amostrados estão apresentados na Tabela 4.

Tabela 4: Perfis de sexo e faixa etária, por grupo de amostragem.

Grupo Etário	Grupo							
	C1.1		C1.2		C2.1		C2.2	
	(n ₁ =100)		(n ₂ =98)		(n ₃ =99)		(n ₄ =100)	
	sexo		sexo		sexo		sexo	
	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.	Fem.	Masc.
até 40 anos	23 (23%)	34(34%)	19 (19%)	35(36%)	21 (21%)	40(41%)	30(30%)	27(27%)
de 41 a 65	14 (14%)	15(15%)	14 (14%)	20(21%)	18(18%)	11(11%)	16(16%)	14(14%)
66 ou mais	6 (6%)	8 (8%)	8 (8%)	2(2%)	4(4%)	5 (5%)	6(6%)	7(7%)

Para cada um dos quatro casos, C1.1 a C2.2, os efeitos do grupo etário e sexo sobre o grau de concordância foram avaliados pelo *modelo logit ordinal*. Para todas as análises estatísticas que se seguem, o *nível de significância* foi fixado em 0,05. Não foram identificados efeitos estatisticamente significativos para nenhuma das perguntas o que significa que as respostas obtidas não forneceram quaisquer evidências estatísticas para a existência de associações entre o grau de concordância, o sexo e faixa etária das amostras da população.

GRAUS DE CONCORDÂNCIA OBTIDOS NOS CASOS C1.1, C1.2, C2.1, C2.2

As frequências dos graus de concordância obtidas através do caso *Uni* e das questões específicas (C.1.1 a C.2.2) estão apresentadas em valores absolutos na tabela 5 e em valores percentuais na tabela 6.

Tabela 5: Frequências do grau de concordância, por grupo.

Caso	Grau de Concordância							Discorda	Concorda
	-3	-2	-1	0	+1	+2	+3	-3: -1	+1: +3
Uni	10	10	9	24	13	18	16	29	47
Q1.1	12	11	5	15	15	12	30	28	57
Uni	12	9	9	20	15	10	23	30	48
Q1.2	30	17	8	9	9	7	18	55	34
Uni	14	7	8	15	19	11	25	29	55
Q2.1	17	7	10	5	19	12	29	34	60
Uni	5	4	7	34	12	13	25	16	50
Q2.2	60	11	6	7	5	4	7	76	16

Tabela 6: Frequências (%) do grau de concordância, por grupo.

Caso	Grau de Concordância							Discorda	Concorda
	-3	-2	-1	0	1	2	3	-3:-1	+1:+3
Uni	10,00%	10,00%	9,00%	24,00%	13,00%	18,00%	16,00%	29,00%	47,00%
Q1.1	12,00%	11,00%	5,00%	15,00%	15,00%	12,00%	30,00%	28,00%	57,00%
Uni	12,24%	9,18%	9,18%	20,41%	15,31%	10,20%	23,47%	30,61%	48,98%
Q1.2	30,61%	17,35%	8,16%	9,18%	9,18%	7,14%	18,37%	56,12%	34,69%
Uni	14,14%	7,07%	8,08%	15,15%	19,19%	11,11%	25,25%	29,29%	55,56%
Q2.1	17,17%	7,07%	10,10%	5,05%	19,19%	12,12%	29,29%	34,34%	60,61%
Uni	5,00%	4,00%	7,00%	34,00%	12,00%	13,00%	25,00%	16,00%	50,00%
Q2.2	60,00%	11,00%	6,00%	7,00%	5,00%	4,00%	7,00%	76,00%	16,00%

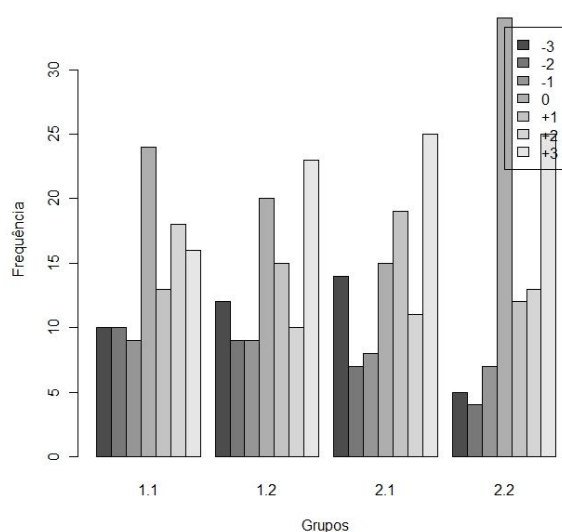
A homogeneidade da distribuição das respostas à questão Uni (entre os grupos) foi avaliada pelo teste de hipótese do qui-quadrado. As observações não mostraram evidência estatística para rejeitar a hipótese nula de homogeneidade das proporções ($\chi^2 = 23,1$, $df = 18$, $p = 0,187$), pelo que não é possível afirmar a existência de diferenças significativas entre as diferentes populações amostradas no que diz respeito à inclinação para a acrasia (veja Gráfico 1). Isso significa que o caso Uni não foi percebido pelos indivíduos dos quatro grupos de maneira

diferente, isto é, tanto os indivíduos dos grupos de controlo (A e B), como os indivíduos dos grupos experimentais (A e B) consideraram de maneira homogénea o caso Uni.

É possível perceber pelo Gráfico 1 que a distribuição das frequências dos valores das respostas no caso Uni é muito parecida nos quatro grupos.

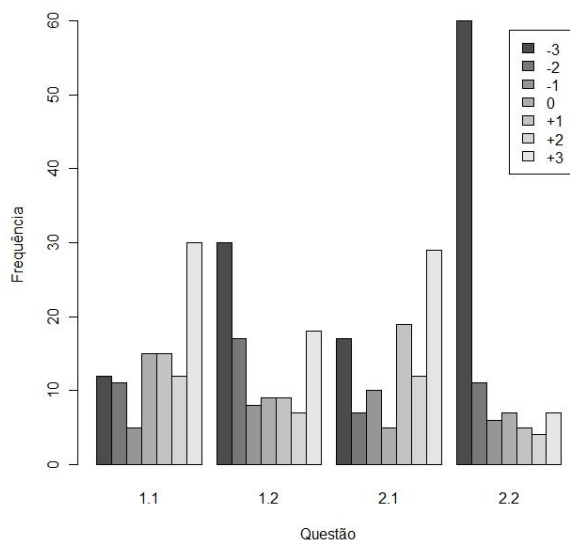
Há apenas duas situações que poderiam gerar alguma diferença, mas que a avaliação pelo teste do qui-quadrado não revelou: a quantidade elevada do valor zero (*neutro* – na escala de concordância) nos grupos 1.1 e 2.2 (veja Gráfico 1)

Gráfico 1: Frequência dos valores das respostas no caso Uni, por grupo (C)



Cabe ressaltar, que o caso Uni não tem, aparentemente, nenhum conteúdo moral. Em particular, este facto valida as conclusões inferidas a partir das distribuições das respostas às perguntas C1.1 à C2.2 que se seguem.

Os resultados obtidos para o grau de concordância de cada questão específica estão representados no gráfico 2.

Gráfico 2: Frequência dos valores dos graus de concordância, por caso específico (C)

A fim de compreender se havia, ou não, o já referido efeito para cada uma das traduções de "acrasia" utilizadas (*fraqueza da vontade* e *força de vontade*), compararam-se as distribuições dos graus de concordância correspondentes às questões C1.1 e C1.2, e C2.1 e C2.2. Os resultados do teste de hipótese do qui-quadrado foram estatisticamente significativos para C1.1 e C1.2 ($\chi^2 = 17$, $df = 6$, $p = 0,009$) e para C2.1 e C2.2 ($\chi^2 = 51,8$, $df = 6$, $p = 0,000$) o que permite concluir que, em cada caso, as duas populações amostradas não são homogêneas no que diz respeito às suas respostas (Gráficos 3 e 4, respectivamente).

Cabe ressaltar que, uma vez que o valor de p no segundo caso é menor, a evidência estatística para a conclusão é mais forte para o par C2.1-C2.2 do que para a primeira comparação.

Gráfico 3: Frequência dos valores do grau de concordância para os casos C1.1. e C1.2

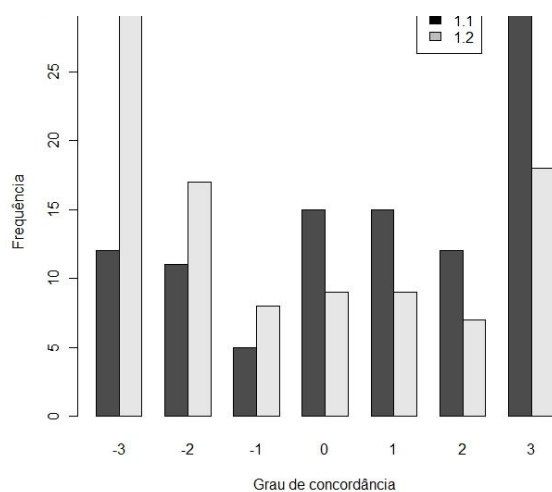
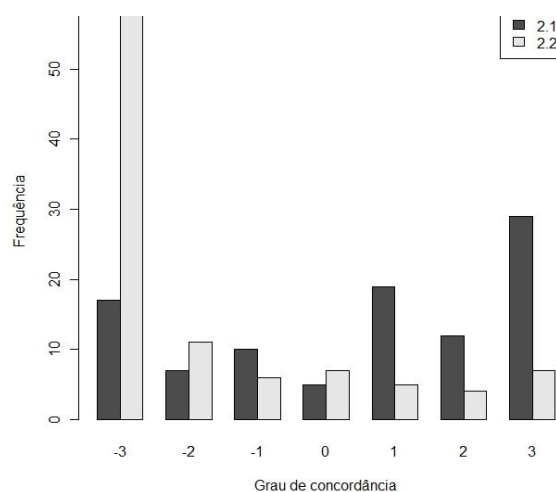


Gráfico 4: Frequências dos graus de concordância para os casos C2.1 e C2.2



Finalmente, investigamos as diferenças na distribuição dos graus de concordância entre as questões específicas, avaliando o mesmo conceito, mas usando a terminologia linguística diferente, ou seja, foram comparadas as respostas da questão específica C1.1 com as de C2.2, e as de C1.2 com C2.1 (veja gráficos 5 e 6). Como as duas traduções utilizadas eram antônimos, os valores da variável *grau de concordância* teve de ser invertida para um dos grupos envolvidos na comparação. Os resultados do teste de hipótese do qui-quadrado foram estatisticamente significativos para o primeiro par de perguntas ($\chi^2 = 21,4$, $df = 6$, $p = 0,002$) e estatisticamente não significativos para a segunda comparação ($\chi^2 = 6,6$, $df = 6$, $p = 0,362$). Em particular, a pergunta "Mata" é percebida de maneira diferente de acordo com as traduções utilizadas (fraqueza, ou força), mas não acontece o mesmo com a pergunta "Procura".

Gráfico 5 : Frequência de valores do grau de concordância para as questões C1.1. e do simétrico de C2.2.

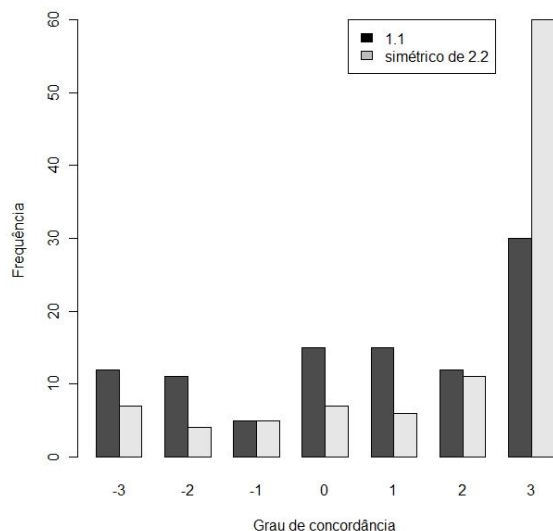
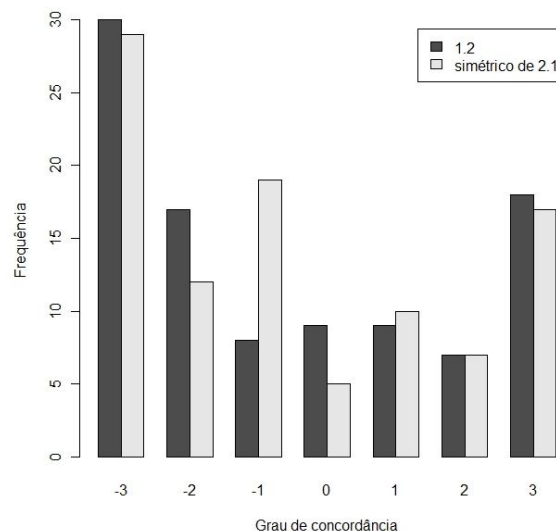


Gráfico 6: Frequência de valores do acordo nível para as questões C1.2. e do simétrico de C2.1



INTERPRETAÇÃO DOS RESULTADOS

GRUPO UNI

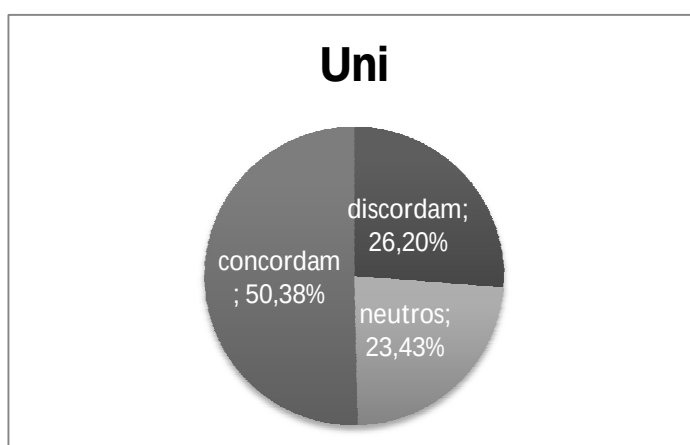
O caso Uni, como dito anteriormente, foi aplicado aos 397 indivíduos e serviu para verificar se havia alguma diferença substancial entre os 4 grupos que pudesse comprometer o resultado do experimento. Felizmente, não houve diferenças relevantes e a população amostrada caracterizou-se por ser homogênea.

Este grupo também nos mostra algo acerca da percepção das pessoas sobre a acrasia. Neste caso, as pessoas tinham que dizer se concordavam, ou não, que o facto de S ter decidido comer X, mas ter comido Y, poderia ser considerado

fraqueza da vontade. No gráfico 7, podemos ver que 50,38% da população amostrada concordou com a afirmação, portanto S seria acrático para metade das pessoas. Porém, os outros quase 50% ou discordaram, ou revelaram-se neutros.

Era esperado um número menor de concordantes e um número maior de neutros. Uma interpretação possível deste número inesperado de concordantes é que as pessoas provavelmente consideram o facto de uma pessoa ter decidido fazer algo e não fazer, sem que nenhum elemento moral esteja em jogo directamente, uma fraqueza moral. Isto é, para um número expressivo das pessoas não agir de acordo com a decisão é sempre um sinal de fraqueza moral, desde que o caso não tenha conteúdo moral. Se tiver o referido conteúdo moral, as pessoas vão responder de acordo com a hipótese levantada neste experimento e também de acordo com os resultados obtidos à frente, que prevê que o conteúdo moral provoca um efeito sobre a percepção das pessoas acerca do não cumprimento de uma decisão. Esta análise só é possível por causa, principalmente, dos resultados comparados de C2.1 e C2.2.

Gráfico 7: Distribuição da Frequência do Grau de Concordância do Grupo *Uni* (%)



EXPERIMENTO A

COMPARAÇÃO ENTRE O GRUPO DE CONTROLO E O GRUPO EXPERIMENTAL (C1.1 E C1.2)

C1.1 é o grupo de controlo do experimento A. Isto é, aqui era esperado um resultado em conformidade com TD. Nesse caso, João decide (segundo o seu melhor juízo) fazer X (procurar um emprego), mas faz Y (mata uma pessoa). Segundo a teoria em questão, uma acção desse tipo deveria ser considerada acrática. Esperava-se que as pessoas comuns indicassem exactamente isso. No entanto, a hipótese é que o fariam, não por concordância estrita à teoria e sim porque “matar uma pessoa” tem um valor moral negativo e “procurar um emprego” tem um valor moral positivo. Assim, as pessoas comuns diriam que fazer Y seria uma acção acrática por causa do seu valor moral negativo.

Por causa disso, seria necessário, então, fazer com que a opinião moral-avaliativa das pessoas comuns não coincidisse com a prescrição da teoria (amoral-avaliativa). Nesse sentido, foi aplicada uma outra questão, chamada questão experimental, a um outro grupo de 100 pessoas, a fim de testar essa intuição.

Os resultados desse grupo confirmaram essas expectativas. O resultado específico está representado no gráfico 8.

