



TRADUÇÃO

“Rational fools”, de Amartya Sen: uma tradução comentada

Amartya Sen's "Rational Fools": A commented translation

"Rational fools", de Amartya Sen: una traducción anotada

Fabrizio Pontin¹

orcid.org/0000-0002-3984-1849

fabrizio.pontin@unilasalle.edu.br

Rafaela Lemos da Silva¹

orcid.org/0000-0002-0135-5327

rafaela.silva0157@unilasalle.edu.br

Recebido em: 22 ago. 2024

Aprovado em: 6 ago. 2025.

Publicado em: 13 nov. 2025.

Resumo: Este artigo apresenta uma tradução comentada da obra seminal de Amartya Sen, “RATIONAL FOOLS: A CRITIQUE OF THE BEHAVIORAL FOUNDATIONS OF ECONOMIC THEORY” (1977), traduzida para o português como “OTÁRIOS RACIONAIS: UMA CRÍTICA DOS FUNDAMENTOS COMPORTAMENTAIS DA TEORIA ECONÔMICA”. O trabalho de Sen é descrito como um manifesto que questiona o conceito de *homo economicus* — um agente que é exclusivamente motivado por autointeresse — e critica as fundações da teoria econômica dominante. A introdução do artigo contextualiza o impacto revolucionário de Sen, explorando sua crítica à teoria das preferências, seu questionamento do egocentrismo metodológico e sua desconstrução do problema da transitividade. O texto também discute as opções de tradução adotadas, com ênfase na escolha de “otários racionais” para o título, que, segundo os tradutores, melhor representa a sátira de Sen sobre a descrição limitada do ator racional na teoria econômica clássica. A escolha do termo “otário” é justificada por descrever uma pessoa deliberadamente desagradável e planejada, em oposição a termos como “tolo” ou “bobo”, que sugerem ingenuidade ou um comportamento irrefletido. Outra decisão crucial de tradução foi a diferenciação entre “simpatia” (*sympathy*) e “compromisso” (*commitment*), que são conceitos centrais na crítica de Sen. Os tradutores optaram por “compromisso” para capturar a ideia de um indivíduo que age a partir de um senso de responsabilidade social e não de uma obrigação formal, frequentemente influenciado pelo receio de retaliação social. O artigo argumenta que a visão de Sen leva a uma concepção de racionalidade mais rica e humana, reintegrando a ética, o contexto e a psicologia na análise econômica e política.

Palavras-chave: ação racional; comportamento; simpatia; empatia; economia comportamental.

Abstract: This article presents a commented translation of Amartya Sen's seminal work, “RATIONAL FOOLS: A CRITIQUE OF THE BEHAVIORAL FOUNDATIONS OF ECONOMIC THEORY” (1977), translated into Portuguese as “OTÁRIOS RACIONAIS: UMA CRÍTICA DOS FUNDAMENTOS COMPORTAMENTAIS DA TEORIA ECONÔMICA”. Sen's work is described as a manifesto that challenges the concept of the *homo economicus*—an agent motivated solely by self-interest—and critiques the foundations of dominant economic theory. The article's introduction contextualizes Sen's revolutionary impact, exploring his critique of preference theory, his questioning of methodological egocentrism, and his deconstruction of the problem of transitivity. The text also discusses the translation choices made, with an emphasis on the selection of “otários racionais” for the title, which, according to the translators, better represents Sen's satire on the limited description of the rational actor in classical economic theory. The choice of the term “otário” is justified as describing a person who is deliberately unpleasant and calculating, as opposed to terms like “tolo” or “bobo,” which suggest naiveté or an unthinking behavior. Another crucial translation decision was the differentiation between “simpatia” (*sympathy*) and “compromisso” (*commitment*), which are central concepts in Sen's critique. The translators chose “comprometimento” to capture the idea of an individual acting out of a sense of social responsibility rather than a formal obligation, often influenced by the fear of social retaliation. The article argues



¹ Universidade La Salle, Canoas, Rio Grande do Sul, Brasil.

that Sen's view leads to a richer and more human conception of rationality, reintegrating ethics, context, and psychology into economic and political analysis.

Keywords: Rational Action; Behavior; Sympathy; Empathy; Behavioral Economics.

Resumen: Este artículo presenta una traducción comentada de la obra seminal de Amartya Sen, "RATIONAL FOOLS: A CRITIQUE OF THE BEHAVIORAL FOUNDATIONS OF ECONOMIC THEORY" (1977), traducida al portugués como "OTÁRIOS RACIONAIS: UMA CRÍTICA DOS FUNDAMENTOS COMPORTAMENTAIS DA TEORIA ECONÔMICA". El trabajo de Sen se describe como un manifiesto que cuestiona el concepto de *homo economicus*—un agente exclusivamente motivado por el autointerés—y critica los cimientos de la teoría económica dominante. La introducción del artículo contextualiza el impacto revolucionario de Sen, explorando su crítica a la teoría de las preferencias, su cuestionamiento del egocentrismo metodológico y su deconstrucción del problema de la transitividad. El texto también discute las opciones de traducción adoptadas, con énfasis en la elección de "otários racionais" para el título, que, según los traductores, representa mejor la sátira de Sen sobre la limitada descripción del actor racional en la teoría económica clásica. La elección del término "otário" se justifica por describir a una persona deliberadamente desagradable y planeada, en oposición a términos como "tolo" o "bobo", que sugieren ingenuidad o un comportamiento irreflexivo. Otra decisión crucial de traducción fue la diferenciación entre "simpatía" (*sympathy*) y "compromiso" (*commitment*), que son conceptos centrales en la crítica de Sen. Los traductores optaron por "comprometimiento" para capturar la idea de un individuo que actúa desde un sentido de responsabilidad social y no desde una obligación formal, a menudo influenciado por el temor a la retaliación social. El artículo argumenta que la visión de Sen conduce a una concepción de la racionalidad más rica y humana, reintegrando la ética, el contexto y la psicología en el análisis económico y político.

Palabras clave: acción racional; comportamiento; simpatía; empatía; economía del comportamiento.

1 Introdução da tradução: para além dos otários racionais: contexto e impacto de "Rational Fools" na teoria da preferência (e algumas notas sobre opções na tradução)

A publicação de "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory", em 1977, por Amartya Sen, não é apenas um artigo acadêmico; é um manifesto. Este trabalho, que se consagraria como um dos mais citados e influentes na história das ciências sociais, e com impacto profundo e duradouro na ciência política e na filosofia, abalou as fundações do paradigma econômico dominante. Em um momento em que a teoria se consolidava em torno da figura do

homo economicus — um agente movido por um autointeresse calculista e implacável —, Sen lança um desafio fundamental: ele questiona o quão "racional" pode ser um agente desprovido de simpatia, compromisso, moralidade e, em última análise, de humanidade. O presente texto, que serve como introdução à tradução brasileira desta obra seminal, busca contextualizar o impacto revolucionário de Sen, destacando seu debate com a teoria das preferências, sua crítica aos modelos comportamentais, a desconstrução do problema da transitividade, o questionamento do egocentrismo metodológico e a centralidade do conceito de "simpatia", elementos já explorados com acuidade nos textos que servem de base para esta análise.

O cenário intelectual que Sen se propõe a desconstruir tem como epicentro a obra de Kenneth Arrow, *Social Choice and Individual Values* (1951). Fortemente influenciado por Vilfredo Pareto, Arrow buscou formalizar a passagem das preferências individuais para uma escolha social coerente, uma "função de bem-estar social". Para isso, ele estabeleceu um modelo de racionalidade individual baseado em premissas de elegância lógica, mas de grande restrição empírica. Um agente racional, para Arrow, é aquele capaz de ordenar suas preferências de forma consistente, obedecendo a dois axiomas principais:

1. **conectividade (ou completude):** para quaisquer duas opções x e y em um conjunto de escolha, o agente pode afirmar que prefere x a y (xPy), prefere y a x (yPx), ou é indiferente entre elas (xly). Formalmente, para todos os objetos de escolha x e y , ou xRy ou yRx (onde R denota a relação "preferível ou indiferente a");
2. **transitividade:** se o agente prefere (ou é indiferente a) x em relação a y , e prefere (ou é indiferente a) y em relação a z , então ele necessariamente prefere (ou é indiferente a) x em relação a z . Formalmente, para todo x, y, z , se xRy e yRz , então xRz .

Nesse arcabouço, a motivação por trás da preferência é um detalhe secundário. O que define a

racionalidade é a consistência interna da escolha observada, uma noção que Arrow considerava muito fraca, mas pragmaticamente superior à hipercontextualização exigida por pensadores como John Stuart Mill, que vinculava a escolha a "hábitos de autoconsciência e auto-observação" (Rawls, 2008). Para Arrow, se um agente declara preferir laranjas a abacaxis e, na prática, escolhe a laranja, ele é racional. A escolha material revela a preferência declarada. Este é o cerne do egocentrismo metodológico: o foco no comportamento observável do indivíduo autointeressado como a única fonte de dados relevante.

É precisamente essa base que Sen ataca. Ele argumenta que essa visão não é apenas uma simplificação, mas uma distorção da natureza humana, e é a partir dessa crítica que os temas centrais de sua obra emergem.

A crítica ao egocentrismo e a introdução da simpatia e do compromisso

O ataque mais direto de Sen em "Rational Fools" é à premissa do egoísmo universal. Reduzir todas as ações humanas a uma busca pela maximização do bem-estar pessoal é, para ele, ignorar as motivações éticas que estruturam a vida em sociedade. Para refinar a análise, Sen introduz a distinção crucial entre simpatia e compromisso.

Sen tem significativa conexão com o consequencialismo: em diferentes momentos, destaca-se a relevância da análise de consequências para a compreensão de processos de escolha. Em outros termos, e principalmente quando disserta a respeito do comprometimento, a avaliação de possíveis ou esperadas consequências faz-se importante, dado que estas podem influenciar o comportamento de um indivíduo, seja negativa ou positivamente. É contemplada a ideia de que, por exemplo, a antecipação de uma mudança de posição social ou de uma influência sobre sua reputação, como consequência de uma ação, pode influenciar o comportamento da pessoa, levando-a a tomar uma decisão que talvez não seja a alternativa que lhe trará maiores ganhos pessoais, porém proporcionará melhor enqua-

ramento dentro de um grupo social. Assim, trazendo a avaliação de consequências ao cálculo inicial, tem-se certa influência do utilitarismo de resultado dentro da abordagem do ator racional para Sen, que não a descarta — pelo contrário, inclusive, assegurando-a como bastante relevante —, mas não a coloca como premissa principal; ou seja, Sen não considera este como sendo o fator mais influente sobre o comportamento de um indivíduo.

Quando Sen apresenta os conceitos de "simpatia" e "comprometimento" (*sympathy* e *commitment*, originalmente) como variáveis relevantes para a interpretação do comportamento humano, ressalta a importância da distinção entre eles, como será demonstrado posteriormente. Consequentemente, faz-se importante a razão para a escolha do termo "comprometimento" em vez de "compromisso": dentro da dissertação de Sen, ele utiliza o "comprometimento" como uma postura na qual o indivíduo é influenciado a tomar uma decisão com base em sua visão de responsabilidade social. Não é, simplesmente, uma questão de o indivíduo fazer sua escolha guiado por uma *obrigação* social propriamente dita, mas, sim, por *sentir-se* pressionado a seguir um comportamento que é socialmente melhor aceito ou mais esperado. Deste modo, o compromisso traz uma ideia de resposta direta a uma força maior, enquanto o comprometimento descreve melhor a relação de *senso* de dever e receio de retaliação social (em consequência).

