UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Direito – Programa de Pós-Graduação

Luiz Felipe Drummond Teixeira

A IMPLEMENTAÇÃO DE *NUDGES* A POLÍTICAS PÚBLICAS: GANHOS DE BEM-ESTAR E RISCOS À AUTONOMIA

BELO HORIZONTE
2022

LUIZ FELIPE DRUMMOND TEIXEIRA

A IMPLEMENTAÇÃO DE *NUDGES* A POLÍTICAS PÚBLICAS: GANHOS DE BEM-ESTAR E RISCOS À AUTONOMIA

Dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Direito da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de mestre pela instituição.
Orientador: Prof. Dr. Renato César Cardoso.
Coorientador: Prof. Dr. Leandro Novais e Silva.

LINHA DE PESQUISA 3: História, Poder e Liberdade

ÁREA DE ESTUDO: Direito, Tecnociências e Interdisciplinaridade

BELO HORIZONTE

2022

Ficha catalográfica elaborada pela bibliotecária Meire Luciane Lorena Queiroz - CRB-6/2233.

Teixeira, Luiz Felipe Drummond
T266i A implementação de *nudges* a políticas públicas [manuscrito]: ganhos de bem-estar e riscos à autonomia / Luiz Felipe Drummond Teixeira. - 2022.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Faculdade de Direito.

1. Direito - Teses. 2. Bem-estar econômico - Teses. 3. Paternalismo - Teses. 4. Políticas públicas - Teses. 5. Autonomia - Teses. I. Cardoso, Renato César. II. Silva, Leandro Novais e. III. Universidade Federal de Minas Gerais - Faculdade de Direito. IV. Título.

CDU: 330.5

UNIVERSIDADE FEDERAL DE MINAS GERAIS

PROGRAMA DE PÓS-GRADUAÇÃO EM DIREITO

UFMG

ATA DA DEFESA DA DISSERTAÇÃO DO ALUNO
LUIZ FELIPE DRUMMOND TEIXEIRA

Realizou-se, no dia 24 de agosto de 2022, às 14:00 horas, Faculdade de Direito da UFMG, da Universidade Federal de Minas Gerais, a defesa de dissertação, intitulada *A implementação de nudges a políticas públicas: ganhos de bem-estar e riscos à autonomia*, apresentada por LUIZ FELIPE DRUMMOND TEIXEIRA, número de registro 2020652344, graduado no curso de DIREITO/DIURNO, como requisito parcial para a obtenção do grau de Mestre em DIREITO, à seguinte Comissão Examinadora: Prof(a). Renato Cesar Cardoso - Orientador (Faculdade de Direito da Universidade Federal de Mi), Prof(a). Leandro Novais e Silva (Faculdade de Direito da Universidade Federal de Mi), Prof(a). Fabio Queiroz Pereira (UFMG), Prof(a). Rute Neto Cabrita e Gil Saraiva (Faculdade de Direito da Universidade de Lisboa), Prof(a). Diogo Rosenthal Coutinho (Faculdade de Direito da Universidade de São Paulo).

A Comissão considerou a dissertação:

(X) Aprovada, tendo obtido a nota 95,0.

( ) Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.
Belo Horizonte, 24 de agosto de 2022.

Renato C. Cardoso
Prof(a). Renato Cesar Cardoso (Doutor) nota 95,0

Leandro Novais e Silva
Prof(a). Leandro Novais e Silva (Doutor) nota 95,0.

Prof(a). Fabio Queiroz Pereira (Doutor) nota 95,0.

Rute Neto Cabrita e Gil Saraiva
Prof(a). Rute Neto Cabrita e Gil Saraiva (Doutora) nota 95,0.

Prof(a). Diogo Rosenthal Coutinho (Doutor) nota 95,0.

LUIZ FELIPE DRUMMOND TEIXEIRA

A IMPLEMENTAÇÃO DE *NUDGES* A POLÍTICAS PÚBLICAS: GANHOS DE BEM-ESTAR E RISCOS À AUTONOMIA

Dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Direito da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de mestre pela instituição.

Belo Horizonte, 24 de agosto de 2022

Componentes da Banca Examinadora

_____

Professor Doutor Renato César Cardoso

Orientador

Universidade Federal de Minas Gerais

_____

Professor Doutor Leandro Novais e Silva

Coorientador

Universidade Federal de Minas Gerais

_____

Professor Doutor Fabio Queiroz Pereira

Universidade Federal de Minas Gerais

_____

Professor Doutora Rute Saraiva

Universidade de Lisboa

_____

Professor Doutor Diogo Rosenthal Coutinho

Universidade de São Paulo

À Clarissa Hees Drummond, Luiz Carlos Ângelo Teixeira e Isabel Drummond Teixeira, que tornaram este trabalho possível.

À Zoê Hees Drummond (In Memorian), que tornou este trabalho provável.

AGRADECIMENTOS

À minha mãe, Clarissa, por ser meu maior exemplo de persistência e amor e nunca colocar barreiras a nenhum dos meus sonhos. Ao meu pai, Luiz Carlos, por sempre ser a minha melhor companhia e por me ensinar a amar a música e a valorizar as amizades. À minha irmã, Isabel, por ser meu braço direito em todos os momentos, mesmo nas horas mais difíceis. Ao Luiz Carlos, por sempre me tratar como filho e vibrar com todas as minhas conquistas, e a André, Alisson e Enzzo, por serem irmãos que ganhei ao longo da vida. Ao meu cunhado Vinícius, amigo que entrou para a família e hoje é mais um irmão que tenho.

À minha avó, Zoê, que nos deixou já na reta final desta dissertação e que sempre ensinou a todos a importância de se comemorar a vida. Ao meu avô, Walter, que há tanto tempo nos faz falta. À minha avó, Lucinha, de quem herdei o gosto pela escrita, por todo o amor e por ser o meu maior exemplo intelectual. Ao meu avô, Luiz, por ser, desde a infância, a figura em quem eu me espelhei e por me ensinar a amar São João del Rei e Minas Gerais. A todos os meus tios, em especial aos meus padrinhos Míriam e Paulinho, por todo o carinho de sempre, e à minha tia Dora, que se tornou uma interlocutora essencial recentemente. A todos os meus primos, em especial à minha prima Ana Clara, por fazer com que a saudade seja sempre proporcional à passagem do tempo. À Aparecida e ao Caio, por sempre me acompanharem com carinho.

Ao meu orientador e amigo, Professor Dr. Renato César Cardoso, por sempre confiar no meu trabalho e por fazer com que as paredes dos departamentos ruíssem em busca de respostas às múltiplas perguntas que surgiam, com sua promissora e pioneira linha de pesquisas. Desde 2014, quando me tornei seu orientando, vejo a interdisciplinaridade como um caminho inevitável ao se analisar problemas jurídicos. Ao meu coorientador e amigo, Professor Dr. Leandro Novais e Silva, pela interlocução diária, pela parceria mantida já há alguns anos e pelas indicações acadêmicas, literárias e musicais que foram fundamentais à realização desta pesquisa. Grande parte dessa dissertação é fruto desse contato, ao longo da graduação, no estágio de docência e nas reuniões e mesas de bares do RAPP. À Professora Dra. Rute Saraiva, por toda a atenção e pela interlocução indispensável à realização dessa pesquisa. Aos Professores Dr. Fabio Queiroz Pereira e Dra. Mariana Alves Lara, pelas importantes lições a respeito da autonomia. Aos Professores Dr. Brunello Stancioli Souza e Dr. Leonardo Wykrota, pelas sugestões imprescindíveis na banca de qualificação. Ao Professor Dr. Diogo Coutinho, pela atenção e por permitir gentilmente que eu participasse de sua disciplina na graduação da FD/USP, em que pude ter bons insights que estão refletidos nesta dissertação.