Gráfico 8: Distribuição da Frequência do Grau de Concordância do Grupo C1.1 (%)

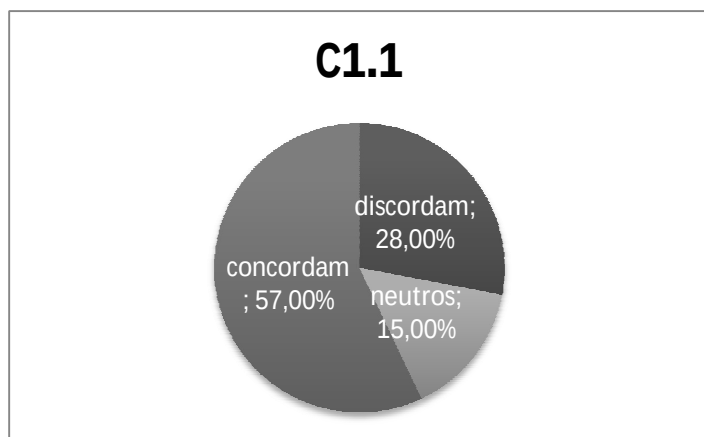
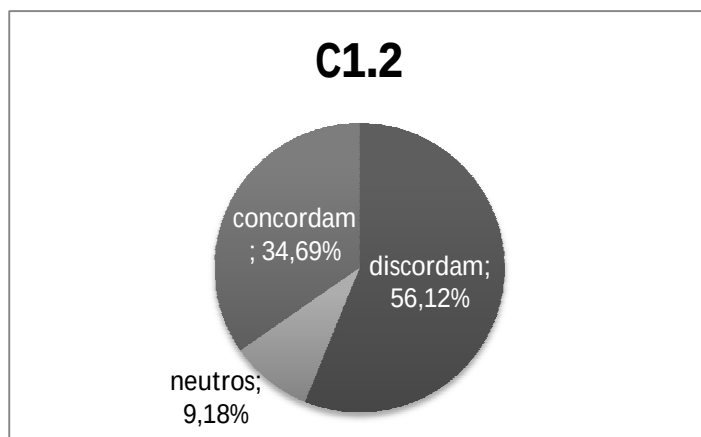


Gráfico 9: Distribuição da Frequência do Grau de Concordância do Grupo C1.2 (%)



Como é possível verificar, 57% da população amostrada consideraram acrasia a acção de João – matar uma pessoa – e 28% não consideraram acrasia. Por outro lado, com a mesma mecânica, mas com a alteração do valor moral da acção, o resultado foi o inverso. Como se pode verificar no gráfico 9, 56,12% da população amostrada discordaram que a acrasia caracterizava a acção de procurar um emprego, quando a decisão tinha sido matar mais uma pessoa por dinheiro, enquanto 34,69 concordaram. Aqui, verifica-se claramente a inversão na percepção das pessoas acerca do fenómeno quando o valor moral é invertido, mesmo que a mecânica do caso continue a mesma (caracterizando sempre uma acção acrática segundo a TD).

EXPERIMENTO B

COMPARAÇÃO ENTRE DOS GRUPO DE CONTROLO E O GRUPO EXPERIMENTAL (C2.1 E C2.2)

O experimento B foi aquele que se mostrou mais expressivo. Neste experimento, troquei a expressão fraqueza da vontade, por força de vontade, para tentar uma abordagem ainda mais próxima da linguagem comum. Fazendo isto, o caso de controlo passou a ser 2.2 e o caso experimental passou a ser o 2.1.

No caso 2.1, João decide matar mais uma pessoa por dinheiro, mas acaba por procurar um emprego – acção acrática segundo TD.

Mesmo sendo uma acção acrática, 60,61% da população amostrada, como podemos ver no gráfico 10, afirmou tratar-se de uma acção caracterizada pela força de vontade. Procurar um emprego tem um valor moral positivo e produziu uma distorção relevante na percepção das pessoas acerca da acção acrática. As pessoas parecem não enxergar nesse tipo de acção aquilo que alguns filósofos afirmam existir – a acrasia.

No caso 2.2, o resultado foi também expressivo. As pessoas foram convidadas a dizer se concordavam que João agiu com força de vontade ao matar mais uma pessoa quando havia decidido procurar um emprego. Como mostra o gráfico 11, 77% das pessoas discordaram desta afirmação. Isto significa que, dado o valor moral negativo de matar mais uma pessoa, as pessoas não concordaram que a acção caracterizava-se pela força de vontade.

Mesmo que pensássemos em afirmar que no caso 2.2 as pessoas concordaram com TD, teríamos que levar em conta que em 2.1 as pessoas fizeram o oposto, dado o valor moral positivo da acção.

Gráfico 10: Distribuição da Frequência do Grau de Concordância do Grupo C2.1 (%)

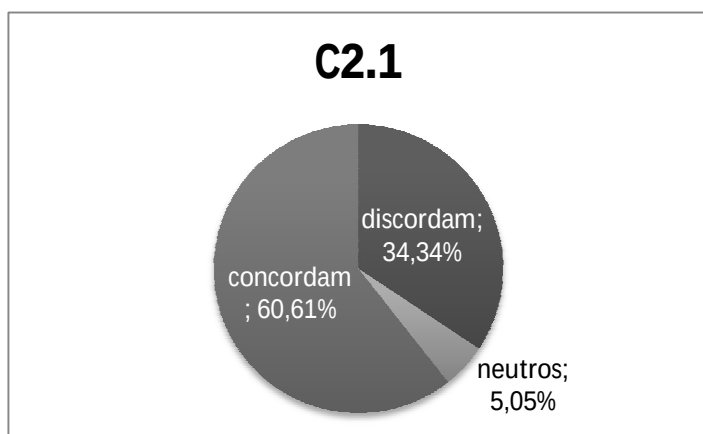
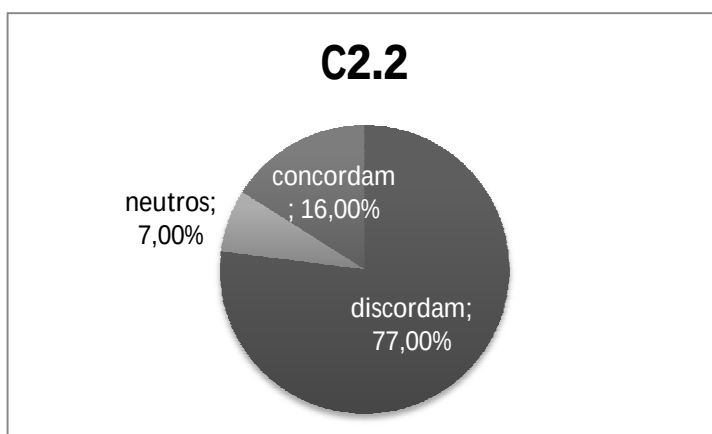


Gráfico 11: Distribuição da Frequência do Grau de Concordância do Grupo C2.2 (%)



3.4.2.3 AVALIAÇÃO FINAL DO EXPERIMENTO

A análise filosófica dos resultados deste experimento é bastante simples, dado o que foi escrito neste capítulo, até aqui. A minha hipótese era que a acrasia é uma questão moral e não um problema de racionalidade. Creio que os resultados da experiência confirmam a intuição de que as pessoas associam o valor moral da acção à percepção da fraqueza da vontade. Isto é, a imputação de acrasia depende do valor moral da acção.

No item 1.1 e ao longo de todo o capítulo 3, argumentei em favor da importância da psicologia popular, da filosofia experimental (moderada) e da importância da concepção popular da acrasia para o trabalho do filósofo que se debruça sobre o tema. Levando isto em consideração, posso afirmar que, dentro daquilo a que me propus fazer, o resultado deste experimento confirma a minha hipótese central sobre a acrasia.

Alguém poderá dizer que este resultado não produz nenhum efeito sobre as teorias filosóficas da acrasia e que apenas mostra que as pessoas não sabem o que é a acrasia.

Esta linha de pensamento é bastante curiosa. Parece que o conceito de acrasia neste caso é encarado como algo que não depende da percepção das pessoas acerca do fenómeno. Quem pensa que o “erro” das pessoas se deve à ignorância acerca do conceito, pensa que o conceito de acrasia é como dizer que a raiz quadrada de 9 é 3. Isto é, o conceito está dado e o que se pode fazer é medir o conhecimento das pessoas sobre ele. Mas o caso da acrasia é diferente. A definição do conceito depende da ocorrência de algo no agente real que possa ser caracterizada como acrasia. Se isto é verdade, é necessário tentarmos investigar como e quando o agente real percebe a ocorrência deste fenómeno. Um experimento directo é muito difícil, isto é, tentar monitorizar um agente e tentar isolar uma ocorrência e descobrir o que se passou, é um experimento que não parece ser operacionalmente possível. A alternativa é tentar perceber como e em quais circunstâncias as pessoas comuns caracterizam uma acção como acrática. Foi exactamente isto que se fez.

Outros poderiam dizer que a acrasia é especificamente o fenómeno definido pela TD e aquilo que o experimento captou é algum outro fenómeno que nenhuma relação tem com a acrasia. Isto seria possível se não levássemos em consideração o uso que as pessoas fazem da linguagem comum quando se referem a fenómenos como a acrasia, ou a estados mentais como crenças e desejos. Se estamos dispostos a investigar este uso da linguagem, temos também que estar dispostos a levar a sério as investigações acerca dele.

OBSERVAÇÕES FINAIS

Aqui, pretendo, objectiva e directamente, recapitular as questões centrais, mostrar como pretendi respondê-las ao longo da tese e, principalmente, dar a conhecer a minha proposta para uma agenda futura de investigação. Não me parece adequado incluir uma secção denominada conclusão, tendo em vista a natureza deste trabalho. Uma concepção deflacionista da racionalidade na acção, como a que aqui foi exposta, assim como o modelo dela derivado, precisam ser testados em outros casos, além do caso da acrasia aqui trabalhado. Por essa razão, esta investigação exploratória deverá ser continuada com o objectivo de dar maior robustez teórica e metodológica ao referido modelo.

Um dos objectivos deste trabalho era propor um modelo funcionalista da racionalidade na acção. Isto foi feito no capítulo 2. No entanto, foi necessário, para lá chegar, passar pelos elementos que constituem as bases deste modelo. Todo o Capítulo 1 foi dedicado a esses fundamentos. Foram tratados os seguintes temas: a *psicologia popular*, o *realismo intencional*, a *racionalidade na acção* na economia e na filosofia da acção. Por um lado, o modelo proposto tem os alicerces no realismo intencional e na psicologia popular, por outro lado, a relevância deste modelo precisa ser testada na filosofia da acção e a sua aplicabilidade testada na teoria económica. O capítulo 3 pretendeu ser um teste deste modelo na filosofia da acção. O teste e adaptação do modelo na teoria económica não eram objectivos deste trabalho, razão pela qual não foram realizados. No entanto, pretende-se cumprir este objectivo na continuação desta investigação.

O diálogo entre a filosofia e a economia é muito fértil. Nesta tese, tentou-se aproximar a teoria da acção da filosofia da teoria da escolha racional – este trabalho precisa ser prosseguido, mas, pelo menos parece ter ficado claro que são duas abordagens que devem ser integradas. A teoria da escolha racional parece responder bem a questões colectivas, no entanto, a teoria da acção da filosofia parece ser mais realista em termos psicológicos individuais. Existem pontos de contacto entre as duas abordagens, além da psicologia popular, e deverão ser explorados a partir daqui.

Um outro tema que foi tratado nesta tese e que me parece importante é a filosofia experimental. Esta abordagem filosófica tem produzido um dos mais intensos debates já vividos na filosofia de inclinação analítica. Nos últimos 10 anos, muito se tem feito a partir desta abordagem directamente experimental, tendo como base a ideia, não tão recente, de uma aproximação entre a filosofia e as ciências. No entanto, nos últimos cinco anos, estes debates intensificaram-se de tal forma que nas discussões metafilosóficas parece impossível um filósofo esquivar-se de ter uma opinião, mesmo que provisória, acerca da filosofia experimental.

Por fim, proponho uma agenda mínima de investigação a partir desta concepção deflacionista da racionalidade na acção:

Desenvolver um conjunto de: a) aplicações práticas, na área das políticas públicas, b) propostas teóricas formuladas no âmbito da filosofia da acção e da filosofia da economia em torno de acção, agência e intencionalidade.

Avançar uma concepção filosófica e económica da explicação da decisão e da acção que seja ao mesmo tempo psicologicamente viável.

Produzir um modelo aplicado às políticas públicas desta concepção filosófica da explicação da decisão e da acção.

Realizar um conjunto sistemático de experimentos acerca da acção intencional, da racionalidade na acção, das explicações da acção, de forma a poder vir a aplicar os resultados às Políticas Públicas.

As bases teóricas centrais desta agenda seriam: (a) a filosofia da acção, (b) a filosofia da economia, (c) a ciência cognitiva, (d) a filosofia política. A base teórico-aplicada seria formada (e) pelas abordagens filosóficas das políticas públicas. As bases metodológicas seriam: (i) método filosófico analítico (não estritamente ligado à filosofia da linguagem), (ii) filosofia experimental (moderada, que aceita as intuições como evidência), (iii) métodos aplicados à ciência cognitiva, (iv) análise filosófico-económica (debates acerca da teoria da escolha racional). A metodologia utilizada na aplicação dos resultados esperados seria constituída por um conjunto de estratégias de investigação que formam a área das políticas públicas, associado ao método filosófico.

Seguindo a definição canónica de Hausman, a filosofia da economia consiste, resumidamente, em: investigações acerca da racionalidade humana (debates acerca da teoria da escolha racional); análise filosófica dos resultados económicos, institucionais e também acerca dos processos económico-sociais; e investigações acerca da ontologia dos fenómenos económicos e acerca das questões epistemológicas ligadas à ciência económica. É possível dizer que as questões da filosofia da economia se enquadram basicamente em três áreas: teoria da acção, ética, e filosofia da ciência. A área das políticas públicas pode ser resumidamente caracterizada como consistindo em estudos analíticos acerca dos possíveis cursos

de acção das entidades governamentais e acerca dos processos de decisão governamental. Isto envolve, no mínimo, estudos acerca da acção individual, da acção colectiva, da racionalidade na acção e da (im)possibilidade de agregação racional das preferências individuais. É uma área constituída primordialmente pelas ciências sociais, especialmente a economia, a política e a sociologia.

Esta agenda propõe-se, num primeiro momento, explicitar as bases comuns da filosofia da acção, filosofia da economia, ciências cognitivas e filosofia política, defendendo que elas se erguem em torno da acção e da racionalidade na acção. Num segundo momento procurar-se-á discutir questões aplicadas, mais ligadas à política, à economia e às políticas públicas. Tenho em mente, mais especificamente, questões com extrema repercussão social e com estreita relação com as explicações da acção e da racionalidade na acção, tais como, por exemplo o combate à corrupção, o problema do *free rider*, a justiça distributiva, a justiça social, a relação entre identidade e violência, os debates acerca do cosmopolitismo, além de questões mais específicas que podem ser definidas contextualmente.

BIBLIOGRAFIA

Com esta secção o meu objectivo foi seleccionar um conjunto grande de publicações que possa servir como guia para outros investigadores que trabalhem, ou que queiram trabalhar, acerca dos temas desta tese. Esta secção inclui as referências bibliográficas e uma bibliografia geral da área.

Adams, F. (1986). Intention and intentional action: The simple view. *Mind and Language* 1: 281-301.

Adams, F., & Steadman, A. (2004). Intentional action in ordinary language: Core concept or pragmatic understanding? *Analysis*, 64(2), 173–181.

Adams, Fred & Steadman, Annie (2004). Intentional action and moral considerations: Still pragmatic. *Analysis* 64 (3):268–276.

Adler, J. E. (2002). Akratic believing? *Philosophical Studies*, 110(1).

Ainslie, G. (2005). Précis of breakdown of will. *Behavioral and Brain Sciences*, 28(5), 635-650.

Alexander, Joshua & Weinberg, Jonathan M. (2007). Analytic epistemology and experimental philosophy. *Philosophy Compass* 2 (1):56–80.

Alston, W. P. (1986). An action-plan interpretation of purposive explanations of actions. *Theory and Decision*, 20: 275-299.

Alt, J. E., & Gilligan, M. (1994). The political economy of trading states: Factor specificity, collective action problems and domestic political institutions. *Journal of Political Philosophy*, 2(2), 165–192.

Andrews, Kristin (2003). Knowing mental states: The asymmetry of psychological prediction and explanation. In Quentin Smith & Aleksandar Jokic (eds.), *Consciousness: New Philosophical Perspectives*. Oxford University Press.

Anscombe, Elizabeth, (2000). *Intention* (reprint), Cambridge, MA: Harvard University Press.

Antony, L. M. (1987). Attributions of intentional action. *Philosophical Studies*, 51(3).

Antunes, C. H., Dias, L. C. (2007). *Decisão: Perspectivas Interdisciplinares*. Coimbra: Imprensa da Universidade de Coimbra.