De fato, uma das grandes críticas de Sen à premissa básica da teoria econômica tradicional, que pressupõe que todo indivíduo age com base em interesse próprio, é, justamente, a ideia de motivação única para o comportamento humano (no caso, o egoísmo). Ao longo do seu trabalho, Sen disserta sobre diferentes sentimentos, conceitos e visões que podem influenciar indivíduos em suas decisões, demonstrando que muito se perde ao limitar o homem como um ser racional puramente autointeressado, baseado em observação e fundamentos comportamentais.

O problema da indiferença e o paradoxo do liberal paretiano

A desconstrução de Sen estende-se também à noção de “indiferença”. Na teoria de Arrow, se um agente se declara indiferente entre as opções A e B (AIB), isso implica que as ordenações AB e BA são perfeitamente intercambiáveis para ele. Sen questiona essa intercambialidade automática. Uma declaração de indiferença pode, na verdade, mascarar uma dupla rejeição (“não quero nem A, nem B”). Nesse caso, a efetivação de qualquer uma das opções não é, de modo algum, indiferente para o agente, que pode ter uma forte preferência por uma delas ser *menos prejudicial* que a outra, mesmo que ambas sejam indesejáveis.

Essas críticas convergem em sua demonstração mais célebre, apresentada em “The Impossibility of a Paretian Liberal” (1970). Sen mostra que mesmo um critério liberal mínimo — a Condição L, que postula que cada indivíduo deve ter uma esfera privada sobre a qual sua preferência é decisiva (por exemplo, escolher a cor das paredes de seu próprio quarto) — pode entrar em conflito direto com o Princípio de Pareto (a Condição P), que afirma que, se todos na sociedade preferem uma opção a outra, a sociedade também deveria preferi-la.

O paradoxo de Arrow, para Sen, não indica a impossibilidade da democracia, mas sim a impossibilidade de um *liberal paretiano*. O problema não está no liberalismo, mas nas premissas comportamentais empobrecidas do modelo paretiano, especialmente na Condição U (domínio irrestrito). Essa condição trata qualquer preferência logicamente coerente — inclusive as mais antissociais, como a de um puritano que prefere que outra pessoa não leia um livro que ele considera lascivo, mesmo que isso vá contra a vontade da própria pessoa — como igualmente válida na agregação social. Para Sen, e mais tarde para Martha Nussbaum, com sua abordagem das capacidades, certas titularidades e direitos são inalienáveis e não podem depender da flutuação de preferências em um “mercado” de escolhas sociais. A justiça exige um domínio restrito, orien-

tado por valores fundamentais que não podem ser objeto de negociação.

A gramática das preferências e a dificuldade na confluência entre comportamento e preferência

No entanto, para Sen, a adição do contexto à função do processo de escolha é muitíssimo relevante, de modo que sugere que a observação crua do comportamento do indivíduo não é suficiente para fazer um ordenamento de suas preferências — e, mais do que isso, sugere que um ordenamento único também não é suficiente para representar a relação de preferências e escolhas. Apesar de ter considerável relação com o viés do utilitarismo devido a sua proximidade com o consequencialismo, Sen não se atém à visão puramente utilitarista, entendendo que uma escolha de um indivíduo pode, sim, maximizar sua utilidade, dando a entender que o objetivo era de fato baseado em resultado, no entanto, Sen argumenta que esse resultado pode ser, de certa forma, uma coincidência: o processo que levou o indivíduo àquela escolha é relevante, pois pode demonstrar que sua intenção não era potencializar seus ganhos, mas pode ter sido baseado em um comportamento que, quando contextualizado, demonstra que a intenção era influenciada na relação do indivíduo com a sociedade ou grupo em que está inserido. Essas possíveis influências, segundo Sen, podem ser causadas por sentimentos de simpatia e comprometimento.

A crítica de Sen ao modelo comportamental vai além da introdução de novas motivações. Ela atinge o próprio núcleo da relação entre escolha e preferência, questionando a “gramática” que a teoria econômica utiliza para descrever a agência humana. A abordagem da “preferência revelada”, popularizada por Paul Samuelson e adotada implicitamente por Arrow, assume que as escolhas observáveis são a única fonte confiável de informação sobre as preferências de um indivíduo. Se uma pessoa escolhe ‘X’ em vez de ‘Y’, quando ambos estão disponíveis, a teoria infere que ela prefere ‘X’ a ‘Y’. Para Sen, essa

confluência é problemática e, em muitos casos, simplesmente falsa.

Em "Behaviour and the Concept of Preference" (1973), Sen desdobra essa crítica, argumentando que a escolha é um ato, enquanto a preferência é uma relação de ordenação mental. A lacuna entre os dois é preenchida por uma miríade de fatores que o modelo padrão ignora. A escolha pode ser influenciada por códigos sociais, senso de dever, medo de reprovação, ou mesmo pela falta de alternativas desejáveis. Uma pessoa pode escolher um prato menos saboroso em um jantar por educação (compromisso com uma norma social), ou um candidato em quem não acredita plenamente para evitar um mal maior (uma escolha estratégica, não uma preferência genuína). Em ambos os casos, o comportamento observável não "revela" a preferência preferida do agente, mas, sim, uma decisão complexa, mediada pelo contexto.

Da indiferença à imparcialidade: o legado de "Otários Racionais" para a teoria da justiça

O trabalho de Sen nos anos 70 sobre a impossibilidade da indiferença neutra e a necessidade de considerar o contexto lança as bases para sua obra tardia sobre justiça, notadamente em *A Ideia de Justiça* (2011). A crítica à indiferença na teoria da preferência ecoa em sua crítica à imparcialidade fechada na teoria da justiça, como a encontrada na posição original de John Rawls.

Para Sen, a imparcialidade alcançada sob o "véu da ignorância" de Rawls é "fechada" porque limita o escopo da argumentação moral aos membros de uma comunidade política específica, ignorando as perspectivas e os interesses de "espectadores imparciais" externos. Assim como a indiferença não é uma posição neutra, a imparcialidade não pode ser alcançada em um vácuo. Sen defende uma imparcialidade aberta, inspirada em Adam Smith, que exige que consideremos os julgamentos de outros, próximos e distantes, para evitar o paroquialismo e o autoengano.

A conexão é profunda: assim como a regra do atrito não pode ser ignorada por um engenhei-

ro ao projetar um prédio real, as "fricções" do mundo — as peculiaridades locais, as tensões culturais e as assimetrias de poder — não podem ser ignoradas na formulação de princípios de justiça. A presunção de uma indiferença ou imparcialidade total, sem posicionamento, é uma ficção perigosa. A cidadania, para Sen, pressupõe o reconhecimento da impossibilidade de uma indiferença não posicionada.

Ao desafiar os pilares do egocentrismo metodológico, da equivalência simplista entre escolha e preferência, e da transitividade acontextual, Amartya Sen não apenas desmonta a figura do "otário racional", mas abre caminho para uma concepção de racionalidade mais rica, complexa e, em última análise, mais humana. Ele reintegra a ética, o contexto e a psicologia no coração da análise econômica e política. A leitura de "Rational Fools" é, portanto, um convite para reconhecer que a racionalidade não se esgota na consistência lógica de um agente isolado; ela floresce na interação social, na capacidade de se comprometer com valores e na sensibilidade para com o bem-estar dos outros. É um passo indispensável para quem busca compreender não apenas as bases de uma sociedade justa, mas o próprio significado da escolha humana.

Ainda, finalmente, é relevante dar ênfase à escolha do termo "otários racionais" (*rational fools*, originalmente). Quando Sen faz sua descrição a respeito do que seria o *rational fool*, não há intenção de ser, de certa forma, agressivo ao retratar o indivíduo como carente de discernimento; na realidade, ao apresentar a ideia, Sen faz uma crítica ao que é esperado de um indivíduo sob luz da teoria tradicional econômica, de modo que, muito elegantemente, satiriza a forma limitada e engessada como teóricos clássicos descrevem um ator racional. O termo "otário" nos parece representar melhor o tipo de personalidade que Sen descreve no artigo — na discussão com colegas, cogitamos o uso de "tolos" ou "bobos" racionais, mas os termos comunicam uma ingenuidade, ou falta de discernimento, que Sen não comunica ao descrever o ator racional como alguém capaz apenas de pensar em si mesmo, sem conside-

ração pelos outros — na passagem clássica do artigo, essa pessoa não é uma “boba”, ou uma “tola”, mas, de fato, uma “otária” —, uma pessoa extremamente desagradável ao convívio, e que é desagradável de forma deliberada e planejada.

Os tradutores gostariam de agradecer contribuições importantes para esta tradução por parte de Julia Leuopolt Campos, Gabriel Fernandes Mafioleti, Nyara Rosana K. de Araújo, Neuro Zambam, Johannes Servan, Carola Freiin von Villiez e, especialmente, Flávio Comim.

Referências

- ARROW, K. J. *Social Choice and Individual Values*. 2. ed. New Haven: Yale University Press, 1963.
- ARROW, K. J. *Social choice and individual values*. Chicago: Cowles Commission for Economics Research, 1951.
- ARROW, K. J. Arrow's theorem. In: EATWELL, J.; MILGATE, M.; NEWMAN, P. (org.). *The New Palgrave: A Dictionary of Economics*. London: Palgrave Macmillan, 1984. p. 26-27.
- BRINK, D. O. *Mill's progressive principles*. Oxford: Clarendon Press, 2015.
- BRINK, D. O. Mill on Justice and Rights. In: MACLEOD, C. (org.). *A companion to Mill*. Hoboken: Wiley Blackwell, 2017.
- CALABRESI, G.; MELAMED, A. D. Property Rules, Liability Rules, and Inalienability: One View of the Cathedral. *Harvard Law Review*, [s. l.], v. 85, n. 6, p. 1089-1128, 1972.
- COMIM, F.; NUSSBAUM, M. C. (org.). *Capabilities, gender, equality: towards fundamental entitlements*. Cambridge: Cambridge University Press, 2014.
- MILL, J. S.; BENTHAM, J.; AUSTIN, J. *Utilitarianism; and, On liberty: including Mill's "Essay on Bentham" and selections from the writings of Jeremy Bentham and John Austin*. Malden: Blackwell, 2003.
- NUSSBAUM, M. C. Capabilities As Fundamental Entitlements: Sen And Social Justice. *Feminist Economics*, [s. l.], v. 9, n. 2-3, p. 33-59, 2003.
- NUSSBAUM, M. C. Poverty and Human Functioning: Capabilities as Fundamental Entitlements. In: NUSSBAUM, M. C. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge: Belknap Press, 2006.
- NUSSBAUM, M. C. *Frontiers of justice: disability, nationality, species membership*. Cambridge: Belknap, 2007.
- PARETO, V.; FINER, S. E. (org.). *Vilfredo Pareto: sociological writings*. London: Pall Mall, 1966.
- PONTIN, F. Economia Política e as raízes da posição original em Rawls. *DoisPontos*, Curitiba, v. 10, n. 1, p. 201-224, 2013.
- PONTIN, F. From the Paradox of Voting to a General Paradox of Choice. In: SILVA, D. M. da et al. (org.). *Contratualismo moral e político na contemporaneidade: Anais do Primeiro Congresso Internacional*. Seropédica: UFRRJ, 2014. p. 27-41.
- RAWLS, J. *A theory of justice*. Cambridge: Belknap, 1999.
- RAWLS, J. *Political liberalism*. New York: Columbia University Press, 2005.
- RAWLS, J.; FREEMAN, S. R. (org.). *Lectures on the history of political philosophy*. Cambridge: Belknap Press of Harvard University Press, 2008.
- RÖSSLER, B.; BRINK, B. van D.; OWEN, D. Work, Recognition, Emancipation. *Recognition and Power*, [s. l.], p. 135-163, 2005.
- SAMUELSON, P. A. A Note on The Pure Theory of Consumer's Behaviour. *Economica*, New Series, v. 5, n. 17, p. 61-71, 1938.
- SEN, A. Quasi-Transitivity, Rational Choice and Collective Decisions. *The Review of Economic Studies*, [s. l.], v. 36, n. 3, p. 381-393, 1969.
- SEN, A. The Impossibility of a Paretian Liberal. *Journal of Political Economy*, [s. l.], v. 78, n. 1, p. 152-157, 1970.
- SEN, A. *On Economic Inequality*. Oxford: Clarendon Press, 1973.
- SEN, A. Behaviour and the Concept of Preference. *Economica*, New Series, v. 40, n. 159, p. 241-259, 1973.
- SEN, A. Choice, Ordering and Morality. In: KÖRNER, S. (org.). *Practical reason*. Oxford: Blackwell Basil, 1974.
- SEN, A. Rationality and morality: A reply. *Erkenntnis*, [s. l.], v. 11, n. 1, p. 225-232, 1977.
- SEN, A. Maximization and The Act of Choice. *Econometrica*, [s. l.], v. 65, n. 4, p. 745-779, 1997.
- SEN, A. *A ideia de Justiça*. São Paulo: Companhia das Letras, 2011.
- SUNSTEIN, C. R. The Ethics of Nudging. *Yale Journal on Regulation*, [s. l.], v. 32, n. 2, 2015.
- SUZUMURA, K. Rational choice and revealed preference. In: SUZUMURA, K. *Rational choice, collective decisions, and social welfare*. Cambridge: Cambridge University Press, 1983. p. 18-61.