À Camila, por debater e discordar de minhas conclusões parciais e por me mostrar o caminho que esta dissertação estava trilhando, no momento mais difícil da escrita. Ao Otávio, por todo o apoio e pelas conversas essenciais sobre esta dissertação. Ao Gabriel, meu primeiro interlocutor acadêmico, pelo constante debate sobre os caminhos desta pesquisa. Ao Pedro, por todo o companheirismo e por sempre me ensinar a valorizar as coisas que realmente importam na vida. À Gabriela, por estar comigo em todos os momentos, independentemente de sua dificuldade. Aos meus amigos inseparáveis de São João del Rei, Richarlyson, Matheus, Guilherme e Diogo, que sempre tornam maior a distância e que sempre estão comigo, apesar da minha constante ausência. Ao Augusto, que me ensina a ver a vida de forma mais lírica e que, desde a infância, me mostra os passos que devo seguir. Ao Leonardo, companheiro dos bares de Belo Horizonte e parceiro de todas as horas. Ao Lucas, amigo com quem divido a paixão pela música e que continua contando comigo, apesar das reiteradas vezes em que deixei de encontrá-lo para escrever. Ao Glacus, pelas trocas acadêmicas e literárias e por contribuir com essa dissertação desde o início. À Jade, pelo auxílio inestimável para que eu definisse o recorte que deveria seguir e entendesse qual era a importância desta pesquisa.

Aos amigos do Grupo de Estudos em Regulação, Antitruste e Políticas (RAPP), por permitirem que eu tivesse ideais essenciais à realização desta pesquisa e por sempre me motivarem. Aos amigos do Grupo de Estudos em Neuroética e Neurodireito (GENe), pelas inúmeras indicações enquanto essa pesquisa se iniciava e pela inspiração constante.

A todos os amigos do Grinberg Cordovil Advogados, que me deram todas as condições necessárias para finalizar esta dissertação. A todos os amigos do TPC Advogados, por sempre me apoiarem e incentivarem esse projeto. Ao Dr. Paulo André Rohrmann, amigo em quem me inspiro desde a época de estagiário e cujas conversas sempre me fazem colocar as coisas em perspectiva.

À Faculdade de Direito da Universidade Federal de Minas Gerais, minha segunda casa, e a todos os seus alunos, professores e servidores, que a transformam em um espaço público de excelência, que possibilitou o meu desenvolvimento intelectual e criou as possibilidades para que esta pesquisa fosse realizada.

"Ainda é tempo. É tempo. Apesar do próprio homem, ainda é tempo. Apesar dessa crosta que cultivas com amianto e medo, ainda é tempo. Apesar da reserva delicada das toneladas cegas mas perfeitas de TNT pousando sobre o centro de cada coração - ainda é tempo".

(MELLO, 1989, p. 3)

RESUMO

Ao longo da última década, *nudges* e o paternalismo libertário tornaram-se reconhecidos como uma nova via de intervenção na vida privada, com enorme potencial e baixo custo. No entanto, desde que o conceito foi criado por Richard H. Thaler e Cass R. Sunstein, a legitimidade da utilização de *nudges* se tornou objeto de uma enorme controvérsia acadêmica. De fato, diversos autores sustentaram que *nudges* não seriam legítimos em termos éticos, práticos e epistemológicos. Embora inúmeras obras a respeito de tal controvérsia tenham sido publicadas, um consenso parece ainda longe de ser atingido. À luz dessas críticas, esta dissertação avaliará em que medida *nudges* podem gerar ganhos efetivos de bem-estar, considerando a efetividade dessas medidas e as justificativas utilizadas para embasá-las, e quais os riscos que *nudges* apresentam à autonomia. Como hipótese inicial, admite-se que, em um cenário marcado por fatores supostamente irrelevantes que impactam as decisões, *nudges* podem gerar bem-estar sem riscos excessivos à autonomia, sobretudo diante de seu bom custo-benefício. Para testar essa hipótese, alguns desafios teóricos precisam ser superados, como, por exemplo, a exata definição de *nudge*, a sua diferenciação de outras técnicas comportamentais, os seus fundamentos comportamentais e os seus efeitos em larga escala. Após enfrentar esses desafios, confirma-se, com parcimônia, a hipótese levantada previamente. Em um primeiro aspecto, *nudges* podem gerar ganhos de bem-estar apesar da dificuldade de se aferir as reais preferências dos decisores, sobretudo se consideradas as distinções entre julgamentos diretos e indiretos e a possibilidade de aferir os julgamentos diretos em condições epistemologicamente favoráveis. Seus efeitos, contudo, são limitados, sobretudo quando o *nudge* for implementado por um arquiteto de escolhas privado, cujos incentivos não se alinham aos objetivos da política proposta. Além disso, *nudges* são compatíveis com atos substancialmente autônomos, desde que não imponham custos excessivos para decidir de outra forma nem sejam implementados de maneira obscura. As restrições institucionais do Estado Democrático de Direito, incluindo o controle jurisdicional, são mecanismos para evitar que *nudges* exploratórios e indevidos sejam implementados.

Palavras-chave: *Nudges*. Paternalismo Libertário. Bem-Estar. Autonomia. Economia Comportamental.

ABSTRACT

Over the last decade, nudges and libertarian paternalism have become known as a new way of intervention in private life, with huge potential and low cost. However, since the concept was created by Richard H. Thaler and Cass R. Sunstein, the legitimacy of using nudges has become the subject of an enormous academic controversy. Indeed, several authors have argued that nudges are not legitimate in ethical, practical, and epistemological grounds. Although numerous papers regarding this controversy have been published, a consensus seems still far from being reached. In light of these criticisms, this dissertation will evaluate to what extent nudges can generate effective welfare gains, considering the effectiveness of these measures and the reasons used to support them, and what risks nudges present to autonomy. As an initial hypothesis, it is admitted that, in a scenario marked by supposedly irrelevant factors that impact decisions, nudges can generate well-being without excessive risks to autonomy, especially given its cost-effectiveness. To test this hypothesis, some theoretical challenges need to be overcome, such as the exact definition of nudge, its differentiation from other behavioral techniques, its behavioral foundations, and its effects on a large scale. After facing these challenges, this dissertation confirms with parsimony the hypothesis previously raised. In a first aspect, nudges can generate welfare gains despite the difficulty of assessing the real preferences of decision-makers, especially against the distinction between direct and indirect judgments and the possibility of assessing direct judgments under epistemologically favorable conditions. Its effects, however, are limited, especially when a nudge is implemented by a private choice architect, whose incentives do not align with the objectives of the proposed policy. In addition, nudges are compatible with substantially autonomous acts, provided that they do not impose excessive costs to decide otherwise or are implemented in an obscure manner. The institutional constraints of the Democratic State of Law, including judicial control, are mechanisms to prevent exploratory and undue nudges from being implemented.

Keywords: Nudges. Libertarian Paternalism. Welfare. Autonomy. Behavioral Economics.

LISTA DE FIGURAS

FIGURA 1 - HEURÍSTICAS COMO SUBSTITUIÇÃO DE ATRIBUTOS 30
FIGURA 2 - HEURÍSTICAS-PROTÓTIPO 36
FIGURA 3 - ALGORITMO TAKE THE BEST 46
FIGURA 4 - GRAUS DE AUTONOMIA DE AÇÕES INTENCIONAIS 108
FIGURA 5 - O CONTINUUM DE INFLUÊNCIAS DE CONTROLADORAS A NÃO CONTROLADORAS 110
FIGURA 6 - O CONTINUUM DA INTENCIONALIDADE 114
FIGURA 7 - CONDIÇÕES NECESSÁRIAS A UM NUDGE NA DEFINIÇÃO ORIGINAL 129
FIGURA 8 - A DEFINIÇÃO AMPLA E AGREGADORA DE NUDGE NA OBRA DE SUNSTEIN . 133
FIGURA 9 - TRÊS DEFINIÇÕES DE NUDGE 143
FIGURA 10 - DEFINIÇÃO DE NUDGE ADOTADA NA DISSERTAÇÃO 153
FIGURA 11 - ÁRVORE DE DECISÃO RÁPIDA E FRUGAL 166
FIGURA 12 - NUDGE VERSUS SLUDGE 179
FIGURA 13 - DEFINIÇÃO DE NUDGE REVISADA COM BASE NA DISTINÇÃO NUDGE/SLUDGE 180
FIGURA 14 - RELAÇÃO ENTRE O PATERNALISMO LIBERTÁRIO E NUDGES EM ACEPÇÃO AMPLA 184
FIGURA 15 - RELAÇÃO ENTRE O PATERNALISMO LIBERTÁRIO E NUDGES EM SENTIDO ESTRITO 185
FIGURA 16 - NUDGES PURAMENTE PRIVADOS, NUDGES IMPUROS E NUDGES PUROS 193
FIGURA 17 - FATORES QUE LEVAM A QUEDAS DE TENSÃO 215
FIGURA 18 - EXEMPLO RELATÓRIO DE ENERGIA DOMÉSTICA (FRENTE E VERSO) 219
FIGURA 19 - A PERSISTÊNCIA DE EFEITOS DO RELATÓRIO DE ENERGIA DOMÉSTICA .. 221
FIGURA 20 - EXCERTOS DE FORMULÁRIO UTILIZADO PELOS BANCOS PARA O OPT-OUT 238