Apostel, L. (1956). The formal structure of action. *Synthese*, 10(1).

Appiah, A. (2008). *Experiments in Ethics*: Harvard University Press.

- Aristóteles (2004). *Ética a Nicómaco*. Tradução de António C. Caeiro. Lisboa: Quetzal Editores.
- Aristóteles (2006). *De Anima*. Tradução de Maria Cecília Gomes dos Reis. São Paulo: Ed.34.
- Arrow, K. et al. (eds.) (1999). *The rational foundations of economic behavior*. Londres: Palgrave Macmillan.
- Audi, R. (1973). Intending. *Journal of Philosophy* 70: 387-402.
- Audi, R. (1979). Weakness of will and practical judgment. *Nous*, 13(2), 173-196.
- Audi, R. (1986). Intending, intentional action, and desire. In *The Ways of Desire*, ed. j. Marks. Precedent.
- Audi, R. (1988). Deliberative intentions and willingness to act: A reply to Professor Mele. *Philosophia* 18: 243-245.
- Audi, R. (1990). Weakness of will and rational action. *Australasian Journal of Philosophy*, 68(3), 270 – 281.
- Audi, R. (1991). Intention, cognitive commitment, and planning. *Synthese* 86: 361-378.
- Audi, R. (1993). *Action, Intention, and Reason*. Ithaca, NY: Cornell University Press.
- Audi, R. (2001). *The Architecture of Reason: The Structure and Substance of Rationality*. Oxford University Press.
- Audi, R. (2008). Intuition, inference, and rational disagreement in ethics. *Ethical Theory and Moral Practice*, 11(5).
- Bach, K. (1978) A Representational Theory of Action. *Philosophical Studies*, 34, 361-379.
- Baert, P. (1997). Algumas Limitações das Explicações da Escolha Racional na Ciência Política e na Sociologia. *RBCS*, 12(35), 63-74.
- Baier, A. C. (1970). Act and intent. *Journal of Philosophy* 19: 648-658.
- Baier, A. C. (1993). Review of *Springs of Action: Understanding Intentional Behavior*, by Alfred Mele. *Mind*, 102, 408: 668.
- Baker, Lynne Rudder (1999). Folk psychology. In Rob Wilson & Frank Keil (eds.), *MIT Encyclopedia of Cognitive Science*. MIT Press.
- Baker, Lynne Rudder (1999). What is this thing called 'commonsense psychology'? *Philosophical Explorations* 2 (1):3-19.
- Bardsley, N. (2007). On collective intentions: Collective action in economics and philosophy. *Synthese*, 157(2).
- Bartsch, Karen & Wright, Jennifer Cole (2005). Towards an intuitionist account of moral development. *Behavioral and Brain Sciences* 28 (4):546-547.
- Baum, W. M. (2006). *Compreendendo o Behaviorismo: comportamento, cultura e evolução*. Porto Alegre: Artmed.
- Bealer, G. (1996). A priori knowledge and the scope of philosophy. *Philosophical Studies*, 81(2-3), 121-142.
- Bealer, G. (1996). A priori knowledge: Replies to William Lycan and Ernest Sosa. *Philosophical Studies*, 81(2-3), 163-174.
- Bealer, G. (1996). On the possibility of philosophical knowledge. *Philosophical Perspectives*, 10, 1-34.

- Bealer, G. (2000). A theory of the a priori. *Pacific Philosophical Quarterly*, 81(1), 1–30.
- Beardman, S. (2007). The special status of instrumental reasons. *Philosophical Studies*, 134(2).
- Beardsley, M. (1980). Motives and Intentions. In *Action and Responsibility*, ed. M. Bradie and M. Brand. Applied Philosophy Program, Bowling Green State University.
- Becker, G. (1976). *The Economic Approach to Human Behavior*. Chicago: University of Chicago Press.
- Bedke, M. S. (2008). Practical reasons, practical rationality, practical wisdom. *Ethical Theory and Moral Practice*, 11(1).
- Bedke, Matthew S. (2008). Practical reasons, practical rationality, practical wisdom. *Ethical Theory and Moral Practice* 11 (1).
- Baker, Lynne Rudder (1999). Folk Psychology. In Rob Wilson & Frank Keil (eds.), *MIT Encyclopedia of Cognitive Science*. MIT Press.
- Bennett, D. (1965). Action, Reason and Purpose. *Journal of Philosophy*, 62: 85-95.
- Bennett, J. (1991). Folk-psychological explanations. In John D. Greenwood (ed.), *The Future of Folk Psychology*. Cambridge University Press.
- Bennett, J. (1995). *The Act Itself*. Oxford: Clarendon Press
- Bermudez, Jose Luis (2002). Rationality and psychological explanation without language. In Jose Luis Bermudez & Alan Millar (eds.), *Reason and Nature*. Clarendon.
- Bermudez, Jose Luis (2003). The domain of folk psychology. In Anthony O'Hear (ed.), *Minds and Persons*. Cambridge University Press.
- Bicchieri, Cristina (2005) *The Grammar of Society: The Nature and Dynamics of Social Norms*. Cambridge: Cambridge University Press.
- Bigelow, J., Dodds, S. M., & Pargetter, R. (1990). Temptation and the Will. *American Philosophical Quarterly*, 27(1), 39-49.
- Bishop, John (1990). *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge University Press.
- Bittner, R. (2001). *Doing Things for Reasons*: Oxford University Press.
- Blaug, M. (1993). *Metodologia da Economia, ou, como os Economistas Explicam*. Tradução de Afonso Medeiros dos Santos Lima - 2. ed. rev. - São Paulo: Edusp.
- Block, Ned & Fodor, Jerry (1972). What psychological states are not. *Philosophical Review* 81 (April):159-81.
- Block, Ned (1971). Are mechanistic and teleological explanations of behaviour incompatible? *Philosophical Quarterly* 21 (April):109-117.
- Blumenfeld, Jean Beer (1976). Goldman's account of intentional action. *Philosophical Studies* 29 (6).
- Boag, S. (2006). Can repression become a conscious process? *Behavioral and Brain Sciences*, 29(5), 513-514.
- Bogdan, Radu J. (1991). *Mind and Common Sense: Philosophical Essays on Commonsense Psychology*. Cambridge University Press.

- Boland, Lawrence A. (1983). On the best strategy for doing philosophy of economics. *British Journal for the Philosophy of Science* 34 (4).
- Boudon, R. (1982). *The Unintended Consequences of Social Actions*. London: The Macmillan Press.
- Boulter, S. J. (2007). The “evolutionary argument” and the metaphilosophy of commonsense. *Biology and Philosophy*, 22(3).
- Bouvier, A. (2002). An epistemological plea for methodological individualism and rational choice theory in cognitive rhetoric. *Philosophy of the Social Sciences*, 32(1).
- Brand, M. (1984). *Intending and acting*. Cambridge, MA.: MIT Press.
- Brand, M. (1989). Proximate Causation of Action. *Philosophical Perspectives*, Vol. Issue Philosophy of Mind and Action Theory, 423-442.
- Bratman, M. (1979). Practical reasoning and weakness of the will. *Nous*, 13(2), 153-171.
- Bratman, M. (1984). Two faces of intention. *Philosophical Review*, 93(3), 375-405.
- Bratman, M. E. (1987). *Intention, Plans, and Practical Reason*. Cambridge, MA: Cambridge University Press.
- Bratman, M. E. (1989). Replies to McCann and Velleman. Manuscript, Stanford
- Bratman, M. E. (1992). Shared cooperative activity. *Philosophical Review*, 101(2), 327-341.
- Bratman, M. E. (1993). Shared intention. *Ethics* 104: 97-113.
- Bratman, M. E. (1997). Responsibility and planning. *Journal of Ethics* 1: 27-43. Bratman, M. E. 1999. *Faces of Intention: Selected Essays on Intention and Agency*. Cambridge University Press.
- Bratman, M. E. (1999). *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge: Cambridge University Press.
- Bratman, M. E. (2006). *Structures of Agency*, Oxford: Oxford University Press.
- Bratman, M. E. (2009). Setiya on intention, rationality and reasons. *Analysis*, 69(3).
- Bratman, M. E. (2009). Intention, practical rationality, and self-governance. *Ethics* 119 (3).
- Breheny, Richard (2006). Communication and folk psychology. *Mind and Language* 21 (1):74-107.
- Brendel, Elke (2004). Intuition pumps and the proper use of thought experiments. *Dialectica* 58 (1):89–108.
- Brewer, B. (1992). Self-Location and Agency. *Mind*, 101: 17-34.
- Brewer, B. (1993). The integration of spatial vision and action. In Eilan, N., McCarthy, R. & B. Brewer, *Spatial Representation*, Oxford: Blackwell, pp. 294-316.
- Bridges, J. Rationality, normativity and transparency.
- Broome, John (1987). Utilitarianism and expected utility. *Journal of Philosophy* 84 (8):405-422.
- Broome, John (1991). Desire, belief and expectation. *Mind* 100 (2):265-267.
- Broome, John (1994). Discounting the future. *Philosophy and Public Affairs* 23 (2):128–156.

- Broome, John (1997). Reasons and motivation. *Aristotelian Society Supplementary Volume* 71 (1):131–146.
- Broome, John (1999). Normative requirements. *Ratio* 12 (4):398–419.
- Broome, John (2001). Normative practical reasoning: Broome, John. *Aristotelian Society Supplementary Volume* 75 (1):175–193.
- Broome, John (2004). *Weighing Lives*. Oxford: Oxford University Press.
- Broome, John (2005). Does rationality give us reasons? *Philosophical Issues* 15 (1):321–337.
- Broome, J. (2007a). Tomar uma decisão através de raciocínio. In Antunes, C. H., Dias, L. C. *Decisão: Perspectivas Interdisciplinares*. Coimbra: Imprensa da Universidade de Coimbra.
- Broome, John (2007b). Does rationality consist in responding correctly to reasons? *Journal of Moral Philosophy* 4 (3).
- Broome, John (2007c). Wide or narrow scope? *Mind* 116 (462).
- Broome, John (2008). Reply to Southwood, Kearns and star, and Cullity. *Ethics*, 119(1).
- Broome, John (2009). Motivation. *Theoria* 75 (2):79–99.
- Brown, Jessica (2008). Knowledge and practical reason. *Philosophy Compass* 3 (6):1135–1152.
- Brunero, John (2007). Are intentions reasons? *Pacific Philosophical Quarterly* 88 (4):424–444.
- Brunero, John (forthcoming). The scope of rational requirements. *Philosophical Quarterly*.
- Buss, S. (1997). Weakness of will. *Pacific Philosophical Quarterly*, 78(1), 13–44.
- Caillé, A. (1998). Nem Holismo nem Individualismo Metodológicos - Marcel Mauss e o paradigma da dádiva. *RBCS*, 13(38), 5–37.
- Carrier, Leonard S. (1986). Free will and intentional action. *Philosophia* 16 (December):355–364.
- Carruthers, Peter & Botterill, George (1999). *Philosophy of Psychology*. Cambridge: Cambridge University Press.
- Carruthers, Peter & Smith, Peter K. (1996). *Theories of Theories of Mind*. Cambridge University Press.
- Carruthers, Peter (1996). Simulation and self-knowledge: A defence of the theory-theory. In Carruthers, Peter & Smith, Peter K. (eds.), *Theories of Theories of Mind*. Cambridge University Press.
- Castañeda, H.-N. (1972). Intentions and intending. *American Philosophical Quarterly* 9: 139–149.
- Castañeda, H.-N. (1975). *Thinking and doing*. Dordrecht, Holland: D. Reidel.
- Castañeda, H.-N. (1982). Conditional intentions, intentional action and aristotelian practical syllogisms. *Erkenntnis*, 18(2).
- Castaneda, H.-N. (1990). Practical thinking, reasons for doing, and intentional action: The thinking of doing and the doing of thinking. *Philosophical Perspectives*, 4, 273–308.
- Castañeda, H.-N. (1992). Indexical reference and bodily causal diagrams in intentional action. *Studia Logica*, 51(3–4).

- Chandler, J. (1973). Act-utilitarianism and collective action. *Ethics*, 84(1), 78-85.
- Chant, S. R. (2007). Unintentional collective action. *Philosophical Explorations*, 10(3), 245 – 256.
- Chant, S. R., & Ernst, Z. (2008). Epistemic conditions for collective action. *Mind*, 117(467).
- Charlton, W. (1988). *Weakness of Will*: B. Blackwell.
- Cherniak, C. (1986). *Minimal Rationality*. Cambridge: MIT Press.
- Christiano, T. (2004). Is normative rational choice theory self-defeating? *Ethics*, 115(1).
- Churchland, Paul M. (1988). Folk psychology and the explanation of human behavior. *Proceedings of the Aristotelian Society* 62:209-21.
- Clarke, Randolph (2008). Autonomous reasons for intending. *Australasian Journal of Philosophy* 86 (2):191 – 212.
- Cohen, J.L. (1981). Can Human Irrationality Be Experimentally Demonstrated?. *Behavior and Brain Sciences*, 4: 317-370.
- Coleman, M. C. (2008). Directions of fit and the Humean theory of motivation. *Australasian Journal of Philosophy*, 86(1), 127 – 139.
- Collins, H. and Kusch, M. (1998). *The Shape of Actions*. Cambridge, MA: MIT Press.
- Cordner, C. (1985). Jackson on weakness of will. *Mind*, 94(374), 273-280.
- Cullity, Garrett (2008). Decisions, reasons, and rationality. *Ethics* 119 (1).
- Cummins, R. (1998). Reflection on reflective equilibrium. In W. Ramsey & M. R. DePaul (Eds.), *Rethinking Intuition: The Psychology of Intuition and Its Role in Philosophical Inquiry*: Rowman and Littlefield Publishers.
- Dahl, N. O. (1984). *Practical Reason, Aristotle, and Weakness of Will*. Minneapolis: University of Minnesota Press.
- Damper, R. I. (2006). The logic of Searle's chinese room argument. *Minds and Machines*, 16(2), 163-183.
- Danto, A. (1973). *Analytical philosophy of action*. Cambridge: Cambridge University Press.
- Darwall, S. (2003). Desires, reasons, and causes. *Philosophy and Phenomenological Research*, 67(2), 436–443.
- Darwiche, A., & Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence*, 89, 1-29.
- Dasaratha Rama, B., J. Milano, S. S., & Che-Hung Liu. Csr implementation: Developing the capacity for collective action. *Journal of Business Ethics*.
- Davenport, J. J. (2002). Fischer and Ravizza on moral sanity and weakness of will. *Journal of Ethics*, 6(3), 235–259.
- Davidson, D. (1963). Actions, reasons, and causes. *Journal of Philosophy*, 60, 685-700. (Reprinted in Davidson 1980)
- Davidson, D. (1970). How Is Weakness of the Will Possible? In J. Feinberg (Ed.), *Moral Concepts*: Oxford University Press.
- Davidson, D. (1973). Freedom to act. In T. Honderich (ed.), *Essays on Freedom of Action*., London: Routledge and Kegan Paul, pp. 137-56. (Reprinted in Davidson 1980, pp. 63-81.)

- Davidson, D. (1978). Intending. In Y. Yovel (ed.), *Philosophy of History and Action*. Dordrecht, Holland: D. Reidel, pp. 41-60. (Reprinted in Davidson 1980, pp. 83-102.)
- Davidson, D. (1980). *Essays on Actions and Events*, Oxford: Oxford University Press
- Davidson, D. (2004). *Problems of Rationality*, Oxford, Oxford University Press.
- Davis, J. K. (2007). Intuition and the junctures of judgment in decision procedures for clinical ethics. *Theoretical Medicine and Bioethics*, 28(1).
- Davis, L. H. (1994). Action, in S. Guttenplan (ed.), *A companion to the Philosophy of Mind*, Oxford: Blackwell, pp. 111-117.
- Davis, W. A. (2005). Reasons and psychological causes. *Philosophical Studies*, 122(1).
- de Bruin, B. (2006). Popper's Conception of the Rationality Principle in the Social Sciences. In I. Jarvie, D. Miller & K. Milford (Eds.), *Karl Popper: A Centenary Assessment: Selected Papers from Karl Popper 2002: Volume III: Science*: Ashgate.
- de Bruin, B. (2009). We and the plural subject. *Philosophy of the Social Sciences*, 39(2).
- Deutsch, Max (2009). Experimental philosophy and the theory of reference. *Mind and Language* 24 (4):445-466.
- Devitt, M. (2006). Intuitions in linguistics. *British Journal for the Philosophy of Science*, 57(3), 481-513.
- Dickenson, J. (2007). Reasons, causes, and contrasts. *Pacific Philosophical Quarterly*, 88(1), 1-23.
- Dilman, Ilham (2000). Psychology and human behaviour: Is there a limit to psychological explanation? *Philosophy* 75 (2):183-201.
- Diver, J. (2008). Attributions of Causation and Moral Responsibility. In Sinnott-Armstrong, W. (Ed.). *Moral Psychology*: MIT Press.
- Dodd, D. (2009). Weakness of will as intention-violation. *European Journal of Philosophy*, 17(1), 45-59.
- Donagan, A. (1987). *Choice: The Essential Element in Human Action*. London: Routledge.
- D'Oro, Giuseppina (2007). Two dogmas of contemporary philosophy of action. *Journal of the Philosophy of History* 1 (1):10-24.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. MIT Press.
- Dretske, F. (1989). Reasons and causes. *Philosophical Perspectives*, 3, 1-15.
- Dupuy, J-P. (1988). *Self-Deception and the Paradoxes of Rationality*. Stanford: CSLI Publications.
- Earlenbaugh, Joshua & Molyneux, Bernard (2009). Intuitions are inclinations to believe. *Philosophical Studies* 145 (1).
- Ehring, Douglas (1985). "Normal" intentional action. *Philosophy and Phenomenological Research* 46 (1):155-157.
- Elster, J. (1979). *Ulysses and the sirens: studies in rationality and irrationality*. Cambridge: Cambridge University Press.
- Elster, J. (1985). Rationality, morality, and collective action. *Ethics*, 96(1), 136-155.
- Elster, J. (1990). Racionalidade e Normas Sociais. *RBCS*, 5(12), 55-69.