2 Tradução:

Otários Racionais: uma crítica dos fundamentos comportamentais da teoria econômica

Em *Mathematical Psychics* (Psíquica Matemática), publicado em 1881, Edgeworth afirmou que "o primeiro princípio de Economia é que todo agente é movido somente pelo próprio interesse"²³. Essa visão dos indivíduos tem sido persistente nos modelos econômicos, e a natureza da teoria econômica parece ter sido muito influenciada por essa premissa básica. Neste trabalho, gostaria de dissertar sobre alguns dos problemas que emergiram dessa noção sobre os seres humanos.

Deveria mencionar que Edgeworth estava ciente de que esse primeiro princípio da Economia não era, particularmente, realista. Além disso, ele sentia que "o homem dito do século XIX é, em grande parte, um egoísta impuro, um utilitarista misto"⁴. Isso reforça o interessante questionamento sobre por que Edgeworth utilizou tanto do seu tempo e de seu talento para desenvolver uma linha de investigação sobre o que ele próprio acreditava ser falso. A questão não é por que abstrações deveriam ser consideradas em questões sobre economia em geral — a natureza da investigação torna isso inevitável —, mas por que alguém optaria por uma suposição que ele mesmo acredita ser não somente imprecisa em detalhes, mas fundamentalmente equivocada? Como veremos, essa questão é, também, de contínuo interesse para a economia moderna.

Parte da resposta, no que tange a Edgeworth, sem dúvida, reside no fato de que ele não assumia que a suposição era fundamentalmente equivocada a respeito dos tipos *específicos* de atividades cujas ele aplicava o que chamava de "cálculo econômico": (i) guerra e (ii) contrato.

"Admitindo que existem, nas partes superiores da natureza humana, uma tendência à e um sentimento por instituições utilitaristas", ele fez a seguinte pergunta retórica: "poderíamos seriamente supor que essas considerações morais foram relevantes para a guerra e o comércio; poderíamos erradicar o cerne incontrolável do egoísmo humano, ou exercitar uma força notável em comparação com o impulso do interesse próprio"⁵. Ele interpretou que Sidgwick dissipou a "ilusão" de que "o interesse de todos é o interesse de cada um", apontando que Sidgwick identificou os "dois princípios supremos, Egoísmo e Utilitarismo", como sendo "irreconciliáveis, exceto se por religião". "Não faz parte do espírito da filosofia do prazer depreciar a importância da religião", afirmou Edgeworth, "mas no presente ensaio, e lidando com elementos inferiores da natureza humana, precisaríamos buscar uma transição mais óbvia, uma passagem mais terrena, do princípio do autointeresse para o princípio, ou ao menos a prática, do utilitarismo"⁶.

Observe que o contexto do debate é importante para esse argumento. Edgeworth identificou que havia se estabelecido a aceitação do "egoísmo" como a suposição comportamental fundamental — da sua própria pesquisa — ao dismantelar a aceitabilidade do "utilitarismo" como uma descrição do comportamento em si. Utilitarismo está, certamente, longe de ser a única abordagem não egoísta. Além disso, entre as reivindicações de si mesmo e as reivindicações de todos, estão as reivindicações de uma variedade de grupos — por exemplo, família, amigos, comunidades locais, grupos homólogos, classes econômicas e sociais. Os conceitos de "responsabilidade familiar", "ética de negócios", "consciência de classes", entre outros, relacionam-se com as áreas intermediárias de interesse, e a destituição do utilitarismo como uma teoria descritiva de comportamento não condiciona o

² Nota dos tradutores: todas as referências da tradução foram mantidas como no texto original de Sen.

³ Edgeworth, F. Y. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. Londres, 1881, p. 16.

⁴ Edgeworth, 1881, p. 104. Na verdade, ele fez algumas observações interessantes sobre os resultados do egoísmo "impuro", admitindo um elemento de simpatia por uns pelos outros. As observações foram investigadas e analisadas por David Collard, "Edgeworth's Propositions on Altruism", *Economic Journal* 85 (1975).

⁵ Edgeworth, 1881, p. 52.

⁶ Edgeworth, 1881, p. 52-53.

egoísmo como única alternativa. Seria difícil negar a relevância de algumas dessas considerações para a economia de negócios e contratos.

Deve-se constatar que a questão de Edgeworth sobre os resultados do contato econômico entre indivíduos puramente autointeressados teve o mérito de ser imediatamente relevante para uma pesquisa abstrata que já durava mais de cem anos, e que foi muito discutida em debates envolvendo Herbert Spencer, Henry Sidgwick e outros grandes pensadores do período. Dois anos antes de Edgeworth publicar *Mathematical Psychics*, Herbert Spencer publicou uma análise elaborada da relação entre egoísmo e altruísmo em *The Data of Ethics*. Ele chegou à confortável — e, de alguma forma, pouco elucidativa — conclusão de que “a felicidade geral deve ser alcançada por meio da busca adequada de indivíduos por sua própria felicidade; enquanto, reciprocamente, a felicidade de indivíduos deve ser alcançada, em parte, pela busca deles pela felicidade geral”⁷. No contexto dessa discussão relativamente abstrata, a análise econômica rigorosa de Edgeworth, baseada em um modelo bem-definido de contratos entre dois indivíduos autointeressados, ou entre dois tipos (idênticos) de indivíduos autointeressados, respondeu uma pergunta hipotética antiga de modo claro.

Sugeriu-se que, no modelo de Edgeworth, baseado no comportamento egoísta, havia uma correspondência notável entre equilíbrio cambial em mercados competitivos e o “núcleo” da economia, como é chamado nos termos da economia moderna. Um resultado é tido como o núcleo da economia se, e apenas se, este se adequa a um *set* de condições de impossibilidade de melhoria. Essas condições, sem grandes pormenores, não apenas ditam que ninguém pode melhorar sua situação sem que outro indivíduo tenha sua situação piorada (essa condição é o que chamamos de “*Pareto optimum*”), mas também que ninguém está numa situação pior do que estaria sem comércio, e que nenhuma coligação de indivíduos,

ao alterar o comércio entre si, poderia por si só melhorar a sua situação. Edgeworth demonstrou que, dadas certas premissas gerais, qualquer equilíbrio que possa emergir em um mercado competitivo deve atender a essas condições e estar no “núcleo”. Deste modo, no modelo de Edgeworth, o equilíbrio do mercado competitivo não é, neste sentido, dominado por qualquer agenda alternativa viável, dada a distribuição inicial de dotações. Mais surpreendentemente em alguns aspectos foi o resultado inverso de que, se o número de indivíduos de cada tipo fosse aumentado sem limites, o núcleo (representando esses resultados não dominados) encolheria em direção ao conjunto de equilíbrios competitivos; ou seja, o núcleo não seria muito mais extenso do que o *set* de equilíbrios competitivos. Esse par de resultados tem sido muito elaborado e estendido na recente literatura a respeito de equilíbrio geral com modelos similares e, essencialmente, as mesmas suposições comportamentais⁸.

Estar no centro, no entanto, não é uma conquista tão importante do ponto de vista do bem-estar social. Uma pessoa que começa mal dotada pode permanecer pobre e desprovida mesmo após as transações, e, se estar no núcleo for tudo que a concorrência oferece, a pessoa despossuída de propriedades pode ser perdoada por não considerar essa conquista como um “grande negócio”. Edgeworth tomou nota disso ao considerar o problema da escolha entre diferentes equilíbrios competitivos. Ele observou que, para a boa sociedade utilitarista, “competitividade exige ser suplementada por arbitragem, e a base da arbitragem entre contratantes autointeressados é a maior soma total possível do utilitário”⁹. Quanto aos aspectos constitucionais dessa arbitragem e suas implicações de longo alcance para a distribuição de propriedade, Edgeworth não se aprofundou de fato, apesar da aparência superficial ao contrário. Com base na conquista da concorrência, por mais limitada que seja, Edgeworth se sentiu no direito de ser “tendencioso

⁷ H. Spencer, *The Data of Ethics* (1879), p. 238.

⁸ Ver, especialmente, K. Arrow e F. Hahn, *General Competitive Analysis* (1971).

⁹ Edgeworth, 1881, p. 56.

para uma cautela mais conservadora na reforma". No cálculo da "utilidade de instituições pré-utilitárias", Edgeworth se impressionou "com a vista da natureza, não, como no quadro deixado por Mill, de todo ruim, mas uma primeira aproximação do melhor"¹⁰.

Não estou preocupado, neste ensaio, em examinar se a aproximação é remota o bastante. (Este acredito ser o caso mesmo entre a estrutura de suposições usadas por Edgeworth, mas isso não é central para o assunto deste trabalho). Estou preocupado, aqui, com o viés do homem que faz parte da análise de Edgeworth e sobrevive mais ou menos intacto em muitas teorias da economia moderna. A visão é uma, certamente, estilizada e voltada especificamente para enfrentar uma disputa relativamente abstrata com a qual Spencer, Sidgwick e muitos outros principais pensadores contemporâneos estavam preocupados — ou seja, em que sentido e até que ponto o comportamento egoísta conquistaria o bem comum? Se o comportamento egoísta é ou não uma suposição precisa, na realidade, não tem, obviamente, qualquer influência na precisão da resposta de Edgeworth à pergunta colocada. Dentro da estrutura de um modelo econômico limitado, isso proporcionou uma resposta clara para a questão sobre egoísmo e bem comum.