LISTA DE TABELAS

TABELA 1 - UM PANORAMA GERAL DAS TRADICIONAIS FALHAS DE MERCADO 75
TABELA 2 - UM PANORAMA GERAL DAS FALHAS DE MERCADO COMPORTAMENTAIS 82
TABELA 3 - NUDGES ANTES DOS NUDGES: PROPOSTAS DE INTERVENÇÃO MAPEADAS NA LITERATURA 125
TABELA 4 - NUDGES E SLUDGES: SOBREPOSIÇÕES E DISTINÇÕES SEGUNDO SUNSTEIN 177
TABELA 5 - REVISÃO BIBLIOGRÁFICA DA EFETIVIDADE DE NUDGES 206

SUMÁRIO

INTRODUÇÃO 15
1. ERROS PREVISÍVEIS E EVITÁVEIS: HEURÍSTICAS COMO ESTRATÉGIAS DE LIMITADA RACIONALIDADE ECOLÓGICA 24
1.1. O PROGRAMA DE HEURÍSTICAS E VIESES: O PARADIGMA DE DANIEL KAHNEMAN E AMOS TVERSKY 28
1.2. RACIONALIDADE ECOLÓGICA E HEURÍSTICAS RÁPIDAS E FRUGAIS: UMA ALTERNATIVA À ABORDAGEM DAS HEURÍSTICAS E DOS VIESES 37
1.2.1. Normas rasas e cegas ao contexto: as críticas normativas de Gigerenzer 37
1.2.2. Vieses cognitivos podem desaparecer: a crítica empírica de Gigerenzer 41
1.2.3. Heurísticas vagas e indefinidas: a crítica metodológica de Gigerenzer 42
1.2.4. Heurísticas rápidas e frugais e modelos mentais probabilísticos: um novo programa de pesquisas..... 43
1.3. HEURÍSTICAS SÃO ÚTEIS EM ALGUNS AMBIENTES, MAS PODEM LEVAR A ERROS QUE CAUSAM PREJUÍZO EM TERMOS DE BEM-ESTAR 47
1.3.1. A calibração como norma no julgamento de probabilidades: maximizar a precisão das crenças e evitar perdas de bem-estar 48
1.3.2. A racionalidade ecológica moderada enquanto medida de racionalidade procedimental compatível com o programa de heurísticas e vieses 53
1.3.3. Modelos inerentemente experimentais: por que os modelos mentais probabilísticos não parecem ser uma alternativa adequada 58
1.3.4. Validação empírica dos vieses comportamentais e fatores supostamente insignificantes que influenciam decisões e julgamentos 61
1.4. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 1 65
2. FALHAS DE MERCADO COMPORTAMENTAIS: NOVAS JUSTIFICATIVAS PARA A INTERVENÇÃO ESTATAL NA VIDA PRIVADA 67
2.1. AS FALHAS DE MERCADO: TRADICIONAIS MOTIVOS PARA A REGULAÇÃO DO COMPORTAMENTO PRIVADO 69
2.2. PREJUÍZOS AOS PRÓPRIOS INDIVÍDUOS EM CURTO E LONGO PRAZO: MAIS MOTIVOS PARA REGULAR O COMPORTAMENTO PRIVADO 76
2.2.1. Os problemas do modelo de preferências reveladas e as falhas de mercado comportamentais 76
2.2.2. Existem falhas de mercado que não são comportamentais? 84
2.3. DIFERENTES ESPÉCIES DE PATERNALISMOS COMPORTAMENTAIS: ENTRE PATERNALISMOS SUAVES E O PATERNALISMO COERCITIVO 87
2.3.1. Paternalismo: conceito e distinções importantes 88
2.3.2. Paternalismos comportamentais suaves: o paternalismo libertário, o paternalismo assimétrico e o paternalismo leve pragmático 94
2.3.3. O paternalismo coercitivo comportamental: fatores supostamente irrelevantes comprometem a noção de autonomia 102

2.4. ATOS SUBSTANCIALMENTE AUTÔNOMOS: NOÇÃO EMPIRICAMENTE FACTÍVEL DE AUTONOMIA..... 106
2.5. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 2 115
3. A ELABORAÇÃO E O DESENVOLVIMENTO DA DEFINIÇÃO DE *NUDGE*: PROPOSIÇÃO ORIGINAL E REFINAMENTOS CONCEITUAIS 117
3.1. *NUDGES* ANTES DOS *NUDGES*: DIFERENTES TÉCNICAS DE INTERVENÇÃO QUE EXPLORAM ELEMENTOS DO COMPORTAMENTO HUMANO 119
3.2. A ELABORAÇÃO DO CONCEITO DE *NUDGE*: UM RÓTULO COMUM ÀS DIFERENTES TÉCNICAS DE INTERVENÇÃO COMPORTAMENTAIS 126
3.3. *NUDGES* EM SENTIDO ESTRITO E DIFERENTES TIPOS DE *NUDGES*: DEBATES CONCEITUAIS EM TORNO DE *NUDGES* 134
3.4. DEFINIÇÃO DE *NUDGE* ADOTADA: ALTERAÇÃO DA ARQUITETURA DE ESCOLHAS PARA EXPLORAR FATORES SUPOSTAMENTE IRRELEVANTES E REDUZIR AS FRICÇÕES ASSOCIADAS A UMA OPÇÃO DE ESCOLHA 148
3.5. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 3 153
4. AVANÇOS TEÓRICOS E A DEFINIÇÃO DE *NUDGE*: ESTRATÉGIAS ALTERNATIVAS, RELAÇÃO COM O PATERNALISMO LIBERTÁRIO E ARQUITETURA DE ESCOLHAS PRIVADA 156
4.1. DESENVIESAMENTO, *BOOSTS*, *THINKS*, *NUDGES PLUS* E *SLUDGES*: ESTRATÉGIAS ALTERNATIVAS AOS *NUDGES*? 158
4.1.1. Técnicas de desenviesamento: a tentativa de corrigir os vieses comportamentais 158
4.1.2. Boosts e o aprimoramento das capacidades dos agentes 164
4.1.3. De thinks a nudges plus: incorporando elementos de deliberação e reflexão na tomada de decisões..... 169
4.1.4. Sludges e a criação de fricções que dificultam determinadas escolhas 174
4.2. *NUDGES VERSUS* PATERNALISMO LIBERTÁRIO: SEPARANDO A TÉCNICA DE INTERVENÇÃO DAS RAZÕES PARA SE INTERVIR 180
4.3. *NUDGES* ENQUANTO TÉCNICA DE REGULAÇÃO ESTATAL E O REARRANJO DA ARQUITETURA DE ESCOLHAS POR AGENTES PRIVADOS 186
4.4. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 4 193
5. *NUDGES* E EFETIVIDADE: FERRAMENTAS ÚTEIS, COMPLEMENTARES E COM BOM CUSTO-BENEFÍCIO PARA PROMOVER A MUDANÇA DE COMPORTAMENTO 196
5.1. REVISÃO DE LITERATURA: *NUDGES* SÃO EFETIVOS COMO FERRAMENTA PARA MUDANÇA DO COMPORTAMENTO 197
5.2. *NUDGES* PERSISTENTES E EM LARGA ESCALA: CAMINHOS PARA EVITAR QUEDAS DE TENSÃO 208
5.2.1. Evitando quedas de tensão: armadilhas que podem limitar a escalabilidade de nudges 208
5.2.2. Nudges em longo prazo: avaliando a persistência dos efeitos de nudges 215
5.2.3. Nudges em larga escala: efeitos significativos, porém menores do que o esperado 224
5.3. FORTES PREFERÊNCIAS ANTECEDENTES, O EQUILÍBRIO DA PESCARIA E REATÂNCIA PSICOLÓGICA..... 233