- Elster, J. (1997). *Economics: analisis entre la interacción entre racionalidad, emoción, preferencias y normas sociales en la economía de la acción individual y sus desviaciones*. Barcelona: Gedisa Editorial.
- Elster, J. (1999). A Possibilidade da Política Racional. *RBCS*, 14(39), 13-40.
- Elster, Jon (2002). *Ulysses Unbound: Studies in Rationality, Precommitment, and Constraints*. Cambridge: Cambridge University Press.
- Ernst, Z., & Chant, S. R. (2007). Collective action as individual choice. *Studia Logica*, 86(3).
- Evrigenis, I. D. (2008). *Fear of Enemies and Collective Action*: Cambridge University Press.
- Farrell, Dan (1989). Intention, Reason, and Action, *American Philosophical Quarterly*, 26: 283-95
- Feltz, Adam & Cokely, Edward (2007). An anomaly in intentional action ascription: More evidence of folk diversity. *Proceedings of the Cognitive Science Society*.
- Feltz, Adam, Cokely, Edward T. & Nadelhoffer, Thomas (2009). Natural compatibilism versus natural incompatibilism: Back to the drawing board. *Mind and Language* 24 (1):1-23.
- Feltz, Adam (2008). Problems with the appeal to intuition in epistemology. *Philosophical Explorations* 11 (2):131 – 141.
- Ferejohn, J. (2002). Symposium on explanations and social ontology 1: Rational choice theory and social explanation. *Economics and Philosophy*, 18(2), 211-234.
- Fiala, Brian & Nichols, Shaun (2009). Confabulation, confidence, and introspection. *Behavioral and Brain Sciences* 32 (2):144-145.
- Fiani, R. (2006). *Teoria dos Jogos*. Rio de Janeiro: Elsevier.
- Finlay, S. (2008). Motivation to the means. In D. Chan (Ed.), *Values, Rational Choice and the Will*.
- Finocchiaro, Maurice A. (1979). The psychological explanation of reasoning: Logical and methodological problems. *Philosophy of the Social Sciences* 9 (3).
- Fleming, Patrick (2006). Berkeley's immaterialist account of action. *Journal of the History of Philosophy* 44 (3).
- Fletcher, G. (1995). *The Scientific Credibility of Folk Psychology*. Lawrence Erlbaum.
- Fodor, Jerry & LePore, Ernest (1992). *Holism: A Shopper's Guide*. Blackwell.
- Fodor, Jerry & Lepore, Ernest (1993). Is intentional ascription intrinsically normative? In B. Dahlbom (ed.), *Dennett and His Critics*. Blackwell.
- Fodor, Jerry (1968). The appeal to tacit knowledge in psychological explanation. *Journal of Philosophy* 65 (October):627-40.
- Fodor, Jerry (1974). Special sciences. *Synthese* 28:97-115.
- Fodor, Jerry (1983). *The Modularity of Mind*. MIT Press.
- Fodor, Jerry (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge (Massachusetts): MIT Press, 1987
- Fodor, Jerry (1991). You can fool some of the people all of the time, everything else being equal: Hedged laws and psychological explanation. *Mind* 100 (397):19-34.
- Fodor, Jerry (1994). *The Elm and the Expert*. MIT Press.

- Fodor, Jerry (1997). Special sciences: Still autonomous after all these years. *Philosophical Perspectives* 11:149-63.
- Føllesdal, D. (1982). The status of rationality assumptions in interpretation and in the explanation of action. *Dialectica* 36: 301-316.
- Frank, R. H. (2006). *Microeconomia e Comportamento*. Lisboa: McGraw-Hill.
- Frankena, W. (1988). Hare on moral weakness and the definition of morality. *Ethics*, 98(4), 779-792.
- Frankfurt, H. (1978). The problem of action. *American Philosophical Quarterly*, 15, 157-162. (Reprinted in H. Frankfurt, 1988. *The Importance of What we Care About*. Cambridge: Cambridge University Press, 69-79.)
- Frankfurt, H. (1988). *The Importance of What we Care About*. Cambridge: Cambridge University Press.
- Frankfurt, H. (1999). *Volition, Necessity, and Love*, Cambridge: Cambridge University Press
- Franklin, J. (1997). Stove's anti-darwinism. *Philosophy*, 72, 133-136.
- Frigerio, A., Giordani, A., & Mari, L. Outline of a general model of measurement. *Synthese*.
- Gallagher, S. (2005). Intentionality and intentional action. *Synthesis Philosophica*, 2(40), 319-326.
- Garzón, F. C. (2003). Nonclassical connectionism should enter the decathlon. *Behavioral and Brain Sciences*, 26(5), 603-604.
- Gauker, Christopher (2003). Attitudes without psychology. *Facta Philosophica* 5 (2):239-56.
- Gauker, Christopher (2005). The belief-desire law. *Facta Philosophica* 7.
- Gendler, T. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. In P. A. French & H. K. Wettstein (Eds.), *Philosophy and the Empirical* (Vol. 31, pp. 68–89): Blackwell Pub. Inc.
- Gert, Joshua (2003). Brute rationality. *Noûs* 37 (3):417–446.
- Gert, Joshua (2004). Intentional action and nearly certain success. *Ratio* 17 (2):150–158.
- Gert, Joshua (2005). A functional role analysis of reasons. *Philosophical Studies* 124 (3).
- Gert, Joshua (2008). Putting particularism in its place. *Pacific Philosophical Quarterly*, 89(3), 312-324.
- Gibbons, J. (2009). You gotta do what you gotta do. *Nous*, 43(1), 157-177.
- Gigerenzer, G. (2002). *Bounded Rationality: The Adaptive Toolbox*. Cambridge: The MIT Press.
- Gigerenzer, Gerd (2004). The irrationality paradox. *Behavioral and Brain Sciences* 27 (3):336-338.
- Gilbert, M. P. (2006). Rationality in collective action. *Philosophy of the Social Sciences*, 36(1).
- Ginet, Carl (1990). *On Action*, Cambridge: Cambridge University Press
- Godfrey-Smith, Peter (1981). Folk psychology as a model. *Philosopher's Imprint* 5 (6):1-16.
- Goldman, A. I. (1970). *A Theory of Human Action*. Englewood Cliffs, NJ: Prentice-Hall.

- Goldman, A. I. (1971). The individuation of action. *Journal of Philosophy*, 68(21), 761-774.
- Goldman, A. I. (1978). Chisholm's theory of action. *Philosophia*, 7(3-4).
- Goldman, A. I. (1979). Action, causation, and unity. *Noûs*, 13(2), 261-270.
- Goldman, A. I. (1989). Interpretation psychologized. *Mind and Language* 4: 161-185.
- Goldman, A. I. (1993). The psychology of folk psychology. *Behavioral and Brain Sciences* 16: 15-28.
- Goldman, A. I. (2006) *Simulating Minds: The Philosophy, Psychology and Neuroscience of Mindreading*. New York: Oxford University Press.
- Goldman, A. I. (2007). Philosophical intuitions: Their target, their source, and their epistemic status. *Grazer Philosophische Studien*, 74(1), 1-26.
- Gomatam, R. (2007). Niels Bohr's interpretation and the copenhagen interpretation—are the two incompatible? *Philosophy of Science*, 74(5).
- Gonnerman, C. (2008). Reading conflicted minds: An empirical follow-up to Knobe and roedder. *Philosophical Psychology*, 21(2), 193 – 205.
- Gopnik, Alison (2003). The theory theory as an alternative to the innateness hypothesis. In
- Gosling, J. C. B. (1990). *Weakness of the Will*: Routledge.
- Goswami, U. (2008). *Cognitive Development: the learning brain*. New York: Psychology Press.
- Greenwood, John D. (1991). *The Future of Folk Psychology: Intentionality and Cognitive Science*. Cambridge University Press.
- Grice, H. (1971). Intention and Uncertainty. *Proceedings of the British Academy* 57: 263-279.
- Griffiths, Paul E. & Stotz ,Karola (2008). Experimental philosophy of science. *Philosophy Compass* 3 (3):507–521.
- Grundmann, T. (2007). The nature of rational intuitions and a fresh look at the explanationist objection. *Grazer Philosophische Studien*, 74(1), 69-87.
- Haggard, Patrick & Clark, S. (2003). Intentional action: Conscious experience and neural prediction. *Consciousness and Cognition* 12 (4):695-707.
- Hahn, U. (2005). Is this what the debate on rules was about? *Behavioral and Brain Sciences*, 28(1), 25-26.
- Haldane, John J. (1988). Understanding folk. *Proceedings of the Aristotelian Society* 62:222-46.
- Hales, S. D. (2000). The Problem of Intuition. *American Philosophical Quarterly*, 37(2), 135-147.
- Hales, S. D. (2004). Intuition, revelation, and relativism. *International Journal of Philosophical Studies*, 12(3), 271 – 295.
- Halliday, Daniel (2005). What explains our intuitions about knowledge ascriptions? *Proceedings of the Aristotelian Society* 105 (3):393–402.
- Hannan, Barbara (1993). Don't stop believing: The case against eliminative materialism. *Mind and Language* 8 (2):165-179.
- Hanser, Matthew (1998). Intention and teleology. *Mind* 107 (426):381-401.

- Hare, R. M. (1996). *A Linguagem da Moral*. São Paulo: Martins Editora.
- Harman, Gilbert (1970). Knowledge, reasons, and causes. *Journal of Philosophy* 67 (21):841-855.
- Harman, Gilbert (1976). Practical reasoning. *Review of Metaphysics*, **79**, 431-63. Reprinted in A. R. Mele (ed.), *The Philosophy of Action*. Oxford: Oxford University Press, pp. 149-177.
- Harman, Gilbert (1986). *Change in View*. Cambridge, MA: MIT Press.
- Harman, Gilbert (1997). 'Practical Reasoning', *Review of Metaphysics*, 79: 431-63; reprinted in Mele.
- Hartmann, G. W. (1944). The strength and weakness of the pacifist position as seen by american philosophers. *Philosophical Review*, 53(2), 125-144.
- Hartshorne, C. (1943). Reflections on the strength and weakness of thomism. *Ethics*, 54(1), 53-57.
- Hauser, L. (1994). Acting, intending and artificial intelligence. *Behavior and Philosophy*, 22, 1: 22-28.
- Hausman, Daniel & McPherson, Michael S. (2009). Preference satisfaction and welfare economics. *Economics and Philosophy* 25 (1):1-25.
- Hausman, Daniel (1981). John Stuart mill's philosophy of economics. *Philosophy of Science* 48 (3):363-385.
- Hausman, Daniel (1984). Causal priority. *Noûs* 18 (2):261-279.
- Hausman, Daniel (1985). Is falsificationism unpractised or unpracticable? *Philosophy of the Social Sciences* 15 (3).
- Hausman, Daniel (1989). Are markets morally free zones? *Philosophy and Public Affairs* 18 (4):317-333.
- Hausman, Daniel (1989). The insufficiency of nomological explanation. *Philosophical Quarterly* 39 (154):22-35.
- Hausman, Daniel (1992). *Essays on Philosophy and Economic Methodology*. Cambridge University Press.
- Hausman, Daniel (1993). Linking causal and explanatory asymmetry. *Philosophy of Science* 60 (3):435-451.
- Hausman, Daniel (1993). Why don't effects explain their causes? *Synthese* 94 (2).
- Hausman, Daniel (1995). Rational choice and social theory: A comment. *Journal of Philosophy* 92 (2):96-102.
- Hausman, Daniel (1995). The impossibility of interpersonal utility comparisons. *Mind* 104 (415).
- Hausman, Daniel (1997). Causation, agency, and independence. *Philosophy of Science* 64 (4):25.
- Hausman, Daniel (1997). The impossibility of interpersonal utility comparisons--a reply. *Mind* 106 (421):99-100.
- Hausman, Daniel (1998). *Causal Asymmetries*. Cambridge University Press.
- Hausman, Daniel (2000). Revealed preference, belief, and game theory. *Economics and Philosophy* 16 (1):99-115.

- Hausman, Daniel (2003). Rational belief and social interaction. *Behavioral and Brain Sciences* 26 (2):163-164.
- Hausman, Daniel (2004). Polling and public policy. *Kennedy Institute of Ethics Journal* 14 (3).
- Hausman, Daniel (2005). Sympathy, commitment, and preference. *Economics and Philosophy* 21 (1):33-50.
- Hausman, Daniel (2006). Economic Analysis, Moral Philosophy, and Public Policy. Cambridge: Cambridge University Press, 2a Edição.
- Hausman, Daniel (2008). Fairness and social norms. *Philosophy of Science* 75 (5).
- Hausman, Daniel (2008). Philosophy of economics. *Stanford Encyclopedia of Philosophy*.
- Hausman, Daniel ed. (2008) The Philosophy of Economics: An Anthology. Cambridge: Cambridge University Press, 3a ed.
- Hayek, F. A. (1991). Individualismo e Coletivismo. tradução Anna Maria Capovilla et alli. In: BICHIR, A.S. *Antologia Liberal*. São Paulo: Editora Inconfidentes, p. 23-28.
- Heath, J. (2008). *Following the Rules: Practical Reasoning and Deontic Constraint*: Oxford University Press.
- Helm, B. W. (2001). Emotions and practical reason: Rethinking evaluation and motivation. *Nous*, 35(2), 190–213.
- Helm, B. W. (2008). Plural agents. *Nous*, 42(1), 17–49.
- Helzner, J. (2009). On the Application of Multiattribute Utility Theory to Models of Choice. *Theory and Decision*, 66(4), 301-315.
- Henden, E. (2004). Intentions, all-out evaluations and weakness of the will. *Erkenntnis*, 61(1).
- Henden, E. (2004). Weakness of will and divisions of the mind. *European Journal of Philosophy*, 12(2), 199–213.
- Herbert A. Simon, Alternative visions of rationality (1983). in Paul K. Moser (ed.), *Rationality in action: contemporary approaches*, Cambridge, Cambridge University Press, 1990, p.201-2.
- Herne, K., & Setälä, M. (2004). A response to the critique of rational choice theory: Lakatos' and Laudan's conceptions applied. *Inquiry*, 47(1), 67 – 85.
- Herschbach, Mitchell (2008). False-belief understanding and the phenomenological critics of folk psychology. *Journal of Consciousness Studies* 15 (12):33-56.
- Herschbach, Mitchell (2008). Folk psychological and phenomenological accounts of social perception. *Philosophical Explorations* 11 (3):223 – 235.
- Heuer, U. (2004). Reasons for actions and desires. *Philosophical Studies*, 121(1).
- Hindriks, Frank (2008). Intentional action and the praise-blame asymmetry. *Philosophical Quarterly* 58 (233):630-641.
- Hodgson, David (1994). Neuroscience and folk psychology: An overview. *Journal of Consciousness Studies* 1 (2):205-216.
- Hoffman, P. (2008). Freedom and weakness of will. *Ratio*, 21(1), 42–54.