Este debate em particular já perdura há bastante tempo e continua a proporcionar motivação para muitos dos exercícios recentes da teoria econômica atual. A natureza limitada dessa discussão tem tido uma influência decisiva sobre a escolha de modelos econômicos e o conceito de seres humanos nestes. Em seu distinto texto sobre a teoria do equilíbrio geral, Arrow e Hahn constataam:

Existe agora uma longa e bastante imponente lista de economistas, de Adam Smith à atualidade, que buscaram demonstrar que uma economia descentralizada motivada pelo autointeresse e guiada por sinais de preço seria compatível com uma disposição coerente de recursos econômicos que seriam considerados, em um senso bem definido, como superior

para uma grande classe de disposições alternativas possíveis. Além disso, os sinais de preço operariam de modo a estabelecer este grau de coerência. É importante entender o quão surpreendente essa alegação deve ser para qualquer um que não esteja exposto à tradição. A resposta de "senso comum" imediata para a questão "Como será uma economia motivada por ganância individual e controlada por um número muito grande de agentes diferentes?" provavelmente será: haverá caos. O fato de uma resposta bastante diferente ter sido considerada verdadeira há muito tempo e, de fato, ter permeado o pensamento econômico de um grande número de pessoas, que não são economistas de forma alguma, é, por si só, motivo suficiente para investigá-la seriamente. Tendo a proposta sido apresentada e considerada muito seriamente, é importante saber não apenas se ela é verdadeira, mas também se *poderia* ser. Grande parte do que se segue diz respeito a essa questão, que nos parece merecer consideravelmente a atenção dos economistas.

A principal preocupação aqui não é com a relação de modelos postulados com o mundo econômico real, mas com a precisão de respostas para questões bem definidas colocadas com suposições pré-selecionadas, que restringem severamente a natureza de modelos que podem ser admitidos à análise. Um conceito específico de homem está enraizado na própria questão, e não há liberdade para se afastar desse conceito enquanto alguém estiver empenhado em responder essa questão. A natureza do homem nos modelos econômicos atuais continua, portanto, a refletir essa formulação particular de certas questões filosóficas gerais levantadas no passado. O realismo do conceito escolhido sobre o homem, simplesmente, não é parte desta pesquisa.

II.

Existe outra razão não empírica — e possivelmente mais simples — de por que o conceito de "homem" nos modelos econômicos tende a ser o de egoísmo autointeressado. É possível definir o interesse de uma pessoa de modo que, não importa o que ela faça, ela possa ser vista como promovendo seus próprios interesses em cada ato isolado de escolha¹¹. Embora formalizada

¹⁰ Edgeworth, 1881, p. 82.

¹¹ Se as ações de uma pessoa hoje afetam seu bem-estar no futuro, então, sob esta abordagem, seus interesses futuros devem ser definidos nos termos do modo como elas são *avaliadas no presente*. No geral, não há motivo para presumir que interesses futuros avaliados

há relativamente pouco tempo no contexto da teoria da preferência revelada, esta abordagem é de antiguidade respeitável, e Joseph Butler já argumentava contra ela em *Rolls Chapel* há dois séculos e meio¹². A redução do homem a um animal autointeressado depende, nesta abordagem, de uma definição cuidadosa. Se é observado que você escolhe x rejeitando y , é declarada a sua preferência "revelada" de x sobre y . Sua utilidade pessoal é, então, definida como simplesmente uma representação numérica dessa "preferência", atribuindo uma utilidade superior a uma alternativa "preferida". Com este *set* de definições, dificilmente pode-se escapar da maximização da sua própria utilidade, exceto por meio de inconsistência. Evidentemente, se você escolhe x e rejeita y em uma ocasião, e então procede prontamente para fazer o exato oposto, você pode prevenir que o teórico de preferência revelada lhe atribua uma ordem de preferência, deste modo, impedindo-o de selar uma função de utilidade a qual você deve parecer estar maximizando. Assim, ele precisará concluir ou que você é inconsistente, ou que suas preferências estão mudando. O teórico de preferência revelada pode ser frustrado por meio de inconsistências mais sofisticadas também¹³. No entanto, se você for consistente, não importará se você é um egoísta obstinado, ou um altruísta delirante, ou um militante com consciência de classe; você parecerá estar maximizando sua própria utilidade neste mundo encantado de definições. Tomando emprestado da terminologia utilizada na conexão com taxaço, se a justificativa de Arrow-Hahn da suposição do egoísmo equivale a tangenciar a questão, a abordagem da preferência revelada parece mais uma peça robusta dessa evasão.

Esta abordagem de egoísmo definicional tam-

bém pode ser denominada de escolha racional, e isso envolve nada mais que consistência interna. As escolhas de uma pessoa são consideradas "racionais" nessa abordagem se, e somente se, essas escolhas puderem ser *todas* explicadas nos termos de alguma relação consistente de preferências com a definição de preferência revelada, isto é, se todas as suas escolhas puderem ser explicadas como a escolha das "mais preferidas" alternativas com respeito à relação de preferências postuladas¹⁴. A justificativa dessa abordagem parece ser baseada na ideia de que a única maneira de entender a real preferência de uma pessoa é examinando suas escolhas de fato, e não existe uma maneira escolha-independente de entender suas atitudes em relação às alternativas. (Este ponto de vista, a propósito, não se limita apenas a economistas. Quando, há muitos anos, tive que fazer meu exame de qualificação em Literatura Inglesa, na Universidade de Calcutá, uma das questões que tivemos que responder sobre *A Midsummer Night's Dream* foi: "Compare as personagens de Hermia e Helena. Qual você escolheria?").

Tentei demonstrar em outro trabalho que, uma vez que evitamos as definições curiosas de preferência e bem-estar, essa abordagem pressupõe, ao mesmo tempo, muito e muito pouco: muito pouco porque há fontes de informação sem escolhas sobre preferência e bem-estar do modo que esses termos são comumente entendidos, e muito porque a escolha pode refletir um compromisso entre uma variedade de considerações das quais o bem-estar pessoal pode ser apenas uma¹⁵.

As questões psicológicas complexas subjacentes à escolha, recentemente, têm sido evidenciadas com vigor por uma série de estudos

no presente coincidirão com os interesses avaliados no futuro. Isto soma uma dimensão adicional ao problema, e sou grato ao Derek Parfit por ter-me convencido quanto à importância conceitual dessa questão.

¹² J. Butler, *Fifteen Sermons Preached at the Rolls Chapel* (1726); ver também T. Nagel, *The Possibility of Altruism* (1970), p. 81.

¹³ H. S. Houthakker, *Revealed Preference and the Utility Function* (1950); P. Samuelson, *The Problem of Integrability in Utility Theory* (1950).

¹⁴ Para os principais resultados analíticos, ver M. Richter, *"Rational Choice"* (1971).

¹⁵ A. Sen, *Behaviour and the Concept of Preference* (1973). Ver também o recente estudo de S. Körner: *Experience and Conduct* (Cambridge, 1972). Também, T. Schwartz *"Von Wright's Theory of Human Welfare: A Critique"*. Após: P. Schlipp *"The Philosophy of Georg Henrik von Wright"*; T. Majumdar, *The Concept of Man in Political Economy and Economics* (1976); e F. Schick, *Rationality and Sociality* (1976).

penetrantes que tratam das decisões do consumidor¹⁶ e atividades de produção¹⁷. É uma questão em aberto saber se essas características comportamentais podem ser captadas dentro dos limites formais de escolha consistente, dos quais depende a abordagem de maximização do bem-estar¹⁸.

III.

Paul Samuelson observou que muitos economistas iriam "separar a economia da sociologia com base no comportamento racional ou irracional, onde esses termos são definidos na penumbra da teoria utilitarista"¹⁹. Esse viés pode muito bem ser rejeitado, por bons motivos, por sociólogos, mas o fardo que economistas têm de lidar neste viés de dicotomia pode ser visto se observarmos que a abordagem do "comportamento racional", como tipicamente interpretado, conduz a uma teoria notavelmente muda. Comportamento, ao que parece, é "explicado em termos de preferências, que, por sua vez, são definidas apenas pelo comportamento". Não surpreendentemente, excursões cíclicas têm sido frequentes. No entanto, Samuelson está, sem dúvidas, correto em afirmar que a teoria "não é, na perspectiva técnica, *sem sentido*"²⁰. O motivo é bem simples. Como já discutido, a abordagem impõe a condição de consistência interna da escolha observada, e isso pode ser refutado por observações de fato, tornando a teoria "sem sentido" da perspectiva cuja declaração de Samuelson se designa.

A exigência de consistência tem um surpreendente poder de corte. Várias características gerais de relações de demanda podem derivar disso.

Entretanto, no presente contexto, a questão principal é a possibilidade de utilizar a necessidade de consistência para *testes* reais. Samuelson especifica a necessidade de "condições observacionais ideais" para as implicações da abordagem serem "refutadas ou verificadas". Isto não é, no entanto, fácil de corresponder dado que, por um lado, nosso amor pela variedade torna ilegítimo considerar atos de escolha individual como unidades adequadas (em vez de *sequência* de escolhas), enquanto, por outro lado, o lapso de tempo torna difícil distinguir entre inconsistências e mudanças de gosto. De fato, houve muito poucas tentativas sistemáticas de testar a consistência do comportamento das pessoas no dia a dia, embora tenha havido experiências inventadas interessantes e úteis sobre as reações das pessoas à incerteza sob condições de laboratório. O que conta como evidência admissível permanece incerto. Se, atualmente, fossem elencados economistas de diferentes escolas, certamente, encontrar-se-ia a coexistência das crenças de que (i) a teoria do comportamento racional é infalsificável, (ii) esta é falsificável e até agora infalsificável, e (iii) é falsificável e, de fato, patentemente falso²¹.

Todavia, para meus objetivos aqui, este não é o objeto central. Mesmo se a consistência requisitada fosse obtida, a questão do egoísmo permaneceria não resolvida no sentido puramente definicional, como já observado. Uma pessoa com escolha consistente pode ter qualquer grau de egoísmo que queiramos especificar. É verdade, certamente, que, no caso especial da pura escolha do consumidor em relação a bens privados, o teórico de preferência revelada busca relacionar a "preferência" ou a "utilidade" da pessoa com seu

¹⁶ Ver T. Scitovsky, *The Joyless Economy: An Inquiry into Human Satisfaction and Consumer Dissatisfaction* (1976). Ver também a crítica geral da suposição do comportamento do consumidor "racional" por J. Kornai: *Anti-Equilibrium* (1971), capítulo 11; e a literatura sobre "modelos psicológicos de escolha", especificamente, D. McFadden: *Economic Applications of Psychological Choice Models* (1975).

¹⁷ Ver H. Liebenstein, *Allocative Efficiency vs. x-Efficiency* (1966). Ver também críticas da suposição tradicional de maximização de lucro no comportamento de negócios, especificamente W. Baumol: *Business Behavior, Value and Growth* (1959); R. Marris, *The Economic Theory of Managerial Capitalism* (1964); O. Williamson, *The Economics of Discretionary Behavior* (1967); e A. Silberston, *Price Behavior of Firms* (1970).

¹⁸ Sobre as condições de consistência exigidas para analisar a escolha em termos de uma relação binária, consultar Sen (1971); H. Herzberger, *Ordinal Preference and Rational Choice* (1973); K. Suzumura, *Rational Choice and Revealed Preference* (1976); S. Kanger, *Choice Based on Preference* (1976).

¹⁹ P. Samuelson, *The Foundation of Economics* (1955), p. 90.

²⁰ P. Samuelson, 1955, p. 91.

²¹ As recentes críticas filosóficas da teoria do comportamento racional incluem, entre outras, M. Hollis e J. Nell, *Rational Economic Man* (1975); S. Wong, *On the Consistency and Completeness of Paul Samuelson's Programme in the Theory of Consumer Behaviour* (1975). Ver também as críticas pragmáticas de Kornai, *Anti-Equilibrium*, capítulo 11.

próprio pacote de *commodities*. Essa restrição surge, contudo, não de alguma garantia de que ela está preocupada somente com seus próprios interesses, mas do fato de que seu próprio pacote de consumos — ou o de sua família — é o único pacote sobre o qual ela tem *controle* direto em seus atos de escolha. A questão do egoísmo permanece completamente aberta.

Acredito que a questão também requer uma formulação mais clara do que tende a receber, e é para esta questão que me voltarei agora.

IV.