5.4. O PAPEL DOS *NUDGES* NO CONJUNTO DE TÉCNICAS REGULATÓRIAS: NECESSIDADE DE COMBINAÇÃO DE DIFERENTES TÉCNICAS DE INTERVENÇÃO 241
5.5. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 5 245
6. *NUDGES* QUE PROMOVEM O BEM-ESTAR SEM PREJUÍZOS À AUTONOMIA: ARGUMENTO EM FAVOR DA UTILIZAÇÃO DE *NUDGES* EM POLÍTICAS PÚBLICAS 248
6.1. CORRESPONDÊNCIA ENTRE PROCEDIMENTOS DE TOMADA DE DECISÃO, ESTRUTURA PSICOLÓGICA E AMBIENTE DE ESCOLHAS: *NUDGES* COMO FERRAMENTAS PROMOTORAS DA RACIONALIDADE ECOLÓGICA 251
6.2. OS *NUDGES* E O PROBLEMA DO CONHECIMENTO: OS JULGAMENTOS DIRETOS E OS INDIRETOS E A ANÁLISE DE CUSTOS E BENEFÍCIOS 256
6.2.1. A limitação das escolhas ativas: concorrência imperfeita, aprendizado e limitada largura de banda 257
6.2.2. O problema do conhecimento e a mensuração de bem-estar: crítica ao critério como julgado por si próprios 264
6.2.3. Para onde empurrar: como o arquiteto de escolhas pode buscar consensos? 268
6.2.4. Saldo de bem-estar e o argumento cético em favor dos nudges 278
6.3. AUTONOMIA EM XEQUE: *NUDGES* TRANSPARENTES E FACILMENTE RESISTÍVEIS COMO FERRAMENTA PROMOTORA DE ATOS AUTÔNOMOS 282
6.3.1. Os riscos apontados pelas críticas: coerção, manipulação, doutrinação e dominação 282
6.3.2. Evitando as preocupações levantadas pelas críticas: mecanismos para afastar a coerção, a manipulação e a doutrinação 288
6.4. RESTRIÇÕES INSTITUCIONAIS E *NUDGES* ILEGÍTIMOS: ESCRUTÍNIO DEMOCRÁTICO E CONTROLE DE LEGALIDADE E INCONSTITUCIONALIDADE 297
6.5. PRINCIPAIS CONCLUSÕES DO CAPÍTULO 6 305
CONSIDERAÇÕES FINAIS 308
REFERÊNCIAS 313

INTRODUÇÃO

Como o governo pode promover a alteração do comportamento das pessoas? Essa pergunta aparentemente simples é o elemento-chave do sucesso das mais diversas políticas públicas. Alguns exemplos podem ilustrar essa afirmação. Um programa de previdência privada subsidiado pelo governo depende das contribuições das pessoas para que possa gerar ganhos de bem-estar individual e coletivo. Por melhor que seja a estrutura de saúde de um determinado município e o seu programa de vacinação, os benefícios dessa política pública dependem da vacinação efetiva da população envolvida. Programas de financiamento estudantil, do mesmo modo, ainda que sejam elaborados de maneira sofisticada, só funcionarão se contarem com a participação dos estudantes. Em suma, resultados efetivos, em termos de políticas públicas, dependem da promoção de comportamentos socialmente positivos e do desestímulo a comportamentos prejudiciais à sociedade.

Na abordagem tradicional, de matriz econômica neoclássica, a resposta a essa pergunta envolveria um conjunto de medidas de comando e controle e alteração de incentivos. Se um comportamento gera prejuízos à sociedade, o governo pode proibi-lo ou, ainda, pode *exigir* que a população se comporte de outra maneira, para evitar o resultado prejudicial. Por outro lado, o governo poderia alterar a estrutura de incentivos dos agentes a partir da imposição de tributos sobre o comportamento danoso ou da instituição de subsídios ao comportamento que se quer estimular.

Ao longo da última década, diversos pesquisadores passaram a utilizar evidências das ciências comportamentais para aprimorar políticas públicas em busca de uma resposta alternativa à pergunta que inicia esta dissertação. Dentre esses pesquisadores, é notável o trabalho de Richard H. Thaler e Cass R. Sunstein (2003a, 2003b, 2009, 2021), que demonstrou a possibilidade de se direcionar as escolhas das pessoas, sem métodos coercitivos ou mudanças substanciais nos incentivos econômicos, a fim de tentar melhorar o seu próprio bem-estar. A

essa estratégia, Thaler e Sunstein deram o nome de *nudge*[1], expressão popularizada em seu livro "*Nudge: Improving Decisions About Health, Wealth, and Happiness*"[2].

Nessa proposta, para alterar o comportamento das pessoas, o governo pode reorganizar deliberadamente a arquitetura de escolhas, isto é, a forma como é organizado o contexto em que as pessoas decidem (THALER; SUNSTEIN, 2009, p. 3). Esse contexto especifica quando, se e como as pessoas devem escolher (SUNSTEIN, 2015a, p. 5) e influencia as opções que são efetivamente escolhidas pelos indivíduos. Em uma livraria, por exemplo, a forma pela qual os livros são organizados influencia a escolha que será feita pelos consumidores. Afinal, os livros mais aparentes serão escolhidos em maior proporção do que aqueles mais distantes dos olhos.

Dois exemplos de *nudges* ilustram essa abordagem. Imagine, a princípio, que um governo presuma o consentimento dos indivíduos para a adesão a determinado programa de previdência. A partir de tal presunção, se o sujeito não se posicionar ativamente para não aderir ao programa (*opt-out*), ele estará automaticamente inscrito e deverá contribuir com o programa. Outro *nudge* que pode ser mencionado é a utilização de técnicas de enquadramento (*framing*) para favorecer uma determinada opção de escolha. Para estimular a utilização de um determinado procedimento cirúrgico, por exemplo, pode-se destacar o percentual de casos bem-sucedidos, ainda que reduzido, ao invés de apresentar o percentual de intercorrências cirúrgicas.

As ciências comportamentais, por outro lado, evidenciaram que uma série de fatores tidos como irrelevantes para os modelos econômicos tradicionais impactam os julgamentos e a tomada de decisões de maneira sistemática. Como as pessoas decidem de maneira automática, com base em processos heurísticos, e têm atenção e autocontrole limitados, em algumas situações, elas não conseguem decidir de maneira adequada[3]. Mais do que isso, o mercado em equilíbrio tende a favorecer a exploração desses fatores supostamente irrelevantes por

[1] A tradução mais adequada para *nudges* seria "*pequenos empurrões*", expressão figurada que representa estímulos comportamentais, independentes de coerção e de alteração significativa da estrutura de incentivos dos indivíduos. Um *nudge* não é qualquer empurrão, já que são mais sutis. Na literatura, costumam se colocar no lado oposto de *shoves*, empurrões mais bruscos, *i.e.*, medidas mais interventivas - *e.g.* (ARNESON, 2015; SUNSTEIN, 2014b), ou, mais recentemente, de *sludges* ou fricções excessivas na tomada de decisões - *e.g.* (MILLS, 2020; SUNSTEIN, 2020a; THALER, 2018a). Além da falta de terminologia perfeitamente adequada para a tradução, a literatura em língua portuguesa tem utilizado a expressão em inglês para denominar esse tipo de intervenção (vide, por exemplo: ARARIPE; MACHADO, 2018; COSTA, 2017; OLIVEIRA; STANCIOLI, 2021; SILVA; RODRIGUES; TIBALDI, 2018). Por essas razões, neste trabalho, a expressão em inglês será utilizada. Na ausência de forma verbal adequada para representar o ato de submeter alguém a um *nudge* (*to nudge*), a forma "empurrar" também será utilizada.

[2] De agora em diante, o livro será referenciado nesta dissertação apenas como *Nudge*.

[3] Cf., por exemplo: (GILOVICH; GRIFFIN; KAHNEMAN, 2002; KAHNEMAN, 2012; TVERSKY; KAHNEMAN, 1982a).

empresas, que buscam auferir lucro a partir de um comportamento imediatista e impulsivo dos consumidores[4].

Nesse contexto, ao lado de uma nova técnica de regulação[5], a literatura também passou a descrever novos fundamentos para a regulação do comportamento privado[6], em contextos em que o arcabouço teórico neoclássico não previa intervenção alguma; nas palavras de Sunstein, esses fundamentos são falhas de mercado comportamentais. De forma geral, esses fundamentos são reunidos sob novas formas de paternalismo suave (*soft paternalism*), que, em variadas roupagens, buscam intervir em decisões privadas, com técnicas sutis como *nudges*, para gerar ganhos de bem-estar individual. Dentre eles, o paternalismo libertário de Thaler e Sunstein é o movimento mais relevante.