- Hollis, M. & Nell, E. J. (1977). *O Homem Econômico Racional: Uma crítica filosófica da economia neoclássica*. Rio de Janeiro: Zahar Editores
- Hollis, Martin & Sugden, Robert (1993). Rationality in action. *Mind* 102 (405):1-35.
- Holton, R. (1999). Intention and weakness of will. *Journal of Philosophy*, 96(5), 241-262.
- Holton, R. (2007). How is strength of will possible? In C. Tappolet & S. Stroud (Eds.), *Weakness of Will and Practical Irrationality*. Oxford.
- Holton, R. (2009). *Willing, Wanting, Waiting*: Oxford University Press.
- Hornsby, Jennifer (1980). *Actions*, London: Routledge & Kegan Paul.
- Hume, D. (2001). *Tratado da Natureza Humana*. Lisboa: Fundação Galouste Gulbenkian.
- Hurley, S. (1991). Newcomb's problem, prisoners' dilemma, and collective action. *Synthese*, 86(2).
- Hurley, P. (1992). How weakness of the will is possible. *Mind*, 101(401).
- Hurley, S. (1997). Non-conceptual self-consciousness and agency; perspective and access. *Communication and Cognition*, 30, 3/4, 207-248.
- Hurley, S. (1998). *Consciousness in Action*, Cambridge, Mass.:Harvard University Press.
- Hursthouse, R. (1991). Arational Actions. *Journal of Philosophy*, 88, 2: 57-68.
- Hutto, Daniel D. (2004). Folk psychological narratives and the case of autism. *Philosophical Papers* 32 (3):345-361.
- Hutto, Daniel D. (2004). The limits of spectatorial folk psychology. *Mind and Language* 19 (5):548-73.
- Ichikawa, J., & Jarvis, B. (2007). Thought-experiment intuitions and truth in fiction. *Philosophical Studies*, 142(2), 221-246.
- Israel, D., Perry, J. & Tutiya, S. (1991). Actions and Movements. *Proceedings of IJCAI-91*, (Sydney, Australia), August 1991.
- Jackman, Henry (2005). Intuitions and semantic theory. *Metaphilosophy* 36 (3):363-380.
- Jackson, Frank & Pettit, Philip (1990). In defense of folk psychology. *Philosophical Studies* 59 (1):31-54.
- Jackson, Frank & Pettit, Philip (1993). Folk belief and commonplace belief. *Mind and Language* 8 (2):298-305.
- Jackson, F. (1984). Weakness of will. *Mind*, 93(369), 1-18.
- Jackson, F. (2000). Psychological explanation and implicit theory. *Philosophical Explorations* 3 (1):83-95.
- Jackson, F. (2008). Review of Knobe & Shaun Nichols (eds.), *Experimental Philosophy*. *Notre Dame Philosophical Reviews*, 2008(12).
- Jerolmack, C., & Porpora, D. (2004). Religion, rationality, and experience: A response to the new rational choice theory of religion. *Sociological Theory*, 22(1), 140-160.
- Jeske, Diane (2009). *Rationality and Moral Theory: How Intimacy Generates Reasons*. Routledge.
- Johnson, J. (1996). How not to criticize rational choice theory: Pathologies of "common sense". *Philosophy of the Social Sciences*, 26(1).

- Johnson, K. (1975). A note on the inapplicability of Olson's logic of collective action to the state. *Ethics*, 85(2), 170-174.
- Jones, K. (2003). Emotion, Weakness of Will, and the Normative Conception of Agency. In A. Hatzimoysis (Ed.), *Philosophy and the Emotions*: Cambridge University Press.
- Jü, Schrö, r., & Der. (1999). What has consciousness to do with explicit representations and stable activation vectors? *Behavioral and Brain Sciences*, 22(1), 166-167.
- Kahneman, Daniel , Slovic, Paul & Tversky, Amos (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.
- Kauppinen, Antti (2007). The rise and fall of experimental philosophy. *Philosophical Explorations* 10 (2):95 – 118.
- Kaufmann, P. (1998) *Dicionário Enciclopédico de Psicanálise: o legado de Freud e Lacan*. Rio de Janeiro: Jorge Zahar Editor
- Kavka, G. (1983). The Toxin Puzzle. *Analysis*, 43: 33-36.
- Kelly, T. (2003). Epistemic rationality as instrumental rationality: A critique. *Philosophy and Phenomenological Research*, 66(3), 612–640.
- Kennett, J. (2001). *Agency and Responsibility: A Common-Sense Moral Psychology*: Oxford University Press.
- Kennett, J., & Smith, M. (1996). Frog and toad lose control. *Analysis*, 56(2), 63–73.
- Kincaid, Harold (1990). Eliminativism and methodological individualism. *Philosophy of Science* 57 (1):141-148.
- Kitcher , P. S. (1984). In defense of intentional psychology. *Journal of Philosophy* 81 (February):89-106.
- Knobe, J. (2003a). Intentional action and side effects in ordinary language. *Analysis*, 63(3), 190–194.
- Knobe, J. (2003b). Intentional action in folk psychology: An experimental investigation. *Philosophical Psychology*, 16(2), 309-325.
- Knobe, J. (2004). Intention, intentional action and moral considerations. *Analysis*, 64(2), 181–187.
- Knobe, J. (2005). Ordinary ethical reasoning and the ideal of 'being yourself'. *Philosophical Psychology*, 18(3), 327 – 340.
- Knobe, J. (2005). Theory of mind and moral cognition: Exploring the connections. *Trends in Cognitive Sciences*, 9, 357-359.
- Knobe, J. (2006). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies*, 130(2), 203-231.
- Knobe, J. (2006). The folk concepts of intention and intentional action: A cross-cultural study. *Journal of Cognition and Culture*, 6(1-2), 113-132.
- Knobe, J. (2007). Experimental philosophy and philosophical significance. *Philosophical Explorations*, 10(2), 119 – 121.
- Knobe, J. (2007). Experimental philosophy. *Philosophy Compass*, 2(1), 81–92.
- Knobe, J. (2007). Folk psychology: Science and morals. In D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Reassessed*: Springer Press.

- Knobe, J. (2007). Reason explanation in folk psychology. *Midwest Studies in Philosophy*, 31(1), 90–106.
- Knobe, J. (2009). Answers to five questions. In Jesús H. Aguilar & Andrei A. Buckareff (eds.), *Philosophy of Action: 5 Questions*. Automatic Press.
- Knobe, J. (forthcoming). What is experimental philosophy? *The Philosophers' Magazine*.
- Knobe, J., & Burra, A. (2006). Experimental philosophy and folk concepts: Methodological considerations. *Journal of Cognition and Culture*, 6(1-2), 331-342.
- Knobe, J., Cohen, A., & Leslie, A. (2006). Acting intentionally and the side-effect effect: 'Theory of mind' and moral judgment. *Psychological Science*, 17, 421-427.
- Knobe, J., Cushman, F., & Sinnott-Armstrong, W. (2008). Moral appraisals affect doing/allowing judgments. *Cognition*, 108(2), 353-380.
- Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (Ed.), *Moral Psychology*: MIT Press.
- Knobe, J., & Malle, B. (1997). The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33, 101-121.
- Knobe, J. & Malle, B. (2002). Self and other in the explanation of behavior: 30 years later. *Psychologica Belgica* 42:113-130.
- Knobe, J., Malle, B. F. & Nelson, S. (2007). Actor-observer asymmetries in explanations of behavior: New answers to an old question. *Journal of Personality and Social Psychology* 9 (4):491-514.
- Knobe, J., & Mendlow, G. (2004). The good, the bad and the blameworthy: Understanding the role of evaluative reasoning in folk psychology. *Journal of Theoretical and Philosophical Psychology*, 24, 252-258.
- Knobe, J. & Nichols, S. (2008) *Experimental Philosophy*. New York: Oxford University Press.
- Knobe, J., & Prinz, J. J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, 7(1), 67-83.
- Knowles, Jonathan (2001). Does intentional psychology need vindicating by cognitive science? *Minds and Machines* 11 (3):347-377.
- Knowles, Jonathan (2002). Is folk psychology different? *Erkenntnis* 57 (2):199-230.
- Kolodny, Niko (2008). The myth of practical consistency. *European Journal of Philosophy* 16 (3):366-402.
- Korkman, P. (2009). Weakness of will from Plato to the present (review). *Journal of the History of Philosophy*, 47(3), pp. 466-467.
- Korsgaard, C. M. The activity of reason. *The Proceedings and Addresses of the American Philosophical Association*.
- Koutoungos, A. (2008). Moral sensitivity and desire attachment: In what sense are they constituents of one's rational profile? *Acta Analytica*, 23(2).
- Kraus, Jody S. & Coleman, Jules L. (1987). Morality and the theory of rational choice. *Ethics* 97 (4):715-749.
- Kuhn, D., & Udell, W. (2007). Coordinating own and other perspectives in argument. *Thinking and Reasoning*, 13(2), 90 – 104.

- Lagier, D. G. (2001). *Las Paradojas de la Acción*. Alicante: Publicaciones Universidad de Alicante.
- Lahav, Ran (1992). The amazing predictive power of folk psychology. *Australasian Journal of Philosophy* 70 (1):99-105.
- Landemore, H. (2004). Politics and the economist-King: Is rational choice theory the science of choice? *Journal of Moral Philosophy*, 1(2).
- Landy, J. (2004). *Philosophy As Fiction: Self, Deception, and Knowledge in Proust*: Oxford University Press.
- Langsam, Harold (2008). Rationality, justification, and the internalism/externalism debate. *Erkenntnis* 68 (1).
- Lehtinen, A., & Kuorikoski, J. (2007). Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences*, 37(2).
- Lenman, J. (2009). Reasons for action: Justification vs. explanation. *The Stanford Encyclopedia of Philosophy*.
- Levin, J. (2004). The evidential status of philosophical intuition. *Philosophical Studies*, 121(3), 193-224.
- Levin, J. (2007). Can modal intuitions be evidence for essentialist claims? *Inquiry*, 50(3), 253 – 269.
- Levine, A., Sober, E. & Wright, E. O. (1989). Marxismo e Individualismo Metodológico. *RBCS*, 4(11), 57-70.
- Levine, J. (1987). The nature of psychological explanation by Robert Cummins: A critical notice. *Philosophical Review* 96 (2):249-274.
- Levy, N. (2008). Review of _Experimental Philosophy_. *Metapsychology*, 12(33).
- Levy, N. Resisting weakness of the will. *Philosophy & Phenomenological Research*.
- Liao, S. (2008). A defense of intuitions. *Philosophical Studies* 140 (2), 247-262.
- Liao, S. (2009). Aristotle's view on "the right of practice": An investigation into Aristotle's theory of action. *Frontiers of Philosophy in China*, 4(2).
- Lichtenberg, J. (1983). Subjectivism as moral weakness projected. *Philosophical Quarterly*, 33(133), 378-385.
- Lippert-rasmussen, K., & Jensen, K. K. (2002). Does particularism solve the moral problem? *Philosophical Explorations*, 5(2), 125 – 140.
- Little, D. (1988). Collective action and the traditional village. *Journal of Agricultural and Environmental Ethics*, 1(1).
- Littlejohn, C. (2009). On Treating Something as a Reason for Action. *Journal of Ethics and Social Philosophy*.
- Livengood, J., & Machery, E. (2007). The folk probably don't think what you think they think: Experiments on causation by absence. *Midwest Studies in Philosophy*, 31(1), 107–127.
- Livingston, Paisley (1991). *Literature and Rationality: Ideas of Agency in Theory and Fiction*. Cambridge University Press.
- Lord, E. (2008). Dancy on Acting for the Right Reason. *Journal of Ethics & Social Philosophy*, 1-7.

- Ludwig, K. (2007). The epistemology of thought experiments : First person versus third person approaches. In P. A. French & H. K. Wettstein (Eds.), *Philosophy and the Empirical* (Vol. 31, pp. 128–159): Blackwell Pub. Inc.
- Lukes, S. (1965). Moral weakness. *Philosophical Quarterly*, 15(59), 104-114.
- Lumer, C., & Nannini, S. (2007). *Intentionality, deliberation and autonomy: the action-theoretic basis of practical philosophy*: Ashgate Publishing.
- Lyons, D. (1966). The weakness of formal equality. *Ethics*, 76(2), 146-148.
- Macdonald, C. (2002). Theories of mind and 'the commonsense view'. *Mind and Language* 17 (5):467-488.
- Macdonald, Cynthia (2006). 'The Metaphysics of Mental Causation'. *The Journal of Philosophy* 103:539-576.
- Machery, E. (2006). The folk concept of intentional action: Philosophical and experimental issues. *Mind and Language*, 23(2), 165–189.
- Machery, E. (2006). Understanding the folk concept of intentional action: Philosophical and experimental issues. *Mind and Language*.
- Machery, E., Mallon, R., Nichols, S., & Stich, S. P. (2004). Semantics, cross-cultural style. *Cognition*, 92(3).
- Malle, Bertrand, Moses L. J., and Baldwin D. A. (eds) (2001). *Intentions and Intentionality - Foundations of Social Cognition*. Cambridge, MA: MIT Press.
- Malle, Bertram & Sara D. Hodges (2005). *Other Minds: How Humans Bridge the Gap Between Self and Others*. Guilford.
- Malle, Bertram (1999). How People Explain Behavior: A New Theoretical Framework. *Personality and Social Psychology Review*. Vol. 3, N. 1, 23-48.
- Malle, Bertram (2004). *How the Mind Explains Behavior: Folk Explanations, Meaning, and Social Interaction*. MIT Press.
- Malle, Bertram (2005). Folk Theory of Mind: Conceptual Foundations of Human Social Cognition. In Ran R. Hassin, James S. Uleman & John A. Bargh (eds.), *The New Unconscious*. Oxford Series in Social Cognition and Social Neuroscience. Oxford University Press.
- Malle, Bertram (2006). Of windmills and straw men: Folk assumptions of mind and action. In Susan Pockett, William P. Banks & Shaun Gallagher (eds.), *Does Consciousness Cause Behavior?* MIT Press.
- Malle, Bertram (2006). The relation between judgments of intentionality and morality. *Journal of Cognition and Culture* 6:61-86.
- Mallon, R. (2008). Knobe versus Machery: Testing the trade-off hypothesis. *Mind and Language*, 23(2), 247-255.
- Manson, N. C. (2004). Reason explanation a first-order rationalizing account. *Philosophical Explorations*, 7(2), 113 – 129.
- Marchi ,Neil de (1986). Mill's unrevised philosophy of economics: A comment on Hausman. *Philosophy of Science* 53 (1):89-100.
- Margolis, Eric & Laurence, Stephen (2003). Should we trust our intuitions? Deflationary accounts of the analytic data. *Proceedings of the Aristotelian Society* 103 (3):299-323.

- Margolis, Joseph (1980). The trouble with homunculus theories. *Philosophy of Science* 47 (June):244-259.
- Margolis, Joseph (1991). The autonomy of folk psychology. In John D. Greenwood (ed.), *The Future of Folk Psychology*. Cambridge University Press.
- Martin, Michael (1971). Neurophysiological reduction and psychological explanation. *Philosophy of the Social Sciences* 1 (1).
- Mas-Colell, A., Whinston, M. D., Green, J. R. (1995). *Microeconomic Theory*. New York: Oxford University Press
- Matthews, G. (1966). Weakness of will. *Mind*, 75(299), 405-419.
- Mauro, C., Miguens, S., Cadilha, S. (Eds) (no prelo a). *Conversations on Human Action and Rationality*.
- Mauro, C. (2002). Algumas considerações sobre Escolha Racional e Corrupção Pública e Política. *Estratégica*, 1 (3), 19-30.
- Mauro, C. & Maçorano, J. P. (2006b). Racionalidade na Acção – uma explicação a partir do modelo crença-desejo. *Revista de Economia e Relações Internacionais*, 5 (8), 2006.
- Mauro, C. & Cadilha, S. (2006a). Por que não pode existir uma acção irracional. In Miguens, Sofia & Mauro, Carlos (Eds). *Perspectives on Rationality*. Porto: FLUP. *O Estatuto do Singular: Estratégias e Perspectivas*, Lisboa: INCM.
- Mauro, C., Miguens, S., Cadilha, S. (Eds) (2009a). *Mente Linguagem e Acção: textos para discussão*. Colecção Campo da Filosofia. Porto: Campo das Letras.
- Mauro, C. (2009b). Será que de facto eu decido?. In Mauro, C.; Miguens, S.; Cadilha, S. (Eds). *Mente Linguagem e Acção: Artigos para Discussão*. Porto: Campo das Letras.
- Mauro, C. A (2009c). Não Inevitabilidade da Identidade Nacional. *Diversitates*, Rio de Janeiro: UFRJ, v.1, n.1.
- Mauro, C., Cadilha, S. (Eds) (no prelo b). *Diálogos sobre a Acção Humana*.
- Miguens, S. & Mauro, C. (Eds) (2006c). *Perspectives on Rationality*. Porto: FLUP.
- Mauro, C. & Cadilha, S. (2009d). Racionalidade na Acção: podemos de facto explicar (racionalizar) as nossas acções?. In M. L. Couto Soares, N. Venturinha e G. C. Santos (org.), Mauro, C. (2005). Somos ou não inevitavelmente racionais? Uma argumentação a partir de elementos da teoria económica. *Intelectu*, nº 11.
- Mauro, C. (2003). Uma análise da Corrupção Pública a partir da Racionalidade e da Acção. *Revista de Economia e Relações Internacionais*, 1 (2), 75-83.
- May, J. (forthcoming) Review of _Experimental Philosophy_ ed. by Knobe and Nichols. *Philosophical Psychology*.
- May, J., Sinnott-Armstrong, W., Hull, J. G., & Zimmerman, A. Z. Practical Interests, Relevant Alternatives, and Knowledge Attributions: An Empirical Study. *Review of Philosophy and Psychology*.
- McCann, Hugh, (1986). 'Rationality and the Range of Intention', *Midwest Studies in Philosophy* 10: 191-211.
- McCann, Hugh (1998). *The Works of Agency: On Human Action, Will and Freedom*. Ithaca, New York: Cornell University Press.