À medida que consideramos os desvios do «antipático isolamento, assumido abstratamente na Economia», utilizando as palavras de Edgeworth, devemos fazer a distinção entre dois conceitos separados: (i) simpatia e (ii) comprometimento. O primeiro corresponde ao caso no qual a preocupação com os outros afeta diretamente o seu próprio bem-estar. Se o conhecimento da tortura de outros faz mal a você, é um caso de simpatia; se isto não lhe faz sentir pessoalmente pior, mas você sabe que é errado e está pronto para fazer algo para impedir, é um caso de comprometimento. Não tenho a intenção de alegar que as palavras escolhidas têm qualquer grande mérito, porém a distinção é, acredito, importante. Pode ser discutido que o comportamento baseado na simpatia é um importante senso de egoísmo, pois alguém fica satisfeito com o prazer do outro e sente dor com a dor do outro, e a busca pela utilidade própria de alguém pode, portanto, ser ajudada por uma ação simpática. É a ação baseada em comprometimento em vez de simpatia que seria não egoísta nesse sentido. (Observe, no entanto, que a *existência* da simpatia não implica que a ação útil para outros deve ser *baseada em simpatia* no sentido que a ação não aconteceria se alguém obtivesse menos ou nenhum conforto do bem-estar dos outros. Esta questão de *causalidade* será abordada no presente).

"Simpatia" é, de certa forma, um conceito mais fácil de analisar do que o "comprometimento". Quando o senso de bem-estar de uma pessoa é psicologicamente dependente do bem-estar

de outro alguém, é um caso de simpatia; em outras palavras, a consciência da melhoria do bem-estar desse outro alguém melhora diretamente a situação dessa pessoa. (Evidentemente, quando a influência é negativa, a relação é melhor nomeada como "antipatia", porém podemos economizar em terminologia e nos ater ao termo "simpatia", apenas salientando que a relação pode ser positiva ou negativa). Enquanto a simpatia conecta coisas semelhantes entre si — ou seja, o bem-estar de diferentes pessoas —, o comprometimento relaciona a escolha a níveis antecipados de bem-estar. Um modo de definir o comprometimento é no sentido de uma pessoa escolher uma ação que ela acredita que lhe renderá um nível mais baixo de bem-estar pessoal do que uma alternativa que também está disponível para ela. Observe que a comparação é entre níveis *antecipados* de bem-estar, e, portanto, esta definição de "comprometimento" exclui atos que vão contra o interesse próprio, resultando puramente de uma falha em prever consequências.

Uma questão mais difícil surge quando a escolha de uma pessoa acaba coincidindo com a maximização do seu nível antecipado de bem-estar pessoal, porém essa não é a *razão* para sua escolha. Se quisermos abrir espaço para isso, podemos expandir a definição de comprometimento para incluir casos em que a escolha da pessoa, maximizando o bem-estar pessoal previsto, não seria afetada sob pelo menos uma condição contrafactual em que a ação escolhida deixaria de ser para maximizar o bem-estar pessoal. Comprometimento neste sentido mais inclusivo pode ser complicado de averiguar, não apenas no contexto da escolha dos outros, mas também na própria escolha de alguém, uma vez que não é sempre claro o que alguém faria se as circunstâncias fossem diferentes. Esse sentido mais amplo pode ter uma relevância em particular quando alguém age com base na preocupação com o dever que, se violado, poderia causar remorso, mas a ação é de fato escolhida por senso de dever e não somente para evitar o mal-estar resultante do remorso que ocorreria

se agisse de outra forma. (Certamente, mesmo o sentido mais restrito de comprometimento cobriria o caso em que o mal-estar resultante do remorso, se houver, for *compensado* pelo ganho em bem-estar).

Até então, não me referi à incerteza relativa ao bem-estar previsto. Quando isto é introduzido, o conceito de "simpatia" não é afetado, no entanto, o de "comprometimento" exigirá uma reformulação. As modificações necessárias dependerão da reação da pessoa à incerteza. O caso mais simples é, provavelmente, aquele no qual a ideia de uma pessoa sobre o que uma "loteria" oferece a ela nos termos de ganhos pessoais é capturada pela "utilidade esperada" do bem-estar pessoal (isto é, adicionando bem-estar pessoal de diferentes resultados ponderados pela probabilidade de ocorrência de cada resultado). Neste caso, a discussão é reformulada por completo simplesmente substituindo bem-estar pessoal por bem-estar pessoal *esperado*; comprometimento, então, envolve escolher uma ação que produza um bem-estar esperado menor do que uma ação alternativa disponível. (O sentido amplo também pode ser modificado correspondentemente).

Na terminologia da teoria econômica moderna, simpatia é um caso de "externalidade". Muitos modelos descartam externalidades, por exemplo, o modelo padrão de estabelecer que cada equilíbrio competitivo é um Pareto *optimum* e pertence ao núcleo da economia. Se a existência da simpatia fosse permitida nesses modelos, alguns desses resultados padrões seriam frustrados, embora de forma alguma todos eles²². No entanto, isso não exigiria uma revisão séria da estrutura básica desses modelos. Por outro lado, comprometimento envolve, em um sentido muito real, escolha contrapreferencial, destruindo a suposição crucial de que uma alternativa escolhida deve ser melhor que (ou ao menos tão boa quanto) as outras para que a pessoa a escolha, e isto certamente exigiria que modelos fossem formulados de um modo essencialmente

diferente.

O contraste entre simpatia e comprometimento pode ser ilustrado pela história de dois meninos que encontram duas maçãs, uma grande e uma pequena. O menino *A* diz ao menino *B*: "Você escolhe". *B* imediatamente escolhe a maçã maior. *A* fica chateado e se permite a observação de que isso foi grosseiramente injusto. "Por quê?", pergunta *B*. "Qual você escolheria, se fosse você a escolher no meu lugar?"; "A menor, é claro", *A* responde. *B* agora está triunfante: "Então de que você está reclamando? É essa que você ficou!". *B* certamente ganhou essa rodada da discussão, mas, de fato, *A* não perderia nada com a escolha de *B* se tivesse feito sua própria escolha hipotética da maçã menor baseada em simpatia, e não em comprometimento. A raiva de *A* indica que este provavelmente não era o caso.

Comprometimento é, de certo, algo intimamente conectado com a moral do indivíduo. No entanto, a moral nesta questão é compreendida em um sentido muito amplo, cobrindo uma variedade de influências desde religiosas a políticas, desde o mal-entendido ao bem argumentado. Quando, em *The Devil's Disciple*, de Bernard Shaw, Judith Anderson interpreta a vontade de Richard Dudgeon de ser enforcada no lugar do marido como decorrente de simpatia por ele ou amor por ela, Richard é inflexível em sua negação: "O que fiz na última noite, fiz a sangue frio, não me importando com seu marido ou com você tanto quanto eu me importo comigo mesmo. Não tive motivo e nem interesse: tudo que posso lhe dizer é que quando chegou o momento de tirar o meu pescoço da forca para colocar o de outro homem, não consegui fazê-lo"²³.

A característica do comprometimento com a qual estou mais preocupado aqui é o fato que ele cria uma barreira entre escolha pessoal e bem-estar pessoal, e muito da teoria econômica tradicional depende da identidade de ambos. Essa identidade, às vezes obscurecida pela ambiguidade do termo "preferência", dado que o

²² Ver A. Sen, *Labour Allocation in a Co-operative Enterprise* (1966); S. Winter, Jr., *A Simple Remark on the Second Optimality Theorem of Welfare Economics* (1969); Collard, *Edgeworth's Propositions*; G. Archibald e D. Donaldson, *Non-paternalism and Basic Theorems of Welfare Economics* (1976).

²³ Ver G. Shaw, *Three Plays for Puritans* (1966), p. 94.

uso comum da palavra permite a identificação de preferência como o conceito de ser melhor que, e, ao mesmo tempo, não é nada natural definir "preferido" como "escolhido". Não tenho opiniões fortes quanto ao uso "correto" da palavra "preferência", e ficaria satisfeito se ambas as utilizações não fossem feitas *simultaneamente*, buscando uma afirmação empírica em virtude de duas definições²⁴. A conexão básica entre comportamento de escolha e conquistas de bem-estar nos modelos tradicionais é cortada assim que o comprometimento é admitido como um ingrediente de escolha.

V.

"Certo", você pode dizer, "mas o quão relevante é isto tudo para o tipo de escolhas com as quais economistas estão preocupados? Economia não tem muito a ver com a marcha de Richard Dudgeon para a força". Acredito que se deve concordar imediatamente que, para muitos tipos de comportamento, é improvável que o comprometimento seja um ingrediente importante. Na aquisição privada de muitos bens de consumo, o escopo para o exercício de comprometimento pode, de fato, ser limitado e pode aparecer muito raramente em atos exóticos, como o boicote aos abacates sul-africanos ou a rejeição dos feriados espanhóis. Portanto, para muitos estudos sobre comportamento do consumidor e interpretações disso, comprometimento pode não representar um grande problema. Mesmo a simpatia pode não ser extremamente importante, as fontes de interdependência interpessoal situadas em outro foco, por exemplo, no desejo de manter as aparências ou ser influenciado pelos hábitos de outras pessoas²⁵.

No entanto, a Economia não está preocupada com o comportamento do consumidor; muito

menos é o consumo confinado a "bens privados". Uma área em que a questão do comprometimento é mais relevante é a dos chamados bens públicos. Esta deve ter o contraste com "bens privados", o quais têm a característica de não poderem ser utilizados por mais de uma pessoa: se você comeu uma fatia de torta de maçã, eu nem mesmo consideraria comê-la também. O mesmo não acontece com bens públicos, por exemplo, uma estrada ou um parque público, dos quais eu e você poderíamos usufruir. Em muitos modelos econômicos, bens privados são os únicos considerados, e este é tipicamente o caso no qual a "mão invisível" tem a tarefa de fazer um bem notável. Contudo, na verdade, os bens públicos são importantes para muitas economias e abrangem uma vasta gama de serviços, desde estradas à iluminação pública e à defesa. Existe muita evidência de que a parcela de bens públicos em consumo nacional aumentou drasticamente na maior parte dos países do mundo.

O problema da alocação ideal de bens públicos também tem sido muito discutido, especialmente na literatura econômica recente²⁶. Muita atenção, em particular, foi destinada ao problema da revelação correta de preferências. Isto surge mais nitidamente no caso de esquemas de assinatura, nos quais uma pessoa é cobrada de acordo com os benefícios recebidos. O principal problema centra-se no fato de que é de interesse geral entender o benefício esperado, mas este eufemismo pode levar à rejeição de um projeto público que teria sido justificado se os benefícios reais fossem conhecidos. A análise dessa dificuldade, às vezes conhecida como o problema do "carona", recentemente causou propostas extremamente engenhosas para contornar essa ineficiência da estrutura da ação egoísta²⁷. O mecanismo de recompensas é configurado com uma astúcia

²⁴ Ver A. Sen, *Behaviour and the Concept of Preference* (1973); e Shick, *Rationality and Sociality*.

²⁵ Ver J. Duesenberry: *Income, Saving and the Theory of Consumer Behavior* (1949); S. Prais e H. Houthakker, *The Analysis of Family Budgets* (1955); W. Gaertner, *A Dynamic Model of Interdependent Consumer Behaviour* (1973); R. Pollak, *Interdependent Preferences* (1976).