A popularização das ciências comportamentais e da noção de racionalidade limitada criou um certo *hype* sobre a utilização de *nudges*, o que fez com que essa técnica tenha sido prescrita como boa prática para a elaboração de políticas públicas em diversos países do mundo desde a elaboração do conceito[7]. A compreensão do comportamento humano e dos fatores supostamente irrelevantes também é essencial para o direito. Ao longo dos últimos anos, incontáveis estudos têm sido publicados na intersecção entre direito e ciências comportamentais, em quase todos os ramos do direito[8]. No Brasil, também começaram a surgir trabalhos na área[9]. Em muitas dessas pesquisas, os *nudges* são utilizados como prescrição para problemas de natureza variada. A validade desta intervenção, no entanto, costuma ser tida como

[4] Cf., por exemplo: (AKERLOF; SHILLER, 2015).

[5] O termo regulação - assim como intervenção - é utilizado nesta dissertação de maneira ampla, para representar qualquer atividade ordenatória do Estado, destinada a controlar, dirigir ou governar em conformidade com uma regra, princípio ou sistema. Trata-se de um sinônimo para a ação do Estado que busca influenciar ou controlar o comportamento (DUNNE, 2015, p. 34). Não se refere, portanto, à regulação em sentido estrito, ligada à intervenção do Estado na operação dos mercados.

[6] Cf., por exemplo: (BUBB; PILDES, 2014; CAMERER *et al.*, 2003; CONLY, 2013; LOEWENSTEIN; HAISLEY, 2008; THALER; SUNSTEIN, 2003b, 2003a).

[7] Barack Obama, por exemplo, nomeou Sunstein como chefe do Escritório de Informação e Assuntos Regulatórios da Casa Branca (*White House Office of Information and Regulatory Affairs - OIRA*). A respeito do período de Sunstein no OIRA, Cf.: (SUNSTEIN, 2014c). Thaler, por sua vez, participou da criação do *Behavioural Insights Team*, agência britânica que busca aplicar *insights* comportamentais a políticas públicas do Reino Unido. A respeito do *Behavioural Insights Team*, Cf.: (HALPERN, 2015). Mais exemplos de aplicação concreta de *nudges* com sucesso podem ser encontrados num extenso relatório da Organização para a Cooperação e Desenvolvimento Econômico (2017).

[8] Para um apanhado geral, Cf., por exemplo: (ZAMIR; TEICHMAN, 2014).

[9] Há exemplos de pesquisas em diversos ramos, como o direito processual civil (COSTA, 2018; NUNES; LUD; PEDRON, 2018; WOLKART, 2019), o direito civil (OLIVEIRA; STANCIOLI, 2021), o direito ambiental (VICTOR COELHO; AYALA, 2018), o direito do consumidor (DAURA, 2018; LEAL; RIBEIRO, 2018), o direito administrativo (ARARIPE; MACHADO, 2018; COSTA, 2017), o direito tributário (MACHADO SEGUNDO, 2018) e a proteção dos direitos humanos (BIJOS; PROLA JÚNIOR; TABAK, 2017; SILVA; RODRIGUES; TIBALDI, 2018).

pressuposta ou discutida apenas de forma breve, sem que as críticas à sua utilização sejam abordadas em profundidade[10].

Embora a utilização de *nudges* tenha se popularizado, a década que se seguiu à publicação da obra de Thaler e Sunstein é marcada na literatura pela proliferação de críticas aos *nudges* e ao paternalismo libertário, com autores apontando que essas intervenções violariam a autonomia[11], seriam antiéticas[12], manipulativas[13], ineficazes[14] e, até mesmo, que as suas premissas teóricas seriam falsas[15]. De maneira geral, a maior parte dessas críticas abalam a legitimidade normativa de *nudges*, especialmente de cunho paternalista, e criam dificuldades para a obtenção de consensos com relação à adoção de tais medidas.

Nesse contexto, à luz de todas essas críticas, esta dissertação procura respostas ao seguinte problema: os *nudges* são ferramentas capazes de gerar ganhos de bem-estar, sem impactos excessivos à autonomia, quando aplicados a políticas públicas?[16]

Certamente, analisar a enorme história filosófica de conceitos como a autonomia e o bem-estar seriam tarefas intermináveis. A análise do problema enfrentado nesta dissertação, contudo, não se debruça por todo esse longo histórico. De fato, o marco teórico adotado para enfrentar este problema é a noção, apresentada por Thaler e Sunstein (2009, 2021) de que fatores supostamente irrelevantes no ambiente em que as escolhas são feitas – como a ordem em que os alimentos são dispostos – podem alterá-las significativamente. Essa premissa parte da constatação de que as pessoas apresentam erros sistemáticos na tomada de decisões, como foi documentado nos inúmeros experimentos de Daniel Kahneman e Amos Tversky[17], assim

[10] Não se pretende dizer que o tema jamais foi abordado no Brasil, mas, apenas, que a abordagem é, ainda, incipiente. Para algumas abordagens de críticas e objeções aos *nudges* publicadas no Brasil, Cf.: (JORDÃO; BOGOSSIAN, 2021; LINS HORTA, 2017; BRACINNI NETO, 2019; PEREIRA RIBEIRO; DOMINGUES, 2021). Para uma abordagem publicada em Portugal sobre essas críticas, Cf.: (SARAIVA, 2020).

[11] Cf., *e.g.*, (ENGELEN; NYS, 2020; HAUSMAN; WELCH, 2010; MILLS, 2015; SCHMIDT; ENGELEN, 2020; VUGTS *et al.*, 2020).

[12] Cf., *e.g.*, (BOVENS, 2009; REBONATO, 2014; SAINT-PAUL, 2011; SUNSTEIN, 2015b, 2016a; WALDRON, 2014; WHITE, 2010, 2016; WRIGHT; GINSBURG, 2015).

[13] Cf., *e.g.*, (BOVENS, 2009; HANSEN; JESPERSEN, 2013; SUNSTEIN, 2015b, 2016a; WILKINSON, 2012).

[14] Cf., *e.g.*, (BURGESS, 2012; SELINGER; WHYTE, 2012; SUNSTEIN, 2019a).

[15] Cf., *e.g.*, (GIGERENZER, 2015; GRÜNE-YANOFF; HERTWIG, 2016; REBONATO, 2014).

[16] No projeto desta dissertação submetido à banca de qualificação, a proposta era analisar a *constitucionalidade dos nudges*, à luz do teste da proporcionalidade. Apesar disso, ao elaborar o trabalho, entendeu-se que esta seria uma questão de difícil avaliação abstrata, como se pretendeu fazer nesta dissertação. Nesse caso, pareceu mais relevante dar um passo atrás e direcionar o foco ao potencial de geração de bem-estar e aos riscos que os *nudges* geram à autonomia. Essas questões podem iluminar a forma como o teste de proporcionalidade pode ser aplicado em um caso concreto.

[17] Cf., por exemplo: (GILOVICH; GRIFFIN; KAHNEMAN, 2002; KAHNEMAN, 2003, 2012; KAHNEMAN; SLOVIC; TVERSKY, 1982; TVERSKY; KAHNEMAN, 1974). Para mais informações sobre a contribuição da amizade entre Tversky e Kahneman para o desenvolvimento de um paradigma pautado na racionalidade limitada e nos vieses comportamentais, Cf.: (LEWIS, 2016).

como de diversos autores das ciências comportamentais, como o próprio Thaler, George Lowenstein, Colin Camerer, Dan Ariely, entre outros, ao longo dos últimos cinquenta anos[18].

Em linha com esse marco teórico, a hipótese do trabalho para o problema proposto é que há justificativas, em termos de bem-estar e autonomia, para a *regulação sutil* do comportamento privado - seja em prol do bem-estar do próprio indivíduo, seja em prol da coletividade -, com o objetivo de corrigir os *meios* adotados pelos indivíduos para tomar as suas decisões. Em outras palavras, com a prevalência de fatores supostamente irrelevantes, de acordo com a hipótese desta dissertação, os *nudges* poderiam gerar ganhos de bem-estar, sem apresentar riscos significativos à autonomia ou prejuízos aos indivíduos de maneira geral.

Para testar essa hipótese, alguns desafios teóricos devem ser superados para a obtenção de bases normativas claras à avaliação de *nudges*. Enfrentar esses desafios também é um dos objetivos e uma das contribuições desta dissertação.