- McCann, Hugh J. (1991). Settled objectives and rational constraints. *American Philosophical Quarterly* 28 (1):25-36.
- McCann, Hugh J. (1998). *The Works of Agency: On Human Action, Will, and Freedom*. Cornell University Press.
- McCann, Hugh J. (2005). Intentional action and intending: Recent empirical studies. *Philosophical Psychology* 18 (6):737-748.
- McDonough, Richard M. (1991). A culturalist account of folk psychology. In John D. Greenwood (ed.), *The Future of Folk Psychology*. Cambridge University Press.
- McGavock, K. L. (2007). Agents of reform?: Children's literature and philosophy. *Philosophia*, 35(2).
- McIntyre, A. (2006). What is Wrong with Weakness of Will? *Journal of Philosophy*, 103(6), 284-311.
- Meeks, R. R. (2004). Unintentionally biasing the data: Reply to Knobe. *Journal of Theoretical and Philosophical Psychology*, 24, 220-223.
- Mele, Alfred & Cushman, Fiery (2007). Intentional action, folk judgments, and stories: Sorting things out. *Midwest Studies in Philosophy* 31 (1):184-201.
- Mele, Alfred & Paul K. Moser (1994). Intentional action. *Noûs* 28 (1):39-68.
- Mele, Alfred & Piers Rawling (2004). *The Oxford Handbook of Rationality*. Oxford University Press.
- Mele, Alfred & Sverdlik, Steven (1996). Intention, intentional action, and moral responsibility. *Philosophical Studies* 82 (3).
- Mele, Alfred (1982). 'Self-deception, action, and will': Comments. *Erkenntnis* 18 (2):159-164.
- Mele, Alfred (1983). Akrasia, reasons, and causes. *Philosophical Studies* 44 (3):345-368.
- Mele, Alfred (1983). Self-deception. *Philosophical Quarterly* 33 (October):366-377.
- Mele, Alfred (1984). Aristotle's Wish. *Journal of the History of Philosophy* 22 (2):139-156.
- Mele, Alfred (1984). Pears on akrasia, and defeated intentions. *Philosophia* 14 (1-2).
- Mele, Alfred (1986). Incontinent believing. *Philosophical Quarterly* 36 (143):212-222.
- Mele, Alfred (1986). Is akratic action unfree? *Philosophy and Phenomenological Research* 46 (4):673-679.
- Mele, Alfred (1987). Are intentions self-referential? *Philosophical Studies* 52 (3):309-329.
- Mele, Alfred (1987). Intentional action and wayward causal chains: The problem of tertiary waywardness. *Philosophical Studies* 51 (1).
- Mele, Alfred (1987). *Irrationality: An Essay on Akrasia, Self-Deception, and Self-Control*. Oxford University Press.
- Mele, Alfred (1988). Against a belief/desire analysis of intention. *Philosophia* 18 (2-3):239-242.
- Mele, Alfred (1988). Effective reasons and intrinsically motivated actions. *Philosophy and Phenomenological Research* 48 (4):723-731.
- Mele, Alfred (1988). Irrationality: A précis. *Philosophical Psychology* 1 (2):173-177.

- Mele, Alfred (1989). Akratic feelings. *Philosophy and Phenomenological Research* 50 (2):277-288.
- Mele, Alfred (1989). Motivational internalism: The powers and limits of practical reasoning. *Philosophia* 19 (4).
- Mele, Alfred (1989). She intends to try. *Philosophical Studies* 55 (1):101-106.
- Mele, Alfred (1990). Exciting intentions. *Philosophical Studies* 59 (3):289-312.
- Mele, Alfred (1990). Irresistible desires. *Noûs* 24 (3):455-72.
- Mele, Alfred (1992). Akrasia, self-control, and second-order desires. *Noûs* 26 (3):281-302.
- Mele, Alfred (1992). Intending for reasons. *Mind* 101 (402):327-333.
- Mele, Alfred (1992). Intentions, reasons, and beliefs: Morals of the toxin puzzle. *Philosophical Studies* 68 (2).
- Mele, Alfred (1992). *The Springs of Action*, New York: Oxford University Press.
- Mele, Alfred (1993). Justifying intentions. *Mind* 102 (406):335-337.
- Mele, Alfred (1995). *Autonomous Agents: From Self-Control to Autonomy*. Oxford University Press.
- Mele, Alfred (1995). Effective deliberation about what to intend: Or striking it rich in a toxin-free environment. *Philosophical Studies* 79 (1).
- Mele, Alfred (1995). Motivation: Essentially motivation-constituting attitudes. *Philosophical Review* 104 (3):387-423.
- Mele, Alfred (1997). Agency and mental action. *Philosophical Perspectives* 11:231-249.
- Mele, Alfred (1997). Real self-deception. *Behavioral and Brain Sciences* 20 (1):91-102.
- Mele, Alfred (1998). Motivated belief and agency. *Philosophical Psychology* 11 (3):353 – 369.
- Mele, Alfred (1998). Motivational strength. *Noûs* 32 (1):23-36.
- Mele, Alfred (1998). Noninstrumental rationalizing. *Pacific Philosophical Quarterly* 79 (3):236–250.
- Mele, Alfred (1998). Synchronic self-control revisited: Frog and toad shape up. *Analysis* 58 (4):305–310.
- Mele, Alfred (1999). Is there a place for intention in an analysis of intentional action? *Philosophia* 27 (3-4).
- Mele, Alfred (2000). Deciding to act. *Philosophical Studies* 100 (1):81–108.
- Mele, Alfred (2000). Reactive attitudes, reactivity, and omissions. *Philosophy and Phenomenological Research* 61 (2):447-452.
- Mele, Alfred (2000). Self-deception and emotion. *Consciousness and Emotion* 1 (1):115-137.
- Mele, Alfred (2001). *Autonomous Agents*, Oxford: Oxford University Press.
- Mele, Alfred (2001). *Self-Deception Unmasked*. Princeton University Press.
- Mele, Alfred (2002). Autonomy and akrasia. *Philosophical Explorations* 5 (3):207 – 216.
- Mele, Alfred (2003) *Motivation and Agency*. Oxford: Oxford University Press.

- Mele, Alfred (2003). Agents' abilities. *Noûs* 37 (3):447–470.
- Mele, Alfred (2003). Emotion and desire in self-deception. *Philosophy* 52:163-179.
- Mele, Alfred (2003). Intentional action: Controversies, data, and core hypotheses. *Philosophical Psychology* 16 (2):325-340.
- Mele, Alfred (2005). Decisions, intentions, and free will. *Midwest Studies in Philosophy* 29 (1):146-162.
- Mele, Alfred (2005). Motivation and agency: Précis. *Philosophical Studies* 123 (3):243–247.
- Mele, Alfred (2006). The folk concept of intentional action: A commentary. *Journal of Cognition and Culture*.
- Mele, Alfred (2007). Persisting intentions. *Noûs* 41 (4):735–757.
- Mele, Alfred (2008). Proximal intentions, intention-reports, and vetoing. *Philosophical Psychology* 21 (1):1 – 14.
- Mele, Alfred (2009). *Effective Intentions: The Power of Conscious Will*. Oxford University Press.
- Mele, Alfred (2009). Moral responsibility and agents' histories. *Philosophical Studies* 142 (2).
- Mele, Alfred (ed) (1997). *The Philosophy of Action*, Oxford: Oxford University Press.
- Mele, Alfred (forthcoming). Moral responsibility and history revisited. *Ethical Theory and Moral Practice*.
- Mele, Alfred (forthcoming). Weakness of will and akrasia. *Philosophical Studies*.
- Miguens, S. (2002). *Uma Teoria Fisicalista do Conteúdo e da Consciência ? D. Dennett e os debates da filosofia da mente*. Porto: Campo das Letras.
- Miguens, S. (2004). *Racionalidade*. Porto: Campo das Letras.
- Miguens, S. (2007). *Filosofia da Linguagem - uma introdução*. Porto: Faculdade de Letras da Universidade do Porto.
- Miguens, S. (2008). *Será que a minha mente está dentro da minha cabeça? Da ciência cognitiva à filosofia (ensaios)*. Porto: Campo das Letras.
- Mill, J. S. (1996). *Princípios de Economia Política*. Vol. II. São Paulo: Nova Cultural, Coleção Os Economistas. (1º edição em 1848)
- Millar, Alan (2004). *Understanding People: Normativity and Rationalizing Explanation*. Oxford University Press.
- Miller, C. (2008). Gert on subjective practical rationality. *Ethical Theory and Moral Practice*, 11(5).
- Miščević, Nenad (2004). The explainability of intuitions. *Dialectica* 58 (1):43–70.
- Mises, L. von. (1996). *Human Action: A Treatise on Economics*. San Francisco: Fox & Wilkes.
- Mongin, P. (1991). Rational choice theory considered as psychology and moral philosophy. *Philosophy of the Social Sciences*, 21(1).
- Moore, M. A. The actor-observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire.

- Moran, Richard (2001). *Authority and Estrangement: An Essay on Self-Knowledge*, Princeton: Princeton University Press
- Moran, Richard (2004). 'Anscombe on Practical Knowledge,' *Philosophy*, 55 (Supp): 43-68
- Morris, Stephen G. (2005). Identifying the explanatory weakness of strong altruism: The needle in the 'haystack model'. *Philosophy of Science*, 72(5), 1124-1134.
- Morse, S. J. (1994). Causation, Compulsion, and Involuntariness. *Bulletin of the American Academy of Psychiatry and Law*, 22, 2: 159-180
- Morton, Adam (1991). The inevitability of folk psychology. In R. Bogdan (ed.), *Mind and Common Sense*. Cambridge University Press.
- Morton, Adam (1996). Folk psychology is not a predictive device. *Mind* 105 (417):119-37.
- Morton, Adam (2003). *The Importance of Being Understood: Folk Psychology As Ethics*. New York: Routledge.
- Morton, Adam (2006). But are they right? The prospects for empirical conceptology. *Journal of Cognition and Culture*, 6.
- Muñoz, A. A. (2002). *Liberdade e Causalidade: acção, responsabilidade e metafísica em Aristóteles*. São Paulo: Discurso Editorial.
- Nadelhoffer, T., & Feltz, A. (2008). The actor–observer bias and moral intuitions: Adding fuel to Sinnott-Armstrong's fire. *Neuroethics*, 1(2).
- Nadelhoffer, T., & Nahmias, E. (2007). The past and future of experimental philosophy. *Philosophical Explorations*, 10(2), 123 – 149.
- Nadelhoffer, T. (2004). On praise, side effects, and folk ascriptions of intentionality. *Journal of Theoretical and Philosophical Psychology* 24:196-213.
- Nadelhoffer, T. (2005). Skill, luck, control, and intentional action. *Philosophical Psychology*, 18(3), 341 – 352.
- Nadelhoffer, T. (2006). On trying to save the simple view. *Mind and Language* 21 (5):565-586.
- Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for juror impartiality. *Philosophical Explorations*, 9(2), 203 – 219.
- Nadelhoffer, T. (2006). Desire, foresight, intentions, and intentional actions: Probing folk intuitions. *Journal of Cognition and Culture*.
- Nagel, J. (2007). Epistemic intuitions. *Philosophy Compass*, 2(6), 792–819.
- Nahmias, Eddy A., Stephen G. Morris & Thomas Nadelhoffer (2005). Surveying freedom: Folk intuitions about free will and moral responsibility. *Philosophical Psychology* 18 (5):561 – 584.
- Nahmias, Eddy, D. Justin Coates & And Trevor Kvaran (2007). Free will, moral responsibility, and mechanism: Experiments on folk intuitions. *Midwest Studies in Philosophy* 31 (1):214–242.
- Nahmias, Eddy, Morris, Stephen G., Nadelhoffer, Thomas & Turner, Jason (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research* 73 (1):28–53.
- Nahmias, Eddy A. (2006). Close calls and the confident agent: Free will, deliberation, and alternative possibilities. *Philosophical Studies* 131 (3):627-667.

- Nahmias, Eddy A. (2006). Folk fears about freedom and responsibility: Determinism vs. reductionism. *Journal of Cognition and Culture*.
- Nahmias, Eddy A. (2006). Free will and the folk: Responses to commentators. *Journal of Cognition and Culture* 6:305-320.
- Nichols, Shaun & Claudia Uller (1999). Explicit factuality and comparative evidence. *Behavioral and Brain Sciences* 22 (5):776-777.
- Nichols, S. and Folds-Bennett, T. (2003). Are Children Moral Objectivists? Children's Judgments about Moral and Response-Dependent Properties. *Cognition*, 90, B23-32
- Nichols, Shaun & Joseph Ulatowski (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind and Language* 22 (4):346–365.
- Nichols, Shaun & Knobe, Joshua (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs* 41 (4):663–685.
- Nichols, Shaun & Stich, Stephen (1994). Folk psychology. *Encyclopedia of Cognitive Science*.
- Nichols, Shaun & Stich, Stephen (1998). Rethinking co-cognition: A reply to Heal. *Mind and Language* 13 (4):499-512.
- Nichols, Shaun & Stich, Stephen (2003). *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press.
- Nichols, Shaun & Stich, Stephen (2003). Reading one's own mind: A cognitive theory of self-awareness. In Quentin Smith & Aleksandar Jokic (eds.), *Consciousness: New Philosophical Perspectives*. Oup.
- Nichols, Shaun & Todd A. Grantham (2000). Adaptive complexity and phenomenal consciousness. *Philosophy Of Science* 67 (4):648-670.
- Nichols, Shaun (2000). The mind's "I" and the theory of mind's "I": Introspection and two concepts of self. *Philosophical Topics* 28:171-99.
- Nichols, Shaun (2001). Mindreading and the cognitive architecture underlying altruistic motivation. *Mind and Language* 16 (4):425-455.
- Nichols, Shaun (2002). Imagination and the puzzles of iteration. *Analysis* 62 (3):182-87.
- Nichols, Shaun (2002). On the genealogy of norms: A case for the role of emotion in cultural evolution. *Philosophy of Science* 69 (2).
- Nichols, Shaun (2004a). After objectivity: An empirical study of moral judgment. *Philosophical Psychology* 17 (1):3 – 26.
- Nichols, Shaun (2004b). Folk concepts and intuitions: From philosophy to cognitive science. *Trends in Cognitive Sciences*.
- Nichols, Shaun (2004c). Imagining and believing: The promise of a single code. *Journal of Aesthetics and Art Criticism* 62 (2):129-39.
- Nichols, Shaun (2004d). Review: Recreative minds. *Mind* 113 (450).
- Nichols, Shaun (2004e). *Sentimental Rules: On the Natural Foundations of Moral Judgment*. Oxford University Press.
- Nichols, Shaun (2004f). The folk psychology of free will: Fits and starts. *Mind and Language* 19 (5):473-502.

- Nichols, Shaun (2005). Innateness and moral psychology. In Peter Carruthers, Stephen Laurence & Stephen Stich (eds.), *The Innate Mind: Structure and Contents*. New York: Oxford University Press New York.
- Nichols, Shaun (2006). Do children think of the self as the soul? *Behavioral and Brain Sciences* 29 (5):481-482.
- Nichols, Shaun (2006). Folk intuitions on free will. *Journal of Cognition and Culture*.
- Nichols, Shaun (2006). Just the imagination: Why imagining doesn't behave like believing. *Mind and Language* 21 (4):459–474.
- Nichols, Shaun (2006). *The Architecture of the Imagination: New Essays on Pretence, Possibility, and Fiction*. Oxford University Press.
- Nichols, Shaun (2007). The rise of compatibilism: A case study in the quantitative history of philosophy. *Midwest Studies in Philosophy* 31 (1):260-270.
- Nichols, Shaun (2008). Imagination and the I. *Mind and Language* 23 (5):518-535.
- Nichols, Shaun (forthcoming). How can psychology contribute to the free will debate? In J. Baer, J. Kaufman & R. Baumeister (eds.), *Psychology and Free Will*. Oxford University Press.
- Nichols, Shaun (forthcoming). Imagination and immortality: Thinking of me. *Synthese*.
- Norrie, A. (1984). Thomas Hobbes and the philosophy of punishment. *Law and Philosophy*, 3(2).
- Nozick, R. (1993). *The Nature of Rationality*. Princeton: Princeton University Press.
- Nudds, Matthew (2001). Common-sense and scientific psychology. *Croatian Journal of Philosophy* 1 (2):171-180.
- O'Brien, Gerard (1987). Eliminative materialism and our psychological self-knowledge. *Philosophical Studies* 52 (July):49-70.
- O'Brien, Gerard (1993). A conflation of folk psychologies. *Prospects for Intentionality Working Papers in Philosophy* 3:42-51.
- Olson, M. (1971). *The Logic of Collective Action*. Cambridge: Cambridge University Press.
- Orenstein, L. (1998). *A estratégia da Ação Coletiva*. Rio de Janeiro: Revan: UCAN, IUPERJ.
- Overskeid, G. (2002). Psychological hedonism and the nature of motivation: Bertrand Russell's anhedonic desires. *Philosophical Psychology*, 15(1), 77 – 93.
- Owens, Joseph (1994). Psychological externalism and psychological explanation. *Philosophy and Phenomenological Research* 54 (4):921-928.
- Pacherie, E. (2000). The content of intentions. *Mind and Language* 15: 400-432.
- Pacherie, E. (2001). The role of Emotions in the explanation of action. *European Review of Philosophy*, to appear.
- Parfit, Derek (1997). Reasons and motivation. *Aristotelian Society Supplementary Volume* 71 (1):99–130.
- Parfit, Derek A. (1984). *Reasons and Persons*. Oxford University Press.
- Peacocke, Christopher (1981). Demonstrative thought and psychological explanation. *Synthese* 49 (2).
- Pears, D. (1984). *Motivated Irrationality*. Oxford University Press.