²⁶ Ver E. Lindahl, *Die Gerechtigkeit der Besteuerung* (1919), traduzido por R. Musgrave e A. Peacock em *Classics in the Theory of Public Finance* (1967); P. Samuelson, *The Pure Theory of Public Expenditure* (1954); R. Musgrave, *The Theory of Public Finance* (1969); L. Johansen, *Public Economics* (1966); D. Foley, *Lindahl's Solution and the Core of an Economy with Public Goods* (1970); E. Malinvaud, *Prices for Individual Consumption. Quantity Indicators for Collective Consumption* (1972)

²⁷ T. Groves e J. Ledyard, *Optimal Allocation of Public Goods: A Solution to the 'Free Rider Problem'* (1975); J. Green e J. Laffont, *On the Revelation of Preference for Public Goods* (1974). Ver também J. Dreze e D. de la Vallee Poussin, *A Tatonnement Process for Public Goods* (1971); E. Malinvaud, *A Planning Approach to the Public Goods Problem* (1971); V. Smith, *Incentive Compatible Experimental Processes for the*

tão ímpia que as pessoas têm um incentivo para revelar exatamente sua verdadeira intenção de pagar pelo bem público em questão. Uma dificuldade nessa solução surge de uma suposta limitação das possibilidades estratégicas abertas ao indivíduo, cuja remoção leva a um resultado de impossibilidade²⁸. Outra dificuldade diz respeito ao fato que, ao dar às pessoas o incentivo para revelar a verdade, o dinheiro é distribuído e a distribuição de renda muda de uma forma que não é guiada por considerações distributivas. Este efeito pode, é claro, ser desfeito por uma redistribuição de doações iniciais e participação nos lucros²⁹, embora tal ação obviamente levante dificuldades próprias.

Central para esse problema é a suposição de que, quando perguntado, o indivíduo dá a resposta que irá maximizar seu ganho pessoal. O quão boa é esta suposição? Duvido que, de modo geral, ela seja muito boa. ("Onde fica a Estação Ferroviária?", ele me pergunta. "Lá", eu digo, apontando para os correios, "e você poderia, por favor, enviar essa carta para mim no caminho?" "Sim", ele diz, determinado a abrir o envelope para checar se contém algo valioso). Mesmo no contexto particular da revelação de preferências por bens públicos, o comportamento de maximização de ganhos pode não ser a melhor suposição. Leif Johansen, um dos principais contribuidores para a economia pública, é, acredito, certo ao questionar a suposição neste contexto:

A teoria econômica nisto, bem como em outros campos, tende a sugerir que as pessoas são honestas somente na medida em que têm incentivos para fazê-lo. Esta é uma suposição do *homo oeconomicus* que está longe de ser inteiramente verdade, e cuja necessita confronto com realidades observadas. De fato, uma simples linha de pensamento sugere que a suposição dificilmente pode ser verdade em sua forma mais extrema. Nenhuma sociedade

seria viável sem algumas normas e regras de conduta. Tais normas e regras são necessárias para viabilidade exatamente em campos onde incentivos estritamente econômicos são inexistentes e não podem ser criados³⁰.

O que está em questão não é se as pessoas invariavelmente respondem de forma honesta a todas as interrogações, mas se elas sempre dão uma resposta que maximiza seus ganhos, ou, em alguma medida, se elas dão respostas de maximização de ganhos com frequência suficiente para fazer disso a suposição geral apropriada à teoria econômica. A presença de respostas que não maximizam ganhos, incluindo as verdadeiras, imediatamente traz comprometimento como parte do comportamento.

O tópico é relevante também para a literatura recente sobre votação estratégica. Uma quantidade de belos resultados analíticos foi estabelecida recentemente, mostrando a impossibilidade de qualquer processo de voto que satisfaça determinados requisitos elementares e fazendo da votação honesta a estratégia de maximização de ganhos para todos³¹. A correção desses resultados não está em debate, mas é apropriado assumir que as pessoas sempre tentam maximizar ganhos pessoais em seu comportamento de voto? De fato, em grandes eleições, é difícil demonstrar que qualquer eleitor tem qualquer perspectiva real de afetar o resultado com seu voto, e, se a votação envolve algum custo, a rede de ganhos esperada da votação pode ser tipicamente negativa. No entanto, a proporção de participação em grandes eleições ainda pode ser bem alta, e tentei dissertar em outro trabalho que, nessas eleições, as pessoas podem frequentemente ser "guiadas não tanto pela maximização da utilidade esperada, mas por algo muito mais simples, isto é, apenas uma vontade de registrar a verdadeira

Provision of Public Goods (1976).

²⁸ J. Ledyard e D. Roberts, *On the Incentive Problem for Public Goods* (1974). Ver também L. Hurwicz, *On Informationally Decentralized Systems* (1972).

²⁹ Ver o Teorema 4.2 em Groves e Ledyard, *Optimal Allocation of Public Goods* (1975).

³⁰ L. Johansen, *The Theory of Public Goods: Misplaced Emphasis* (1976). Ver também J. Laffont, *Macroeconomic Constraints, Economic Efficiency and Ethics* (1974); P. Bohm, *Estimating Demand for Public Goods: An Experiment* (1972).

³¹ A. Gibbard, *Manipulation of Voting Schemes: A General Result* (1973); M. Satterthwaite, *Strategy-proofness and Arrow's Conditions* (1975); D. Schmeidler e H. Sonnenschein, *The Possibility of Non-manipulable Social Choice Functions* (1974); B. Dutta e P. Pattanaik, *On Nicely Consistent Voting Systems* (1975); P. Pattanaik, *Strategic Voting without Collusion under Binary and Democratic Group Decision Rules* (1975); B. Peleg, *Consistent Voting Systems* (1976); A. Gibbard, *Social Decision, Strategic Behavior, and Best Outcomes: An Impossibility Result* (1976).

preferência de um indivíduo"³². Se esta vontade refletir um senso de comprometimento, então o comportamento em questão estaria em desacordo com a visão do homem na teoria econômica tradicional.

VI.

O ponto do comprometimento é importante em vários outros contextos econômicos³³. Isto é central para o problema da motivação no trabalho, cuja importância para o desempenho de produção dificilmente pode ser ignorada.

Certamente, é custoso e pode ser impossível conceber um sistema de supervisão com recompensas e punições tais que todo mundo tenha o incentivo para se inserir nele. Todo sistema econômico, portanto, tende a confiar na existência de atitudes em relação ao trabalho que substituem o cálculo da renda líquida de cada unidade de inserção. O condicionamento social tem um peso extremamente importante aqui³⁴. Estou convencido de que as dificuldades econômicas britânicas atuais têm muito a ver com problemas de motivação no trabalho que estão por fora da economia de recompensas e punições, e uma razão pela qual economistas parecem ter tão pouco a contribuir nessa área é a negligência da teoria econômica tradicional para com toda essa questão de comprometimento e as relações sociais em volta disso³⁵.

Essas questões estão conectadas, é claro, com ética, desde que razões morais influenciam as ações de um indivíduo, mas, em um sentido mais amplo, essas são questões de cultura, das quais a moralidade faz parte. De fato, para apontar um caso extremo, na "revolução cultural" chinesa,

um dos objetivos principais era o aumento do senso de comprometimento visando resultados econômicos: "o objetivo da Grande Revolução Cultural do Proletariado é revolucionar a ideologia do povo e, como consequência para alcançar resultados maiores, mais rápidos, melhores e mais econômicos em todas as áreas de trabalho"³⁶. Obviamente, a China estava experimentando reduzir drasticamente o papel dos incentivos materiais na produção, o que teria, certamente, aumentado o papel que o comprometimento deveria desempenhar, entretanto, mesmo entre os sistemas tradicionais de pagamento, normalmente, muita confiança é depositada nas regras de conduta e modos de comportamento que vão além dos incentivos estritamente econômicos³⁷. Dirigir uma organização *inteiramente* com base em incentivos de ganho pessoal é uma tarefa praticamente impossível.

Terei um pouco mais a dizer em breve sobre o que pode estar por trás do senso de comprometimento, porém gostaria de enfatizar, neste ponto, que a moralidade ou a cultura subjacentes a este podem muito bem ser de um tipo limitado — muito distante da grandeza de abordagens como o utilitarismo. Os "conluícos implícitos" que foram observados no comportamento de negócios em oligopólios parecem funcionar com base em um sistema de confiança mútua e senso de responsabilidade, os quais têm limites bem definidos, e as tentativas de "universalização" do mesmo tipo de comportamento em outras esferas de ação podem não combinar com isso. Esta é, estritamente, uma questão de ética empresarial que se aplica a um domínio bastante limitado.

Similarmente, em negociações salariais e em

³² Ver A. Sen, *Collective Choice and Social Welfare* (1970), p. 195.

³³ Ver a discussão de Ragnar Frisch sobre a necessidade por "uma fundação teórica realista para política social" em sua obra *Samarbeid mellom Politikere og Økonometrikere om Formuleringen av Politiske Preferenser* (1971). Ver também J. Mirrlees, *The Economics of Charitable Contributions* (1973).

³⁴ Ver A. Fox, *Beyond Contract: Work, Power and Trust Relations* (1974); H. Nutting, *The Firm as a Social Institution: The Failure of a Contractarian Viewpoint* (1976).

³⁵ Comparando: "Nem... devemos esquecer a extensão pela qual a teoria convencional ignora como e por que o trabalho é organizado dentro da empresa e o estabelecimento da forma como é, o que pode ser chamado de "relações sociais" do processo de produção", R. Gordon, *Rigor and Relevance in a Changing Institutional Setting* (1976). Ver também R. Dahrendorf, *Class and Class Conflict in Industrial Society* (1959); O. Williamson, *The Evolution of Hierarchy: An Essay on the Organization of Work* (1976); e S. Marglin, *What Do Bosses Do? The Origins and Functions of Hierarchy in Capitalist Production* (1974).

³⁶ "A Decisão do Comitê Central do Partido Comunista Chinês sobre a Grande Revolução Cultural do Proletariado", adotada em 8 de agosto de 1966, reproduzido em Joan Robinson, *The Cultural Revolution in China* (1969). Ver também A. Sen, *On Economic Inequality* (1973); e C. Riskin, *Maoism and Motivation: A Discussion of Work Motivation in China* (1973).

³⁷ Ver Williamson, *The Evolution of Hierarchy*, para uma análise crítica da literatura recente desta área.

barganha coletiva, o senso de solidariedade em ambos os lados pode ter limites bem definidos, e pode não se enquadrar de modo algum em uma abordagem como a do utilitarismo geral. A suposição implícita de Edgeworth, sobre a qual comentei anteriormente, de que egoísmo e utilitarismo exaurem motivações alternativas possíveis, será especialmente inútil nesse contexto. Enquanto o campo do comprometimento pode não ser tão amplo, o do comprometimento baseado no utilitarismo e outros sistemas morais universalizados pode muito bem constituir uma parte relativamente pequena deste.

VII.

A teoria econômica da utilidade, que se relaciona com a teoria do comportamento racional, é criticada algumas vezes por ter demasiada estrutura; seres humanos são supostamente mais simples na realidade. Se nosso argumento até agora estiver correto, precisamente o oposto parece ser o caso: a teoria tradicional tem *pouquíssima* estrutura. É dado *um* ordenamento de preferências para um indivíduo, e como e quando a necessidade surge, este deve refletir seus interesses, representar seu bem-estar, sumarizar sua ideia do que deveria ser feito, descrever suas escolhas de fato e seu comportamento. Pode um ordenamento de preferências fazer isso tudo? Um indivíduo assim descrito pode ser "racional" no sentido limitado de não revelar inconsistências em seu comportamento de escolha, mas, se ele não tiver utilidade para essas distinções entre conceitos bastante diferentes, ele deve ser um pouco otário. O homem *puramente* econômico está, realmente, próximo de ser um idiota social. A teoria econômica tem estado muito preocupada com esse otário racional enfeitado com a glória de sua ordem de preferência única para todos os fins. Para abriremos espaço aos diferentes conceitos relacionados ao seu comportamento,

precisamos de uma estrutura mais elaborada.