Inicialmente, o primeiro desafio teórico se refere aos fatores supostamente irrelevantes e à medida em que os processos heurísticos de tomada de decisão podem ser considerados inadequados ou impróprios. Essa discussão se liga ao Grande Debate da Racionalidade, enfrentado na psicologia cognitiva entre os programas de pesquisas de Heurísticas e Vieses, capitaneado por Kahneman e Tversky, e de Heurísticas Rápidas e Frugais, liderado pelo psicólogo alemão Gerd Gigerenzer[19]. Debater essa questão é fundamental à compreensão dos fundamentos que justificam a adoção de *nudges*. Afinal, se os processos heurísticos são adequados e funcionam de maneira satisfatória, como Gigerenzer sustenta, talvez os *nudges* não fossem assim necessários, como sugerem autores como Thaler e Sunstein.

Além disso, outra questão teórica fundamental é tentar tornar o conceito de *nudge* mais claro, em uma perspectiva estrita e alinhada à maior parte da literatura. Há, na literatura, importantes obras que se dedicam a essa tarefa[20]. Apesar disso, uma etapa posterior necessária a essa depuração do conceito de *nudge* seria situá-lo em comparação às demais estratégias que

[18] Cf., sobre o surgimento da economia comportamental: (THALER, 2015).
[19] Cf., nesse sentido: (GIGERENZER, 1991, 1996; GIGERENZER; GOLDSTEIN, 1996; GIGERENZER; SELTEN, 2002; GIGERENZER; TODD; ABC RESEARCH GROUP, 1999; KAHNEMAN; TVERSKY, 1996).
[20] Cf., por exemplo: (BALDWIN, 2014; CONGIU; MOSCATI, 2021; FRENCH, 2011; HANSEN, 2016; HANSEN; JESPERSEN, 2013; HAUSMAN, 2018; HAUSMAN; WELCH, 2010; MONGIN; COZIC, 2018; OLIVER, 2015; SAGHAI, 2013).

começam a surgir na literatura, como *boosts*[21], *thinks*[22], *sludges*[23] e *nudges plus*[24]. Afinal, a adequada diferenciação entre todas essas medidas é um pressuposto importante para comparar a sua efetividade e, por consequência, a medida em que elas são necessárias.

Ainda com relação à definição de *nudge*, verifica-se que os *nudges* e o paternalismo libertário são frequentemente tratados como sinônimos na obra de Thaler e Sunstein. Não por outra razão, Sunstein (2014a, 2015b, 2016a; 2009) afirma, em diversas oportunidades, que os *nudges* não são problemáticos porque tentam influenciar as pessoas de maneira a melhorar o seu bem-estar segundo os seus próprios critérios (*as judged by themselves*). Apesar disso, os *nudges* são uma técnica de intervenção, enquanto o paternalismo libertário apresenta razões para se intervir na vida privada. Separar técnicas de intervenção de razões para se intervir é essencial na avaliação normativa de *nudges*, já que essa técnica também pode ser empregada para finalidades *welfaristas*[25] ou, mesmo, *exploratórias*, em detrimento do próprio decisor.

Outra questão teórica relevante se refere à natureza da relação jurídica existente entre o agente responsável por implementar o *nudge*, o que Thaler e Sunstein (2009, p. 3) denominam como arquiteto de escolhas, e o decisor que será influenciado pela medida. A literatura trata quase que como sinônimos os *nudges* implementados por agentes do mercado – livremente ou por ordem do governo – e os *nudges* determinados e implementados pelo próprio governo. A natureza do arquiteto de escolhas, porém, gera desafios normativos diversos na aplicação de *nudges*. Um *nudge* privado aplicado livremente, certamente, terá o lucro como objetivo. Já os *nudges* públicos – aplicados diretamente pelo governo ou por um terceiro – precisam de mais justificativas e são implementados como política pública.

---

[21] *Boosts* são tentativas de aumentar a competência na tomada de decisões dos indivíduos, com o objetivo de desenvolver as suas habilidades e o seu conhecimento, aprimorar o conjunto disponível de ferramentas para tomada de decisões ou reestruturar o ambiente em que as escolhas são feitas, para que essas habilidades e ferramentas possam ser aplicadas com maior efetividade (HERTWIG; GRÜNE-YANOFF, 2017).

[22] Também a partir da premissa de racionalidade limitada, os *thinks* tentam apresentar oportunidades de debate de questões importantes para os indivíduos, para que eles ganhem recursos e motivação para agir (JOHN *et al.*, 2011, p. 24). O objetivo é tentar motivar uma orientação cívica a partir da oportunidade de pensar de maneiras originadas que é proporcionada pela deliberação.

[23] De acordo com Thaler (2018a, p. 431), os *sludges* seriam atividades que prejudicam o ambiente de escolhas e tornando mais difícil a realização de boas escolhas e de atividades benéficas para a sociedade. Os *sludges* podem ocorrer tanto ao desencorajar comportamentos que estão no melhor interesse do decisor, quanto ao encorajar comportamentos autodestrutivos. Como explica Sunstein (2020a, p. 4), os *sludges* são fricções que criam dificuldades à tomada de decisões, normalmente pelo acréscimo de etapas complexas aos processos necessários para que uma escolha seja feita.

[24] Os nudges *plus* são intervenções que incorporam uma estratégia reflexiva à concepção de um *nudge* (BANERJEE; JOHN, 2021, p. 2), numa tentativa de adaptar o contraste entre *nudges* – estratégias baseadas em processos intuitivos – e *thinks* – estratégias que tentam estimular a deliberação.

[25] Aqui, o bem-estar do indivíduo fica em segundo plano. O fundamento da intervenção é o aumento do bem-estar coletivo. Os *nudges* que buscam estimular doações de órgãos são um exemplo claro. Sobre essa distinção, Cf.: (SARAIVA, 2020, p. 21-24).

Do ponto de vista prático, por outro lado, avaliar ganhos de bem-estar e riscos à autonomia ligados aos *nudges* - isto é, testar a hipótese proposta – depende, essencialmente, da compreensão da sua efetividade. Apresentar um panorama geral da literatura a esse respeito também é um dos objetivos desta dissertação. Mais do que isso, é necessário pensar em políticas aplicadas na prática: implementar *nudges* em larga escala não é como testá-los em laboratório. Por essa razão, também serão avaliados os efeitos de larga escala e em longo prazo dos *nudges*, assim como alguns dos desafios à sua efetividade.

Todas essas questões, assim como a própria hipótese que será testada, serão avaliadas a partir de uma ampla revisão da literatura. Foi realizada uma análise teórico-interpretativa de dados secundários, especialmente de obras bibliográficas internacionais. A seleção foi naturalmente enviesada, mas foram privilegiados artigos publicados em periódicos, com o objetivo de documentar o atual estado da arte da literatura que discute a aplicação de *nudges*. A literatura cinzenta, isto é, trabalhos publicados fora dos meios tradicionais do mercado editorial acadêmico, como relatórios, dissertações, teses e resumos apresentados em conferências (HIGGINS *et al.*, 2019, p. 78), foi deixada em segundo plano, embora algumas obras tenham sido incluídas nesta dissertação.

A relevância das obras selecionadas foi aferida, inicialmente, em uma perspectiva qualitativa, a partir da sua inclusão no artigo "*The Ethics of Nudging*", de Sunstein (2015b), já que, neste artigo, o autor aborda as principais críticas direcionadas à aplicação de *nudges*. Também foi utilizada como referência a recente revisão dessas críticas feita por Andreas T. Schmidt e Engelen (2020). Em ambos os casos, os dados bibliográficos foram levantados a partir "da chamada amostragem por referência em cadeia (ou amostragem "bola de neve")" (GUSTIN; DIAS; NICÁCIO, 2020, p. 190). Naturalmente, outras obras foram selecionadas a partir da própria experiência prévia do autor desta pesquisa com o tema.

Há que se ressaltar, nesse sentido, que esta dissertação representa a continuidade das seguintes pesquisas prévias, das quais se originou parte da bibliografia coletada: (a) atividades do Grupo de Estudos em Neuroética e Neurodireito (GENe), coordenado pelo Professor Dr. Renato César Cardoso, entre 2014 e 2019; (b) atividades do Grupo de Estudos em Regulação, Antitruste e Políticas Públicas (RAPP), coordenado pelo Professor Dr. Leandro Novais e Silva, entre 2016 e 2022; (c) projeto intitulado "Livre-arbítrio e tomada de decisões: um diálogo entre Gazzaniga, Thaler e Sunstein", vinculado ao Programa Institucional de Iniciação Científica Voluntária e realizado entre 2016 e 2017; e (d) trabalho de conclusão de curso intitulado "Os limites dos nudges: como as críticas ao paternalismo libertário influenciam a elaboração de políticas públicas", sob a orientação do Professor Dr. Renato César Cardoso, em 2019.