- Pears, D. (1985). Intention and Belief. In *Essays on Davidson*, ed. B. Vermazen and M. Hintikka. Clarendon.
- Pena-Ruiz, H. (2009). *Grandes Mitos da Cultura*. Lisboa: Temas e Debates.
- Pennock, R. T. (1995). Moral darwinism: Ethical evidence for the descent of man. *Biology and Philosophy*, 10(3).
- Pettigrew, R. (2009). On interpretations of bounded arithmetic and bounded set theory. *Notre Dame Journal of Formal Logic*, 50(2), 141-152.
- Pfeifer, K. (1989). *Actions and Other Events*. New-York: Peter Lang
- Phelan, M. T., & Sarkissian, H. (2008). The folk strike back; or, why you didn't do it intentionally, though it was bad and you knew it. *Philosophical Studies*, 138(2).
- Phelan, M., & Sarkissian, H. (2009). Is the 'trade-off hypothesis' worth trading for? *Mind and Language*, 24(2), 164-180.
- Pierre-Charles, P., Teira, D., & Sophie, J. (2008). Facts, Norms and Expected Utility Functions. *History of the Human Sciences*, 21(2), 45-62.
- Pietroski, Paul (2000). *Causing Actions*, New York: Oxford University Press.
- Platão (1999). *Protágoras*. Tradução Ana da Piedade Elias Pinheiro. Lisboa: Relógio d'Água.
- Pollard, Bill (2005). Naturalizing the space of reasons. *International Journal of Philosophical Studies* 13 (1):69 – 82.
- Proust, J. (1999). Are there nonconceptual representations of action?. Manuscript.
- Proust, J. (1999). Experience, action and theory of mind. *Developmental Science*, 2, 3, 286-287.
- Proust, J. (1999). Indexes for action. *Revue Internationale de Philosophie, Neurosciences*, 3, 321-345.
- Proust, J. (2000). Awareness of agency: Three levels of analysis. In T. Metzinger (ed.), *The Neural Correlates of Consciousness*, Cambridge: Cambridge University Press, 307-324.
- Pust, J. (2000). *Intuitions as Evidence*: Garland Pub.
- Pust, J. (2001). Against explanationist skepticism regarding philosophical intuitions. *Philosophical Studies*, 106(3).
- Qizilbash, Mozaffar (2007). The parity view and intuitions of neutrality. *Economics and Philosophy* 23 (1):107-114.
- Quillen, Keith (1986). Propositional attitudes and psychological explanation. *Mind and Language* 1:133-57.
- Quillen, Keith (1989). Perceptual belief and psychological explanation. *Philosophical Quarterly* 39 (July):276-293.
- Rapaport, E. (1975). Describing moral weakness. *Philosophical Studies*, 28(4), 273-280.
- Ratcliffe, Matthew (2006). "Folk psychology" is not folk psychology. *Phenomenology and the Cognitive Sciences* 5 (1):31-52.
- Rawls, John (1999) *A Theory of Justice*. Cambridge, Massachusetts: Belknap Press of Harvard University Press.

- Raz, J. (1999). Explaining normativity: On rationality and the justification of reason. *Ratio*, 12(4), 354–379.
- Raz, Joseph (2005). The Myth of Instrumental Rationality. *Journal of Ethics and Social Philosophy* 1 (1):28.
- Reis, B. P. W. (1990). Reflexões sobre a Epistemologia de Popper e o Individualismo Metodológico. Rio de Janeiro: IUPERJ. (Série Estudos, nº 77).
- Rena, R. Educational breakthrough in eritrea: Some expectations and outcomes.
- Richman, R. J. (1969). Reasons and causes: Some puzzles. *Australasian Journal of Philosophy*, 47(1), 42 – 50.
- Riggs, W. D. (1997). The weakness of strong justification. *Australasian Journal of Philosophy*, 75(2), 179 – 189.
- Risjord, M. (1999). No strings attached: Functional and intentional action explanations. *Philosophy of Science*, 66(3), 313.
- Rodríguez, A. G. (2008). Rational choice theory and economic laws: The role of shared values. *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 96(1), 191-205.
- Roelofsma, P. H. M. P. (1999). Rational choice theory. *Philosophical Psychology*, 12(3), 309.
- Rorty, A. (1983). Akkratic Believers. *American Philosophical Quarterly*, 20(2), 175-183.
- Ross, Glenn (1982). Knowledge and intentional action. *Philosophical Studies* 41 (2).
- Rosthal, R. (1967). Moral weakness and remorse. *Mind*, 76(304), 576-579.
- Roth, Abraham (2000). 'Reasons Explanation of Actions: Causal, Singular, and Situational', *Philosophy and Phenomenological Research*, 59: 839-74
- Roth, Arbraham (2004). 'Shared Agency and Contralateral Commitments,' *Philosophical Review*, 113 July: 359-410
- Rowbottom, Darrell Patrick (2008). An alternative account of epistemic reasons for action: In response to Booth. *Grazer Philosophische Studien* 76 (1):191-198.
- Rubinstein, A. (2006). *Lecture Notes in Microeconomic Theory: The Economic Agent*. Princeton: Princeton University Press.
- Rundle, B. (1997). *Mind in Action*. Oxford: Clarendon Press.
- Saarinen, R. (1994). *Weakness of the Will in Medieval Thought: From Augustine to Buridan*: E.J. Brill.
- Samuels, Richard & Stich, Stephen P. (2004). Rationality and psychology. In *The Oxford Handbook of Rationality*. Oxford: Oxford University Press.
- Samuels, Richard , Stich, Stephen & Faucher, Luc (2004). Reason and rationality. In M. Sintonen, J. Wolenski & I. Niiniluoto (eds.), *Handbook of Epistemology*. Kluwer.
- Samuels, Richard, Stich, Stephen & Bishop, Michael (2002). Ending the rationality wars: How to make disputes about human rationality disappear. In Renee Elio (ed.), *Common Sense, Reasoning and Rationality*. Oxford University Press.
- Sandven, T. (1995). Intentional action and pure causality: a critical discussion of some central conceptual distinctions in the work of Jon Elster. *Philosophy of the Social Sciences* 25, 3, 286-318.

- Sankey, Howard (1995). The problem of rational theory-choice. *Epistemologia* 18 (2):299-312.
- Santas, G. (1966). Plato's protagoras and explanations of weakness. *Philosophical Review*, 75(1), 3-33.
- Santos, R. (1997). A Explicação Davidsoniana da Acrasia, *Cadernos de Filosofia*, 2, pp. 87-104
- Santos, R. (1998). The Philosophy of Action: An Interview with Alfred R. Mele, *Cadernos de Filosofia*, 4, 1998, pp. 53-84.
- Santos, R. (2006). Acção e Explicação Causal", in *Analyses / Análises*, editado por S. Miguens, J.A. Pinto e C. Mauro, Porto: Faculdade de Letras da Universidade do Porto, pp. 113-122.
- Santos, R. Santos, R. (2007). Escolhas inconsistentes ou compreensões diferentes? Reflexão filosófica sobre um problema de teoria da decisão, *Arquipélago - Filosofia*, 8, pp. 255-272.
- Santos, R. (no prelo). Acrasia. In *Dicionário de Filosofia Moral e Política*, editado por António Marques e Diogo Pires Aurélio.
- Sawyer, R. K. (2003). Nonreductive individualism part II—social causation. *Philosophy of the Social Sciences*, 33(2).
- Sayre-McCord, Geoffrey (1989). Functional explanations and reasons as causes. *Philosophical Perspectives* 3:137-164.
- Sayward, C. (1995/6). Taking Actions Seriously. *Behavior and Philosophy*, 23, 3/24, 1: 51-60
- Scanlon, T. (1998). *What We Owe to Each Other*: Belknap Press of Harvard University Press.
- Scanlon, T. M. (2003). Replies. *Ratio*, 16(4), 424–439.
- Schechtman, Marya (1996). The story of the mind: Psychological and biological explanations of human behavior. *Zygon* 31 (4):597-614.
- Scheer, R. K. (1996). What I will do and what I intend to do. *Philosophy*, 71, 278: 531-40.
- Schroeder, M. (2004). The scope of instrumental reason. *Philosophical Perspectives*, 18(1), 337–364.
- Schroeder, M. (2007). Review of Michael Bratman, *_Structures of Agency: Essays_*. *Notre Dame Philosophical Reviews*, 2007(5).
- Schroeder, M. (2007). *Slaves of the Passions*: Oxford University Press.
- Schroeder, M. (2008). Having reasons. *Philosophical Studies*, 139(1).
- Schroeder, M. (2009). Means-end coherence, stringency, and subjective reasons. *Philosophical Studies*, 143(2).
- Schueler, G. F. (1995). *Desire: Its Role in Practical Reason and the Explanation of Action*. Cambridge, MA: MIT Press.
- Schueler, G. F. (2003). *Reasons and Purposes: Human Rationality and the Teleological Explanation of Action*: Oxford University Press.
- Schwitzgebel, Eric (1996). Theories in children and the rest of us. *Philosophy of Science Association* 3 (3):S202-S210.

- Schwitzgebel, Eric (1997). *Words About Young Minds: The Concepts of Theory, Representation, and Belief in Philosophy and Developmental Psychology*. Dissertation, University of California Berkeley
- Schwitzgebel, Eric (1999). Children's theories and the drive to explain. *Science and Education* 8:457-488.
- Schwitzgebel, Eric (1999). Gradual belief change in children. *Human Development* 42 (6):283-296.
- Scott, M. (1996). Wittgenstein's philosophy of action. *Philosophical Quarterly*, 46, 184: 347-64.
- Scott-Kakures, D. (1996). Self-Deception and Internal Irrationality. *Philosophy and Phenomenological Research*, 56: 31-56.
- Scott-Kakures, D. (2005). Folk Psychology. In (Borchert) *Encyclopedia of Philosophy*. New York: Thomson Gale.
- Searle, John (1979). The intentionality of intention and action. *Inquiry*, 22, 253-280.
- Searle, John (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Searle, John (1990). 'Collective Intentions and Actions,' in P. Cohen, J. Morgan, and M. Pollak (eds.), *Intentions in Communication*, Cambridge, MA: MIT Press
- Searle, John (2003). *Rationality in Action*. MIT Press.
- Searle, John, Barry Smith, Leo Zaibert & Josef Moural (2001). Rationality in action: A symposium. *Philosophical Explorations* 4 (2):66 – 94.
- Sehon, S. (1994). Teleology and the Nature of Mental States. *American Philosophical Quarterly*, 31: 63-72.
- Sehon, S. R. (2000). An argument against the causal theory of action explanation. *Philosophy and Phenomenological Research*, 60(1), 67-85.
- Sellars, W. (1966). Thought and action. In K. Lehrer (ed.), *Freedom and determinism*, . New-York: Random House, pp. 105-139
- Sellars, W. (1973). Action and events. *Nous*, 7, 179-202.
- Sen, A. (1987). *On Ethics and Economics*. Nova York: Blackwell Publishers.
- Setiya, K. (2003). Explaining action. *Philosophical Review*, 112(3), 339-393.
- Setiya, K. (2004). Against internalism. *Nous*, 38(2), 266–298.
- Setiya, K. (2005). Review of 'Weakness of Will and Practical Irrationality' edited by Sarah Stroud and Christine Tappolet. *Philosophical Review*, 114.
- Setiya, K. (2007). Cognitivism about instrumental reason. *Ethics*, 117.
- Setiya, K. (2007). *Reasons without Rationalism*: Princeton University Press.
- Setiya, K. (2009). Reply to Bratman and Smith. *Analysis*, 69(3).
- Setiya, K. (2009). Summary of Reasons without Rationalism. *Analysis*, 69(3).
- Shafer-Landau, R. (2009). A defence of categorical reasons. *Proceedings of the Aristotelian Society (Hardback)*, 109(1pt2), 189-206.
- Shani, I. (2005). Computation and intentionality: A recipe for epistemic impasse. *Minds and Machines*, 15(2), 207-228.

- Sharpe, R. A. (1987). The very idea of a folk psychology. *Inquiry* 30 (December):381-93.
- Sharpe, R. A. (1990). *Making the Human Mind*: Routledge.
- Shepsle, K. A. & bonchek, M. S. (1997). *Analyzing Politics*: Rationality, Behavior, and Institutions. New York: W. W. Norton & Company. Cap. 2: Rationality: The Model of Choice, p. 15-35.
- Silva, M. F. G. da. (1996) Fronteiras da Nova Economia Institucional. São Paulo: NPP-EAESP/FGV. (Série Relatórios de Pesquisa - relatório nº 5).
- Simon, H. A. (1982). *Models of Bounded Rationality vol. I, II*. Cambridge: MIT Press.
- Sinnott-Armstrong, W. (Ed.) (2008). *Moral Psychology*: MIT Press.
- Skipper, R. A. (2004). Calibration of laboratory models in population genetics. *Perspectives on Science*, 12(4).
- Sloane, S. B. (2008). Creative rebellion and moral efficiency as elements of managerial ideology. *Journal of Business Ethics*, 81(3).
- Smith, A. (1975). *Theory of Moral Sentiments*. Oxford: Clarendon Press.
- Smith, M. (1987). The Humean theory of motivation. *Mind*, 96(381), 36-61.
- Smith, M. (1988). On Humeans, anti-Humeans, and motivation: A reply to Pettit. *Mind*, 97(388), 589-595.
- Smith, M. (1994). *The Moral Problem*: Blackwell.
- Smith, M. (1995). Internal Reasons. *Philosophy and Phenomenological Research*, 55: 109-131.
- Smith, M. (1996). Normative reasons and full rationality: Reply to Swanton. *Analysis*, 56(3), 160–168.
- Smith, M. (2003). Humeanism, psychologism, and the normative story. *Philosophy and Phenomenological Research*, 67(2), 460–467.
- Smith, M. (2003). Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion. In S. Stroud & C. Tappolet (Eds.), *Weakness of Will and Practical Irrationality*: Oxford: Clarendon Press.
- Smith, M. (2004). *Ethics and the a Priori: Selected Essays on Moral Psychology and Meta-Ethics*: Cambridge University Press.
- Smith, M. (2007). Is there a nexus between reasons and rationality? *Poznan Studies in the Philosophy of the Sciences and the Humanities*, 94(1), 279-298.
- Smith, M. (2009). Reasons with rationalism after all. *Analysis*, 69(3).
- Smith, M. (2009). The explanatory role of being rational. In D. Sobel & S. Wall (Eds.), *Reasons for Action*: Cambridge University Press.
- Sober, Elliott (1978). Psychologism. *Journal for the Theory of Social Behavior* 8 (July):165-91.
- Son, T. C., Tu, P. H., & Zhang, X. (2005). Reasoning about sensing actions in domains with multi-valued fluents. *Studia Logica*, 79(1).
- Sosa, D. (2006). Scepticism about intuition. *Philosophy*, 81(4), 633-648.
- Sosa, Ernest (2007a). Experimental philosophy and philosophical intuition. *Philosophical Studies* 132 (1):99-107.