Que tipo de estrutura precisamos? É fornecido um pouco mais de espaço no topo pela importante distinção de John Harsanyi entre as preferências "éticas" e as preferências "subjetivas" de um indivíduo: "a primeira deve expressar o que este indivíduo prefere (ou melhor, preferiria) com base unicamente em considerações sociais impessoais; a segunda deve expressar o que ele realmente prefere, com base em seus interesses pessoais ou em qualquer outra base"³⁸. Essa estrutura dupla permite a distinção entre o que uma pessoa acredita ser bom do ponto de vista social e o que considera ser bom do seu próprio ponto de vista. Presumidamente, a simpatia entra diretamente na suposta preferência subjetiva, mas o papel do comprometimento fica um pouco confuso. À medida que as preferências "subjetivas" de um indivíduo são tomadas para "definir sua função de utilidade", a intenção parece ser de excluir o comprometimento desta, porém uma ambiguidade surge do fato de que estes são definidos para "expressar suas preferências no sentido pleno da palavra como elas realmente são". Seria este no sentido da escolha, ou no sentido da sua noção do próprio bem-estar? Talvez Harsanyi buscasse o último, já que preferências "éticas" são, em contraste, incumbidas de expressar "o que ele prefere apenas nos momentos possivelmente raros em que ele força uma atitude especial imparcial e impessoal em si mesmo"³⁹. Contudo, e se ele se afastar da maximização de seu bem-estar pessoal (incluindo qualquer simpatia), não por uma preocupação imparcial por todos⁴⁰, mas por um senso de comprometimento a algum grupo em particular, seja à vizinhança ou ao grupo social ao qual ele pertence? O fato é que ainda temos pouca estrutura.

Mesmo ao expressar julgamentos morais de um ponto de vista impessoal, uma estrutura *dupla* é deficiente. Certamente, uma ordem de preferência pode ser *mais* ética que outra, mas *menos*

³⁸ J. Harsanyi: "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility" (1955), p. 315.

³⁹ J. Harsanyi, 1955, p. 315-316.

⁴⁰ Observe que, para Harsanyi, "as preferências de um indivíduo satisfazem esse requisito de impersonalidade se elas indicarem qual situação social ele escolheria se não soubesse qual seria a sua posição geral na nova situação escolhida (e em qualquer uma de suas alternativas), mas tivesse chances iguais de obter qualquer uma das posições sociais existentes nessa situação, da mais alta à mais baixa" (p. 316).

que uma terceira. Precisamos de maior estrutura a respeito disso também. Propus em outro lugar — na conferência sobre “razão prática” em Bristol, 1972 — que precisamos considerar *rankings* de *rankings* de preferência para expressar nossos julgamentos morais⁴¹. Gostaria de discutir um pouco essa estrutura. Uma moralidade particular pode ser vista não apenas nos termos do ranqueamento mais moral de um *set* de ações alternativas, mas também como um ranqueamento moral de *rankings* de ações (indo muito além da mera identificação do ordenamento “mais moral” de ações). Sendo X o conjunto de alternativas e combinações de ações mutuamente exclusivas em consideração, e sendo Y o conjunto de ordenamentos dos elementos de X , um ordenamento do conjunto Y (consistindo em *rankings* de ações) será chamado de ordenamento-meta de um conjunto de ações X . É minha afirmação que um ordenamento em particular do conjunto de ações X não é articulado o suficiente para dizer muita coisa sobre dada moralidade, e um formato mais robusto é fornecido ao escolher um ordenamento-meta de ações (ou seja, um ordenamento de Y em vez de X). Obviamente, tal ordenamento-meta pode incluir, entre outros, a especificação de um ordenamento de ações em particular como o “mais moral”, mas, à medida que o comportamento de fato pode ser baseado em um compromisso entre alegações de moralidade e a busca por vários outros objetivos (incluindo autointeresse), é preciso olhar também para as posições morais relativas daqueles ordenamentos de ações que *não* são os «mais morais».

Para ilustrar, considere um conjunto X de combinações de ação alternativas e as três ordenações seguintes desse conjunto de ações X : classificação A representando meu ordenamento de bem-estar pessoal (assim, em certo sentido, representando meus interesses pessoais); classificação B refletindo meus interesses pessoais “isolados”, ignorando a simpatia (quando tal sepa-

ração for possível, o que nem sempre acontece)⁴²; classificação C em termos de quais escolhas são feitas de fato (quando tais escolhas são representáveis por um ordenamento, o que, novamente, nem sempre acontece). O ranqueamento “mais moral” M pode ser, concebivelmente, qualquer um dos ordenamentos A , B ou C ; ou, então, pode ser algum outro ordenamento bastante distinto desses três. (Este será o caso se as reais escolhas de ações não forem as “mais morais” no sentido do sistema moral em questão, e se, além disso, o sistema moral exigir o sacrifício de algum interesse próprio e do interesse próprio “isolado”). No entanto, mesmo quando algum ordenamento M for distinto de A , B e C , ele é identificado como estando no topo da mesa moral, o que ainda deixa em aberto a questão de como A , B e C podem ser ordenados entre eles mesmos. Se, para dar um exemplo específico, isto acontece de modo que a busca pelo autointeresse, incluindo o prazer e a dor da simpatia, é colocada moralmente acima da busca pelo autointeresse “isolado” (assim levando a uma coincidência parcial do interesse próprio com a moralidade), e as escolhas de fato refletem uma posição moralmente superior à busca pelo autointeresse (talvez devido a um compromisso com a direção moral), então a moralidade em questão precipita o ordenamento-meta M , C , A , B , em ordem decrescente. Isto, é claro, vai muito além de especificar que M é «moralmente melhor».

A técnica do ordenamento-meta permite uma extensão variável da articulação moral. Não se afirma aqui que um ordenamento-meta moral deve ser um ordenamento *completo* do conjunto Y , ou seja, deve ordenar completamente todas as classificações de X . Pode ser um ordenamento parcial, e suponho que muitas vezes incompleto, mas devo acreditar que, na maioria dos casos, não haverá problema em ir além da expressão limitada permitida pela especificação dupla de preferências “ética” e “subjetiva”.

⁴¹ A. Sen: “Choice, Orderings and Morality” (Oxford, 1974). Ver também a tréplica de J. Watkins e minha resposta no mesmo volume, e R. Jeffrey, *Preferences among Preferences* (1974); K. Binmore, *An Example in Group Preference* (1975); e B. Weisbrod, *Toward a State-Preference Model of Utility Function Preferences: A Conceptual Note* (1976).

⁴² Isto pressupõe certa “independência” entre os diferentes elementos que influenciam o nível de bem-estar geral, implicando alguma “separabilidade”. Ver W. Gorman, *Tricks with Utility Functions* (1975).

Os ranqueamentos de ações podem, certamente, ser ordenados também por outros motivos além de um sistema particular de moralidade: o ordenamento-meta é uma técnica geral utilizável sob interpretações alternativas de relações de ordenamento-meta. Pode ser utilizado para descrever uma ideologia específica, ou um conjunto de prioridades políticas, ou um sistema de interesses de classe. Em um contexto bastante diferente, pode fornecer o formato para expressar quais preferências que alguém teria preferido ter ("Eu queria gostar mais de comida vegetariana", ou "Eu queria não gostar tanto de fumar"). Ou pode ser usado para analisar os conflitos envolvidos no vício ("Dados meus gostos atuais, estou melhor com a heroína, mas usar heroína me leva ao vício, e eu preferiria não ter estes gostos"). A ferramenta dos ordenamentos-meta pode ser utilizada de muitas maneiras diferentes em contextos distintos.

Esta, claramente, não é a ocasião para entrar em uma análise detalhada sobre como essa estrutura mais ampla permite um entendimento melhor sobre preferência e comportamento. Uma estrutura não é, certamente, uma teoria, e teorias alternativas podem ser formuladas usando tal estrutura. Devo mencionar, entretanto, que a estrutura demanda muito mais informação do que o coletado pela observação das escolhas reais das pessoas, a qual apenas revelaria no máximo o ordenamento *C*. Isso confere um papel à introspecção e à comunicação. Para ilustrar um uso do instrumento, posso me referir a alguns resultados técnicos. Suponha que estou tentando investigar seu conceito do seu próprio bem-estar. Você primeiro especifica o ordenamento *A*, que representa sua classificação de bem-estar. Contudo, eu quero ir além e ter uma ideia da sua função de utilidade *cardinal*, ou seja, grosso modo, não

apenas qual ordenação traz mais bem-estar, mas também o quanto mais. Agora, peço que você ordene as diferentes classificações nos termos da sua "proximidade" ao seu ordenamento real de bem-estar *A*, tanto quanto um policial utiliza a técnica do reconhecimento por foto: esse é mais como ele, ou é aquele? Se as suas respostas refletem o fato de que inverter uma intensidade de preferência mais forte torna o resultado mais distante do que invertendo uma intensidade mais fraca de preferência, suas respostas satisfarão certas propriedades de consistência, e a ordem das classificações nos permitirá comparar suas *diferenças* de bem-estar entre pares. Na realidade, ao considerar classificações de ordem cada vez maiores, podemos determinar sua função cardinal de bem-estar tão precisamente quanto você desejar especificar⁴³. Não estou dizendo que esse tipo de diálogo é o melhor meio para descobrir sua função de bem-estar, mas isso ilustra que uma vez que abandonamos a suposição que observar escolhas é a única fonte de informações sobre bem-estar, um mundo totalmente novo se abre, libertando-nos das algemas informativas da abordagem tradicional.

Essa estrutura mais ampla tem muitos outros usos, por exemplo, permitindo uma análise mais clara da *akrasia* — a fraqueza da vontade — e esclarecendo algumas considerações conflitantes na teoria da liberdade, a qual tentei discutir sobre em outro lugar⁴⁴. Isso também ajuda a analisar o desenvolvimento do comportamento envolvendo o comprometimento em situações caracterizadas por jogos tais como o Dilema dos Prisioneiros⁴⁵. Esse jogo é frequentemente retratado, com justiça, como o clássico caso de fracasso da racionalidade individualista. Há dois jogadores e cada um tem duas estratégias, as quais podemos chamar de *egoísta* e *altruísta* para

⁴³ Este resultado e outros relacionados surgiram em discussões com Ken Binmore, em 1975, mas um texto conjunto projetado relatando-as ainda, infelizmente, não foi escrito. Mais trabalhos sobre isso estão, atualmente, sendo feitos também por R. Nader-Ispahani.

⁴⁴ Ver A. Sen: "Choice, Orderings and Morality" (1976); ver também A. Sen, *Liberty, Unanimity and Rights* (1976). Observe também a relevância desta estrutura ao analisar a incompletude do conceito de liberdade nos termos da habilidade de fazer o que *realmente deseja*. Comparando: "Se eu achar que não sou capaz de fazer nada ou muito pouco do que desejo, preciso apenas contratar ou extinguir meus desejos, e torno-me livre. Se o tirano (ou "persuasor oculto") conseguir condicionar seus súditos (ou clientes) a perderem seus desejos originais e abraçarem ("internalizar") a forma de vida que ele inventou para eles, ele irá, nesta definição, ter sucesso ao libertá-los." I. Berlin, *Two Concepts of Liberty* (1969).