Por fim, esta dissertação está estruturada da seguinte forma: o Capítulo 1 expõe as premissas comportamentais adotadas nesta dissertação, que formam o alicerce desta pesquisa. Inicialmente, serão apresentados – em caráter eminentemente descritivo – as características gerais da obra de Kahneman e Tversky e do Programa de Heurísticas e Vieses, assim como da obra de Gigerenzer e do Programa de Heurísticas Rápidas e Frugais. Em seguida, em linha com autores como Andreas T. Schmidt (2019) Jennifer M. Morton (2011), será demonstrado, de forma mais propositiva, os motivos pelos quais se entende que o Programa de Heurísticas e Vieses e os prejuízos advindos de fatores supostamente irrelevantes são compatíveis com a ideia de racionalidade ecológica.

No Capítulo 2, será demonstrado como a identificação de fatores irrelevantes levou ao desenvolvimento de novos fundamentos para a regulação do comportamento privado. Inicialmente, serão comparadas as falhas de mercado tradicionais com as falhas de mercado comportamentais, em uma perspectiva descritiva. Também em caráter descritivo, serão apresentados os diferentes tipos de paternalismo comportamental, assim como algumas noções prévias ligadas ao paternalismo, que são relevantes ao debate proposto nesta dissertação. Por fim, será apresentada a noção de autonomia adotada neste trabalho.

O Capítulo 3 se dedica à exposição do desenvolvimento do conceito de *nudge* como técnica de regulação do comportamento privado. Serão apresentadas, de forma descritiva, revisões de três momentos distintos da literatura: a) o momento que antecede a elaboração do conceito de *nudge* e que coincide com o surgimento dos paternalismos libertários; b) a elaboração do conceito de *nudge* por Thaler e Sunstein e as diferentes formas pelas quais essa técnica foi definida por esses autores; e c) os refinamentos conceituais e diferentes categorias de *nudges* identificados na literatura após a elaboração do conceito. Feito esse levantamento, o Capítulo 3 se encerra com a apresentação da definição estrita de *nudge* adotada nesta dissertação, associada à neutralidade racional.

O Capítulo 4 explora algumas questões adicionais do estágio atual da literatura a respeito dos *nudges*, para enfrentar alguns dos desafios teóricos citados nesta Introdução. Inicialmente, serão demonstrados os traços definidores de estratégias potencialmente alternativas aos *nudges*, como técnicas de desenviesamento, *boosts*, *thinks*, *nudges plus* e *sludges*, além de se explorar a distinção entre os *nudges* e essas técnicas. Em seguida, em caráter mais propositivo, os *nudges* serão diferenciados do paternalismo libertário, de forma a se evidenciar os diferentes fundamentos que podem estar por trás de um *nudge*. Finalmente, será apresentada uma distinção entre *nudges* adotados por agentes privados – livremente ou por ordem do governo – e *nudges* públicos.

No Capítulo 5, será apresentada uma revisão bibliográfica da efetividade dos *nudges*. Em um primeiro momento, a revisão bibliográfica se volta às meta-análises e revisões sistemáticas de experimentos que implementaram os *nudges*. Em seguida, serão discutidos os desafios à implementação dos *nudges* em larga escala e em longo prazo. Também serão avaliadas questões adicionais que podem resultar na inefetividade dos *nudges*, como fortes preferências antecedentes, incentivos aos agentes do mercado para prejudicar a efetividade do *nudge* e a reatância psicológica. O Capítulo 5 se encerra com um argumento pela utilização de *nudges* como uma ferramenta complementar no arcabouço regulatório, sem a necessidade de funcionar como alternativa às técnicas tradicionais e a outras técnicas comportamentais.

O Capítulo 6, então, enfrenta a discussão normativa que envolve a aplicação de *nudges*. Dois conjuntos mais amplos de críticas à utilização de *nudges* serão enfrentados. Inicialmente, com relação aos fundamentos que justificam a implementação de um *nudge*, dois aspectos serão discutidos: a) a necessidade de utilização de *nudges* em um contexto de racionalidade ecológica e de problemas econômicos estruturais; e b) a legitimidade de aplicação de *nudges* paternalistas diante do problema do conhecimento. Em seguida, serão discutidas as críticas aos *nudges* ligadas à autonomia, segundo três critérios distintos: a) liberdade de escolhas; b) agência; e c) autodeterminação. Por fim, serão apresentadas breves considerações a respeito de questões institucionais que envolvem a aplicação de *nudges*, como a possibilidade de que os arquitetos de escolha errem ou estejam mal-intencionados.

Por fim, nas Considerações Finais, será demonstrado por que a hipótese levantada previamente foi confirmada com parcimônia. A análise das críticas à utilização dos *nudges*, de um lado, e da sua efetividade, de outro lado, indica que *nudges* podem gerar ganhos modestos de bem-estar, com um custo-benefício aparentemente significativo. Os *nudges*, no entanto, não são nenhuma bala de prata e os seus efeitos são limitados, sobretudo quando o *nudge* específico for implementado por um arquiteto de escolhas, cujos incentivos não se alinham aos objetivos da política proposta. Além disso, por mais que sejam sutis, os *nudges* podem impor custos excessivos para que se possa decidir de outra forma e podem ser implementados de maneira obscura. Em tais casos, eles podem não ser compatíveis com atos substancialmente autônomos, de modo que esses elementos devem ser evitados em sua aplicação, para que a sua legitimidade normativa seja mantida. Finalmente, embora *nudges* exploratórios e indevidos possam ser implementados, as restrições institucionais do Estado Democrático de Direito, incluindo o controle jurisdicional de tais *nudges*, são ferramentas para suavizar esses problemas.

1. ERROS PREVISÍVEIS E EVITÁVEIS: HEURÍSTICAS COMO ESTRATÉGIAS DE LIMITADA RACIONALIDADE ECOLÓGICA[26]

"*A consciência é um farol aceso à frente da locomotiva.*"

Boris Pasternak
(2017, p. 80)

"Can not reasoning be both rational and psychological?"

Gerd Gigerenzer e Daniel Goldstein
(1996, p. 666)

Em seu clássico "Ensaio filosófico sobre as probabilidades", o matemático, astrônomo e físico francês Pierre Simon, o Marquês de Laplace (1951), pede ao leitor que considere, por um instante, uma inteligência capaz de compreender todas as forças que animam a natureza e as respectivas situações dos seres que a compõe. Para essa inteligência, nada seria incerto, já que o futuro, assim como o passado, estaria à sua disposição. A mente[27] humana, em contrapartida, não tem tamanha capacidade. Sem a possibilidade de conhecer todos os aspectos do mundo com a certeza da superinteligência descrita por Laplace, os meros mortais baseiam-se em inferências incertas (GIGERENZER; TODD, 1999, p. 8).

Embora essa superinteligência não seja concebível em sistemas reais, seu espírito sobrevive em muitas disciplinas e teorias que visualizam seres humanos divinos (GIGERENZER, 2004), ou, na expressão de Gerd Gigerenzer e Peter M. Todd (1999), demônios laplaceanos. De fato, em inúmeros exemplos concretos, visões irrealistas do comportamento humano podem ser encontradas. Em economia, tradicionalmente, adota-se como pressuposto a ideia de que as pessoas são agentes racionais, isto é, dadas as oportunidades disponíveis, as pessoas fazem o melhor para alcançar os seus objetivos de maneira sistemática e objetiva (MANKIW, 2013, p. 6). Na abordagem econômica neoclássica de Gary Becker

[26] Algumas passagens deste Capítulo, em especial as Seções 1.1 e 1.2 foram desenvolvidas no artigo "Entre heurísticas rápidas e frugais e vieses cognitivos: diferentes abordagens de racionalidade limitada", escrito como trabalho de conclusão de curso da disciplina Temas de Antropologia Jurídica: Neuroética e Neurodireito, oferecida pelos Professores Drs. Renato César Cardoso, Brunello de Souza Stancioli e Leonardo Wykrota.

[27] Por mente, neste trabalho, não se pretende adotar qualquer tipo de perspectiva dualista. O termo mente é utilizado para representar os processos e estados mentais que surgem a partir de processos cerebrais.