- Sosa, Ernest (2007b). Intuitions: Their nature and epistemic efficacy. *Grazer Philosophische Studien* 74 (1):51-67.
- Southwood, N. (2008). Vindicating the normativity of rationality. *Ethics*, 119(1).
- Speight, A. (2002). Arendt and Hegel on the tragic nature of action. *Philosophy and Social Criticism*, 28(5).
- Spohn, W. (1983). Deterministic and probabilistic reasons and causes. *Erkenntnis*, 19(1-3).
- Sripada, Chandra Sekhar (forthcoming). The deep self model and asymmetries in folk judgments about intentional action. *Philosophical Studies*.
- Sripada, C. S. & Stich, Stephen (2004). Evolution, culture, and the irrationality of the emotions. In D. Evans & Pierre Cruse (eds.), *Emotion, Evolution, and Rationality*. Oxford University Press.
- Sripada, C. S. The deep self model and asymmetries in folk judgments about intentional action. *Philosophical Studies*.
- Stampe, D. W. (1990). Desires as reasons--discussion notes on Fred Dretske's explaining behavior: Reasons in a world of causes. *Philosophy and Phenomenological Research*, 50(4), 787-793.
- Stampe, Dennis W. (1990). Desires as reasons--discussion notes on Fred Dretske's explaining behavior: Reasons in a world of causes. *Philosophy and Phenomenological Research* 50 (4):787-793.
- Stein, E. (1996). *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*. Oxford: Clarendon Press.
- Sterelny, Kim (1993). Refuting eliminative materialism on the cheap? *Mind and Language* 8 (2):306-15.
- Sterelny, Kim (1998). Intentional agency and the metarepresentation hypothesis. *Mind and Language* 13 (1):11-28.
- Stewart, J. (2002). Genetics, biology and multifactorial diseases. *Acta Biotheoretica*, 50(4).
- Stich, Stephen, Daniel Kelly, Kevin Haley, Serena Eng & Daniel Fessler (2007). Harm. Affect & the moral / conventional distinction. *Mind & Language* 22 (2):117-131.
- Stich, Stephen & Jonathan M. Weinberg (2001). Jackson's empirical assumptions. *Philosophy and Phenomenological Research* 62 (3):637-643.
- Stich, Stephen, John M. Doris & Erica Roedder (forthcoming). Altruism. In Moral Psychology Research Group (ed.), *The Oxford Handbook of Moral Psychology*. Oxford University Press.
- Stich, Stephen, Nichols, Shaun & Jonathan M. Weinberg (2001). Normativity and epistemic intuitions. *Philosophical Topics*, 29 (1-2):429-460.
- Stich, Stephen, Nichols, Shaun & Jonathan M. Weinberg (2003). Metaskepticism: Meditations in ethno-epistemology. In S. Luper (ed.), *The Sceptics*. Ashgate Publishing.
- Stich, Stephen & Nichols, Shaun (1993). Folk psychology: Simulation or tacit theory? *Mind and Language* 7 (1-2):35-71.
- Stich, Stephen & Nichols, Shaun (1995). Second thoughts on simulation. In Martin Davies & Tony Stone (eds.), *Mental Simulation*. Blackwell.

- Stich, Stephen & Nichols, Shaun (1997). Cognitive penetrability, rationality, and restricted simulation. *Mind and Language* 12 (3-4):297-326.
- Stich, Stephen & Nichols, Shaun (1998). Theory theory to the Max. *Mind and Language* 13 (3):421-449.
- Stich, Stephen & Nichols, Shaun (2000). A cognitive theory of pretense. *Cognition* 74 (2):115-147.
- Stich, Stephen & Nichols, Shaun (2002). Folk psychology. In Stich, Stephen & Ted A. Warfield (eds.), *Blackwell Guide to Philosophy of Mind*. Blackwell.
- Stich, Stephen & Nichols, Shaun (2004). Reading one's own mind: Self-awareness and developmental psychology. In R. Stanton, M. Ezcurdia & C. Viger (eds.), *New Essays in Philosophy of Language and Mind*, Canadian Journal of Philosophy, Supplementary Volume 30. University of Calgary Press.
- Stich, Stephen & R. Ravenscroft (1994). What is folk psychology? *Cognition* 50:447-68.
- Stich, Stephen & Richard E. Nisbett (1980). Justification and the psychology of human reasoning. *Philosophy of Science* 47 (2):188-202.
- Stich, Stephen & Ron Mallon (2000). The odd couple: The compatibility of social construction and evolutionary psychology. *Philosophy of Science* 67 (1):133-154.
- Stich, Stephen & Shaun Nichols (2003) *Mindreading*. Oxford: Oxford University Press.
- Stich, Stephen & Stephen Laurence (1994). Intentionality and naturalism. *Midwest Studies in Philosophy* 19:159-82.
- Stich, Stephen & Ted A. Warfield (1995). Reply to Clark and Smolensky: Do connectionist minds have beliefs? In C. Macdonald & Graham F. Macdonald (eds.), *Connectionism: Debates on Psychological Explanation*. Blackwell.
- Stich, Stephen (1970). Dissonant notes on the theory of reference. *Noûs* 4 (4):385-397.
- Stich, Stephen (1971). What every speaker knows. *Philosophical Review* 80 (4):476-496.
- Stich, Stephen (1972). Grammar, psychology, and indeterminacy. *Journal of Philosophy* 64 (22):799-818.
- Stich, Stephen (1973). What every grammar does: A reply to prof. Arbin. *Philosophia* 3 (1).
- Stich, Stephen (1975). Logical form and natural language. *Philosophical Studies* 28 (6).
- Stich, Stephen (1978). Autonomous psychology and the belief/desire thesis. *The Monist* 61 (October):573-91.
- Stich, Stephen (1978). Beliefs and subdoxastic states. *Philosophy of Science* 45 (December):499-518.
- Stich, Stephen (1978). Empiricism, innateness, and linguistic universals. *Philosophical Studies* 33 (3).
- Stich, Stephen (1978). The recombinant DNA debate. *Philosophy and Public Affairs* 7 (3):187-205.
- Stich, Stephen (1979). Between chomskian rationalism and Popperian empiricism. *British Journal for the Philosophy of Science* 30 (December):329-47.
- Stich, Stephen (1979). Do animals have beliefs? *Australasian Journal of Philosophy* 57 (March):15-28.

- Stich, Stephen (1981). Can Popperians learn to talk? *British Journal for the Philosophy of Science* 32 (2):157-164.
- Stich, Stephen (1982). A philosophical inquiry into ambiguity, vagueness and metaphor in language. *Linguistics and Philosophy* 5 (2).
- Stich, Stephen (1985). Could man be an irrational animal? *Synthese* 64 (1):115-35.
- Stich, Stephen (1988). Reflective equilibrium, analytic epistemology and the problem of cognitive diversity. *Synthese* 74 (3).
- Stich, Stephen (1990). Building belief: Some queries about representation, indication, and function. *Philosophy and Phenomenological Research* 50 (4):801-806.
- Stich, Stephen (1992). What is a theory of mental representation? *Mind* 101 (402):243-61.
- Stich, Stephen (1993). Moral philosophy and mental representation. In R. Michod, L. Nadel & M. Hechter (eds.), *The Origin of Values*. Aldine de Gruyter.
- Stich, Stephen (1993). Naturalizing epistemology: Quine, Simon and the prospects for pragmatism. In C. Hookway & D. Peterson (eds.), *Philosophy and Cognitive Science, Royal Institute of Philosophy, Supplement no. 34*. Cambridge University Press.
- Stich, Stephen (1994). The virtues, challenges and implications of connectionism. *British Journal for the Philosophy of Science* 45 (4).
- Stich, Stephen (1996). Deconstructing the mind. In *Deconstructing the Mind*. Oxford University Press, 1996.
- Stich, Stephen (2001). Plato's method meets cognitive science. *Free Inquiry* 21 (2):36-38.
- Stich, Stephen (2004). Some questions from the not-so-hostile world. *Australasian Journal of Philosophy* 82 (3):503 – 511.
- Stich, Stephen (2006). Is morality an elegant machine or a kludge? *Journal of Cognition and Culture* 6 (1-2):181-189.
- Stich, Stephen (2006). Review of Bishop & Tout, *Epistemology and the psychology of human judgement*. *Mind* 115 (458):390-393.
- Stich, Stephen (2007). Evolution, altruism and cognitive architecture: A critique of Sober and Wilson's argument for psychological altruism. *Biology and Philosophy* 22 (2).
- Stich, Stephen (2008). Some questions about *The evolution of morality*. *Philosophy and Phenomenological Research* 77 (1):228-236.
- Stone, M. Partial order reasoning for a nonmonotonic theory of action.
- Stotz, K. (2007). Experimental philosophy of biology: Notes from the field.
- Stroud, S. (2007). Moral worth and rationality as acting on good reasons. *Philosophical Studies*, 134(3).
- Stroud, S., & Tappolet, C. (2007). *Weakness of Will and Practical Irrationality*. Oxford Clarendon Press.
- Stueber, K. R. (2008). 2. reasons, generalizations, empathy, and narratives: The epistemic structure of action explanation. *History and Theory*, 47(1), 31–43.
- Susko, M. (2003). The fragility of evolution: Part one. *World Futures*, 59(6), 421 – 462.

- Sutton, J. (2005). Moving and Thinking Together in Dance. In R. Grove, K. Stevens & S. McKechnie (Eds.), *Thinking in Four Dimensions: creativity and cognition in contemporary dance*: Melbourne UP.
- Symons, J. (2008). Intuition and philosophical methodology. *Axiomathes*, 18(1).
- Sytsma, J. (2007). The proper province of philosophy: Ordinary language meets experimental philosophy.
- Sytsma, J. (2008). Phenomenological obviousness and the new science of consciousness.
- Sytsma, Justin & Machery, Edouard (2009). How to study folk intuitions about phenomenal consciousness. *Philosophical Psychology* 22 (1):21 – 35.
- Tang, L. (2008). An integral model of collective action in organizations and beyond. *Journal of Business Ethics*, 80(2).
- Tappolet, C. (2003). Emotions and the intelligibility of akratic action *Weakness of Will and Practical Irrationality*: Oxford: Clarendon Press.
- Taylor, Charles (1964). *The Explanation of Behavior*, London: Routledge & Kegan Paul.
- Taylor, C.C.W. (1980). Plato, Hare and Davidson on Akrasia. *Mind*, Vol. LXXXIX, 499-518.
- Taylor, M. (1996). Good government: On hierarchy, social capital, and the limitations of rational choice theory. *Journal of Political Philosophy*, 4(1), 1–28.
- Teigen, Karl Halvor (2000). Intuitive versus analytic abilities: The case of words versus numbers. *Behavioral and Brain Sciences* 23 (5):698-699.
- Teira, D. (2006). On the Normative Dimension of St. Petersburg Paradox. *Studies in History and Philosophy of Science*, 37, 210-223.
- Tenenbaum, S. (1999). The judgment of a weak will. *Philosophy and Phenomenological Research*, 59(4), 875-911.
- Tersman, Folke (2008). The reliability of moral intuitions: A challenge from neuroscience. *Australasian Journal of Philosophy* 86 (3):389 – 405.
- Thalberg, I. (1984). Do our intentions cause our intentional actions? *American Philosophical Quarterly* 21: 249-260.
- Thalberg, I. (1977). *Perception, emotion and action: A component approach*. Oxford: Blackwell.
- Thomas, L. (2003). Equality and the mantra of diversity.
- Thomson, J. (1971). Individuating actions. *Journal of Philosophy*, 68, 774-81.
- Thornton, T. (1997). Reasons and causes in philosophy and psychopathology. *Philosophy, Psychiatry, and Psychology*, 4(4), 307-317.
- Tilley, J. J. (2004). Justifying reasons, motivating reasons, and agent relativism in ethics. *Philosophical Studies*, 118(3).
- Tollefsen, Christopher (2006). Reasons for action and reasons for belief. *Social Epistemology* 20 (1):55 – 65.
- Tomasello, Michael & Rakoczy, Hannes (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language* 18 (2):121-147.
- Treisman, Michel (1962). Psychological explanation: The 'private data' hypothesis. *British Journal for the Philosophy of Science* 13 (August):130-143.

- Tuomela, R. (1977). *Human Action and its Explanation*, Dordrecht, D. Reidel.
- Tuomela, R. (1989). Collective action, supervenience, and constitution. *Synthese*, 80(2).
- Tuomela, R., (2003). 'The We-Mode and the I-Mode,' in Frederick Schmitt (ed.), *Socializing Metaphysics — the Nature of Social Reality*. Lanham, MD: Rowman & Littlefield: 93-127
- Turner, Jason (2004). Folk intuitions, asymmetry, and intentional side effects. *Journal of Theoretical and Philosophical Psychology* 24:214-219.
- Turri, J. (2009). The ontology of epistemic reasons. *Nous*, 43(3), 490-512.
- Vailati, E. (1990). Leibniz on Locke on weakness of will. *Journal of the History of Philosophy*, 28(2).
- Varian, Hal R (1999). *Intermediate Microeconomics*. 5th Edition, New York: Norton.
- Varoufakis, Y. (1997). Moral rhetoric in the face of strategic weakness: Empirical clues for an ancient puzzle. *Erkenntnis*, 46(1).
- Velleman, J. D. (1996). The Possibility of Practical Reason. *Ethics*, 106: 694-726
- Velleman, J. D. (1997). How to share an intention. *Philosophical and Phenomenological Research* 57: 29-50.
- Velleman, J. David (1989). *Practical Reflection*, Princeton: Princeton University Press.
- Velleman, J. David (2000). *The Possibility of Practical Reason*, Oxford: Oxford University Press
- Verbeek, B. (2008). Consequentialism and rational choice: Lessons from the Allais paradox. *Pacific Philosophical Quarterly*, 89(1), 86–116.
- Vlastos, G. (1969). Reasons and causes in the phaedo. *Philosophical Review*, 78(3), 291-325.
- Vogler, C. A. (2002). *Reasonably Vicious*: Harvard University Press.
- Vollmer, Fred (1993). Intentional action and unconscious reasons. *Journal for the Theory of Social Behaviour* 23 (3):315-326.
- Von Eckardt, Barbara (2004). Folk Psychology as Theory. *Stanford Encyclopedia of Philosophy*.
- Walker, A. F. (1989). The problem of weakness of will. *Nous*, 23(5), 653-676.
- Wall, D. (2009). Akrasia and self-control. *Philosophical Explorations*, 12(1), 69 – 78.
- Wallace, R. J. (1999). Addiction as defect of the will: Some philosophical reflections. *Law and Philosophy*, 18(6), 621–654.
- Watson, G. (1977). Skepticism about weakness of will. *Philosophical Review*, 86(3), 316-339.
- Watson, Gary. (2004). *Agency and Answerability: Selected Essays*, Oxford: Oxford University Press.
- Weatherston, B. (2003). What good are counterexamples? *Philosophical Studies*, 115(1), 1-31.
- Wedgwood, R. (2002). Practical reason and desire. *Australasian Journal of Philosophy*, 80(3), 345 – 358.

- Wedgwood, R. (2009). Democracies, human rights, and collective action. *Ethics and International Affairs*, 23(1), 27-37.
- Weinberg, J. M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy*, 31(1), 318–343.
- Weinberg, J. M. (2009). On doing better, experimental-style. *Philosophical Studies*, 145(3).
- Weinberg, Jonathan M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy* 31 (1):318–343.
- Weinberg, Jonathan M. (2008). Naturalism and intuitions: Commentary on Steven Hales, relativism and the foundations of philosophy. *International Journal of Philosophical Studies* 16 (2):263 – 270.
- Weinberg, Jonathan M. (2009). On doing better, experimental-style. *Philosophical Studies* 145 (3).
- Wiener, P. P. (1932). The experimental philosophy of Robert Boyle (1626-91). *Philosophical Review*, 41(6), 594-609.
- Wilding, A. (2005). Why we don't remain in the provinces. *Philosophy and Social Criticism*, 31(1).
- Wilkes, Kathleen V. (1991). The relationship between scientific psychology and common-sense psychology. *Synthese* 89 (October):15-39.
- Williamson, T. (2004). Philosophical 'intuitions' and scepticism about judgement. *Dialectica*, 58(1), 109–153.
- Wilson, George (1989). *The Intentionality of Human Action*, Stanford, CA: Stanford University Press
- Wilson, George (2000). 'Proximal Practical Foresight', *Philosophical Studies*, 99: 3-19
- Wolff, E. (2006). From phenomenology to critical theory: The genesis of adorno's critical theory from his reading of Husserl. *Philosophy and Social Criticism*, 32(5).
- Wright, C. (1995). Can there be a rationally compelling argument for anti-realism about ordinary ("folk") psychology? *Philosophical Issues* 6:197-221.
- Wright, Cory & Bechtel, William P. (2006). Mechanisms and psychological explanation. In Paul Thagard (ed.), *Philosophy of Psychology and Cognitive Science*. Elsevier.
- Wright, J. C., & Bengson, J. (2009). Asymmetries in judgments of responsibility and intentional action. *Mind and Language*, 24(1), 24-50.
- Wright, W. (2002). Fodor's epistemic intuitions of analyticity. *Sorites*, 14(October), 110-116.
- Wringe, B. (2006). Collective action and the peculiar evil of genocide. *Metaphilosophy*, 37(3-4), 376–392.
- Zaragoza, K. (2006). What happens when someone acts compulsively? *Philosophical Studies*, 131(2).
- Zilhão, A. (2007). Heurísticas Rápidas e Frugais, Encontro de Probabilidades e Incontinência. *Cadernos de Filosofia – Revista Internacional de Filosofia*, 21, pp. 43-65.
- Zilhão, A. (2009). Efeito dos Custos Afundados, incontinência e racionalidade. In. Mauro, C., Miguens, S., Cadilha, S. (Eds) *Mente Linguagem e Acção: textos para discussão*. Colecção Campo da Filosofia. Porto: Campo das Letras.