⁴⁵ Ver R. Luce e H. Raiffa, *Games and Decisions* (1958); A. Rapoport e A. Chammah, *Prisoner's Dilemma: A Study in Conflict and Cooperation* (1965); W. Runciman e A. Sen, *Games, Justice and the General Will* (1965); N. Howard, *Paradoxes of Rationality* (1971).

tornar mais fácil de lembrar sem necessidade de muitos detalhes. Cada jogador fica melhor pessoalmente ao utilizar a estratégia egoísta, *não importa* o que o outro faça, mas ambos ficam em melhor situação se ambos escolhem a estratégia altruísta em vez da egoísta. É individualmente melhor escolher a opção egoísta: um só pode afetar suas próprias ações e não as do outro, e, dada a estratégia do outro — não importa qual —, cada jogador fica melhor sendo egoísta. No entanto, essa combinação de estratégias egoístas, que resulta do autointeresse de ambos, produz um resultado que é pior para eles do que o resultado de ambos escolherem a estratégia altruísta. Pode-se demonstrar que esse conflito pode existir mesmo se o jogo é repetido várias vezes.

Algumas pessoas acham intrigante que o autointeresse individual de cada um deveria produzir um resultado inferior para todos, porém, é claro, esse é um conflito conhecido, e tem sido discutido em termos gerais há bastante tempo. De fato, esta foi a base da famosa distinção de Rousseau entre a “vontade geral” e a “vontade de todos”⁴⁶. Todavia, o enigma, do ponto de vista do comportamento racional, está no fato de que, em situações reais, as pessoas muitas vezes não seguem a estratégia egoísta. Exemplos da vida real desse tipo de comportamento em circunstâncias complexas são bem conhecidos, mas, mesmo em experimentos controlados em condições de laboratório, pessoas jogando o Dilema dos Prisioneiros frequentemente tomam a decisão altruísta⁴⁷.

Ao interpretar esses resultados experimentais, o teórico dos jogos é tentado a atribuir isso à falta de inteligência dos jogadores: “Evidentemente, os jogadores medianos não são estrategicamente sofisticados o suficiente para entender que a estratégia DD [a estratégia egoísta] é a única estratégia racionalmente defensável, e esta deficiência intelectual os impede de perder”⁴⁸. Uma abordagem mais frutífera pode consistir em

permitir a possibilidade de que a pessoa é *mais* sofisticada do que a teoria permite, e de que ela se perguntou qual tipo de preferência ela gostaria que o outro jogador tivesse, e, em bases um tanto kantianas, considerou o caso dela própria tendo tais preferências, ou se comportando *como se* as tivesse. Esta linha de raciocínio requer que tal pessoa considere as modificações do jogo que seriam provocadas pela ação por comprometimento (no sentido de “preferência revelada”, isso pareceria *como se* ela tivesse preferências diferentes das quais teria de fato), e ela tem que avaliar normas de comportamento alternativas sob essa luz. Discuti sobre isso em outro momento⁴⁹; deste modo, devo, simplesmente, observar aqui que o instrumento do ordenamento de ordenamentos auxilia o raciocínio que envolve considerar os méritos de ter diferentes tipos de preferências (ou de agir como se os tivesse).

VIII.

Admitir o comportamento baseado em comprometimento teria, certamente, consequências de longo alcance na natureza de muitos modelos econômicos. Tentei demonstrar por que essa mudança é necessária e por que as consequências podem ser muito sérias. Muitas questões permanecem sem solução, incluindo a importância empírica do comprometimento como parte do comportamento, o qual variaria, como argumentei anteriormente, de área para área. Além disso, indiquei por que a evidência empírica para isso não pode ser baseada na mera observação de escolhas de fato, e deve envolver outras fontes de informação, incluindo introspecção e debate.

Permanece, no entanto, a questão de se esta visão do homem equivale a vê-lo como uma criatura irracional. Depende muito do conceito de “racionalidade” utilizado e muitas caracterizações alternativas existem. No sentido da *consistência* de escolha, não há motivo para acreditar que admitir comprometimento deve implicar qual-

⁴⁶ Ver Runciman e Sen.

⁴⁷ Ver, por exemplo, L. Lave, *An Empirical Approach to the Prisoner's Dilemma Game* (1962), e Rapoport e Chammah, *Prisoner's Dilemma*.

⁴⁸ Rapoport e Chammah, p. 29.

⁴⁹ A. Sen, *Choice, Orderings and Morality*. Ver também K. Baier, *Rationality and Morality*, e A. Sen, *Rationality and Morality: A Reply*; K. Baier, *The Moral Point of View* (1958); e a análise de Fred Schick, *Rationality and Sociality*.

quer distanciamento da racionalidade. Isto é, entretanto, um sentido fraco da racionalidade.

O outro conceito de racionalidade predominante na economia identifica isso como a possibilidade de justificar cada ato nos termos do autointeresse: quando ação x é escolhida pela pessoa i e ação y é rejeitada, isso implica que i espera que seus interesses pessoais sejam mais bem atendidos por x do que por y . Existem, ao que me parece, três elementos distintos nessa abordagem: primeiro, é a visão consequencialista: julgar ações somente pelas consequências⁵⁰; segundo, é uma abordagem da avaliação de ações em vez da avaliação de regras; terceiro, as únicas consequências consideradas ao avaliar ações são aquelas que atendem aos próprios interesses do indivíduo, todo o resto sendo, no máximo, um produto intermediário. É claramente possível contestar as reivindicações de cada um desses elementos como sendo uma parte necessária do conceito de racionalidade no sentido do dicionário de "o poder de ser capaz de exercitar a razão de alguém". Além disso, argumentos para rejeitar as amarras de cada um desses três princípios não são difíceis de encontrar. O caso de ações baseadas em comprometimento pode surgir da violação de qualquer um desses três princípios. O comprometimento, às vezes, está relacionado a um senso de obrigação que vai além das consequências. Às vezes, a falta de ganho pessoal em ações específicas é aceita por considerar o valor das regras de comportamento. Contudo, mesmo dentro de uma estrutura consequencialista de avaliação de atos, a exclusão de qualquer consideração, que não seja o autointeresse, parece impor uma limitação totalmente arbitrária da noção de racionalidade.

Henry Sidgwick observou a natureza arbitrária da suposição do egoísmo:

Se o Utilitarismo tiver que responder à questão "por que eu deveria sacrificar minha própria felicidade pela felicidade maior de outro?", certamente deve ser admissível perguntar ao Egoísta: "Por que eu deveria sacrificar um

prazer atual por um no futuro? Por que eu deveria me preocupar mais com meus próprios sentimentos do que com os sentimentos das outras pessoas?". Sem dúvidas, parece paradoxal ao Senso Comum perguntar por uma razão pela qual um indivíduo deveria buscar por sua própria felicidade num geral; porém não vejo como a demanda possa ser repudiada como absurdo por aqueles que adotam visões da escola extremamente empírica de psicólogos, embora comumente se suponha que essas visões tenham uma afinidade estreita com o Egoísmo Hedonista. Garanta que o Ego seja meramente um sistema de fenômenos coerentes, que o "eu" permanente idêntico não seja um fato, mas uma ficção, como Hume e seus seguidores mantêm; por que, portanto, deveria uma parte da série de sentimentos, sobre a qual o Ego está decidido, preocupar-se com outra parte da mesma série, assim como com qualquer outra série⁵¹?

A visão da racionalidade que a identifica com a avaliação de ações consequencialistas usando autointeresse pode ser questionada de qualquer um desses três ângulos. Admitir o comprometimento como parte do comportamento implica a negação da avaliação fundamentada como base para agir.

Não há muito mérito em se esforçar demasiado debatendo a definição "propícia" de racionalidade. O termo é utilizado em muitos sentidos diferentes, e nenhuma das críticas dos fundamentos comportamentais da teoria econômica apresentadas aqui depende de uma definição escolhida. A questão principal é a aceitabilidade da suposição da busca invariável pelo autointeresse em cada ação. Chamar esse tipo de comportamento racional, ou afastá-lo do irracional, não muda a relevância dessas críticas, apesar de produzir uma definição arbitrariamente estreita de racionalidade. O presente artigo não se preocupou com a questão de saber se o comportamento humano é mais bem descrito como racional ou irracional. A tese principal foi a necessidade de acomodar o comprometimento como parte do comportamento. Comprometimento não pressupõe raciocínio, porém não o exclui: na realidade, à medida que as consequências para os outros devem ser mais claramente compreendidas e

⁵⁰ Sobre a natureza do "consequencialismo" e os problemas gerados por este, ver B. Williams, *A Critique of Utilitarianism* (1973).

⁵¹ H. Sidgwick, *The Method of Ethics* (1907). Ver também a exposição contundente de Nagel da tese que "altruismo por si só depende do reconhecimento da realidade de outras pessoas, e da equivalente capacidade de considerar-se apenas um indivíduo em meio a muitos." *The Possibility of Altruism* (1970).

avaliadas nos termos dos valores e instintos de um indivíduo, o escopo do raciocínio pode muito bem expandir-se. Busquei analisar as extensões estruturais no conceito de preferências feitas necessárias pelo comportamento baseado na avaliação fundamentada do comprometimento. Preferências como ordenamentos devem ser substituídos por uma estrutura muito mais elaborada, envolvendo ordenamentos-meta e conceitos relacionados.

Ademais, argumentei contra a visão de comportamento nos termos da dicotomia tradicional entre egoísmo e sistemas de moral universalizados (tal como o utilitarismo). Grupos intermediários entre si e todos, tais como classe e comunidade, providenciam o foco de muitas ações envolvendo o comprometimento. A rejeição do egoísmo como descrição de motivação, portanto, não implica a aceitação de alguma moralidade universalizada como base do comportamento de fato. Sequer torna os seres humanos excessivamente magnânimos.

Tampouco, certamente, o uso do raciocínio implica uma notável sabedoria. "É tão verdade quanto que o nome de César seja Kaiser, Que nenhum economista jamais foi tão sábio", disse Robert Frost em um elogio brincalhão ao economista contemporâneo. Talvez, uma homenagem similarmente dúbia possa ser feita ao homem econômico em nosso conceito modificado. Se ele brilha, ele brilha em comparação — em contraste — à dominante imagem do otário racional.

Referências

ARROW, Kenneth. *Social Choice and Individual Values*. Chicago: Cowles Commission for Economics Research, 1951.

ARROW, Kenneth; HAHN, Frank. *General Competitive Analysis*. São Francisco: Holden-Day, 1971.

EDGEWORTH, Francis. *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. Londres: C. Kegan Paul and Co., 1881.

NAGEL, Thomas. *The Possibility of Altruism*. Princeton: Princeton University Press, 1970.

SAMUELSON, Paul. The Problem of Integrability in Utility Theory. *Economica*, New Series, v. 17, n. 68, p. 355-385, Nov. 1950.

SEN, Amartya. Behavior and The Concept of Preference. *Economica*, New Series, v. 40, n. 159, p. 241-259, Aug. 1973.

SEN, Amartya. Maximization and The Act of Choice. *Econometrica*, [s. l.], v. 65, n. 4, p. 745-779, July 1997.

SEN, Amartya. Rational Fools: A Critique of the Behavioral Foundations of Economic Theory. *Philosophy and Public Affairs*, [s. l.], v. 6, n. 4, p. 317-344, 1977.

Fabrizio Pontin

Doutor em Filosofia, (Southern Illinois University, 2013), Professor no Programa de Pós-Graduação em Educação e no Programa de Pós-Graduação em Direito da Universidade LaSalle.

Rafaela Lemos da Silva

Bacharela em Relações Internacionais (Universidade LaSalle, 2023).

Endereço para correspondência

FABRICIO PONTIN

Av. Victor Barreto, 2288, Centro, 92010-000
Canoas, Rio Grande do Sul, Brasil

RAFAELA LEMOS DA SILVA

Av. Victor Barreto, 2288, Centro, 92010-000
Canoas, Rio Grande do Sul, Brasil

Os textos deste artigo foram revisados por Araceli Pimentel Godinho e submetidos para validação dos autores antes da publicação.