(1976, p. 14), por exemplo, todo o comportamento humano é visto como se envolvesse participantes que maximizam sua utilidade a partir de um rol estável de preferências. Em uma perspectiva filosófica, Immanuel Kant (2013, p. 19-20), por exemplo, descreve o arbítrio humano como aquele que pode ser afetado, mas não determinado, pelos impulsos e que pode ser determinado às ações por uma vontade pura, isto é, pela faculdade da apetição que tem por fundamento interno a própria razão do sujeito, livre da influência dos impulsos.

Segundo Gigerenzer e Todd (1999, p. 9), essas teorias irrealistas do comportamento humano deixam de levar em consideração as limitações de tempo, conhecimento e capacidades computacionais de humanos reais. Adotam a visão de uma racionalidade ilimitada (*unbounded rationality*). Nessa abordagem, para resolver problemas de julgamento, as pessoas buscariam informações indefinidamente até que pudessem calcular qual seria a estratégia racional a ser adotada. Os proponentes dessa visão, segundo Gigerenzer e Todd, até chegam a admitir que sua abordagem exigiria habilidades mentais irrealistas. No entanto, defendem-na mesmo assim ao argumento de que, em uma perspectiva normativa, humanos deveriam agir *como se* tivessem racionalidade ilimitada.[28]

A realidade, no entanto, é mais complexa, e as capacidades do ser humano são mais modestas do que preveem as teorias que adotam a racionalidade ilimitada como pressuposto. Com efeito, "[n]osso cérebro funciona principalmente no piloto automático, e a mente consciente tem pouco acesso à fábrica gigantesca e misteriosa que funciona por baixo dela" (EAGLEMAN, 2012, p. 13). Como afirma Leonard Mlodinow (2013, p. 29), "nós não somos como os computadores, que processam dados de uma forma relativamente direta e calculam o resultado". Na verdade, "nosso cérebro compreende uma coleção de módulos que funcionam em paralelo, como interações complexas, a maior parte delas operando fora da consciência" (MLODINOW, 2013, p. 29-30). Em outras palavras, para fazer menção à frase de Iuri Andreievitch Jivago - o Doutor Jivago do clássico de Pasternak citado em epígrafe -, a consciência não é mais do que o farol de uma locomotiva: ilumina parte do caminho que está sendo percorrido, sem, contudo, esclarecer todos os aspectos que fazem a locomotiva se movimentar.

A compreensão de processos cognitivos rápidos, muitas vezes baseados em elementos emocionais e heurísticos, levou ao surgimento de abordagens que buscam uma visão mais

[28] Milton Friedman (1966, p. 14), por exemplo, afirma que as teorias mais significativas normalmente apresentam premissas extremamente irrealistas. Para que uma hipótese seja importante, ela deveria ser capaz de abstrair os elementos comuns e cruciais das complexas circunstâncias que circundam o fenômeno a ser explicado e permitir previsões válidas baseadas nos elementos abstraído. Assim, a conformidade das premissas com a realidade não seria algo importante para a validade de uma teoria. A atenção deveria se voltar para o seu poder explicativo.

realista do comportamento. Essas abordagens, é necessário ressaltar, não são propriamente novas, já que retomam ideias expostas em obras clássicas do Iluminismo, como a Teoria dos Sentimentos Morais de Adam Smith (2015), em que são identificadas explicações psicológicas para o comportamento humano, com base em sentimentos como a compaixão. De todo modo, durante boa parte do século XX, essa visão *smithiana* foi preterida, enquanto a sua famosa passagem acerca do autointeresse de açougueiros, cervejeiros e padeiros[29] serviu como uma das bases para o desenvolvimento do arquétipo do *Homo economicus*[30].

Coube ao polímata Herbert Simon (1955, p. 99) a reintrodução da noção de racionalidade limitada na economia. Simon defendia que os agentes decisores não teriam informações e capacidade computacional suficientes para decidir da forma prevista pelos modelos racionais da economia tradicional. Por isso, em vez de modelos de racionalidade global, Simon (1955, p. 110) introduziu um modelo de racionalidade limitada (*bounded rationality*)[31] que buscava se aproximar o máximo possível da tomada de decisões real.

Nos anos 1970, a amizade e a colaboração entre dois psicólogos israelenses fizeram surgir uma nova abordagem de pesquisas que, seguindo os passos de Simon[32], viria a revolucionar a forma como a tomada de decisões é compreendida nos mais diversos ramos da ciência[33]. De fato, a partir de uma indagação a respeito da possibilidade de as pessoas serem dotadas de uma percepção intuitiva para princípios básicos da estatística, Amos Tversky e Daniel Kahneman estudaram e documentaram vieses de pensamento intuitivo em tarefas variadas (KAHNEMAN, 2012, p. 13-15). Nesse contexto, surgiu o Programa de Heurísticas e Vieses. A partir da obra de Tversky e Kahneman, a noção de racionalidade limitada se popularizou e contribuiu, inclusive, para a formação de um novo ramo da economia: a economia comportamental[34].

---

[29] Em uma das mais célebres passagens de "A Riqueza das Nações: Investigação sobre a sua Natureza e as suas Causas", Adam Smith (1996, p. 74) afirma: "Não é da benevolência do açougueiro, do cervejeiro ou do padeiro que esperamos nosso jantar, mas da consideração que eles têm pelo seu próprio interesse. Dirigi-mo-nos não à sua humanidade, mas à sua auto-estima, e nunca lhes falamos das nossas próprias necessidades, mas das vantagens que advirão para eles."

[30] O *Homo economicus* é o arquétipo utilizado pela economia neoclássica para descrever o indivíduo racional maximizador, cujas preferências são ordenadas a partir da utilidade que se espera auferir com cada opção de escolha.

[31] No sentido proposto por Simon (1955), as pessoas têm racionalidade limitada, uma vez que suas capacidades computacionais não são perfeitas e irrestritas, e que, ainda que fossem, as pessoas não conseguiriam ter acesso a todas as informações que permitiriam uma escolha racional.

[32] Em entrevista para o *podcast Conversations with Tyler*, produzido pelo economista Tyler Cowen (2018), Kahneman afirmou não ter sido inspirado diretamente por Simon, isto é, originalmente, no momento em que elaborava o programa de heurísticas e vieses com Amos Tversky, Daniel Kahneman não seguia intencionalmente os passos de Simon. Apesar disso, Kahneman reconhece que, de forma indireta e de maneira retrospectiva, a sua pesquisa continuava a linha apresentada inicialmente por Simon.

[33] Vide nota 17 *supra*.

[34] Sobre o surgimento da economia comportamental, Cf.: (THALER, 2015).

Essa visão, porém, não é unanimidade na literatura da psicologia cognitiva. Com efeito, a partir de meados dos anos 1980, uma torrente de estudos capitaneada pelo proeminente psicólogo alemão Gigerenzer passou a questionar a abordagem das heurísticas e dos vieses segundo critérios empíricos, metodológicos e normativos[35]. A partir dessas críticas, Gigerenzer e seus seguidores desenvolveram um novo programa de pesquisas, que enfatiza a *importância* da utilização de heurísticas simples para fazer boas inferências e o caráter ecológico da racionalidade[36].

O embate entre, de um lado, o programa de heurísticas e vieses de Kahneman e Tversky e, de outro lado, o programa de heurísticas rápidas e frugais de Gigerenzer e seus seguidores é conhecido como "O Grande Debate da Racionalidade" (*The Great Rationality Debate*) (KELMAN, 2011; STANOVICH; WEST, 2000). Esse debate não é uma simples minúcia restrita às cadeiras da psicologia cognitiva. De fato, a opção por um ou outro programa de pesquisas implica a adoção de diferentes modelos de comportamento humano. Esses modelos, por sua vez, podem estar no cerne das justificativas para um maior nível de intervenção estatal na vida privada – tema que será mais bem abordado no Capítulo 2.

A ideia deste Capítulo é expor quais são as premissas comportamentais adotadas nesta dissertação. Estas premissas formam o alicerce desta pesquisa, uma vez que, como já demonstrado na Introdução, o marco teórico adotado é a noção, apresentada por Thaler e Sunstein (2009), de que fatores supostamente irrelevantes do ambiente em que escolhas são feitas podem alterá-las significativamente. Essa noção parte da constatação de que as pessoas apresentam erros sistemáticos na tomada de decisões, o que poderia justificar intervenções para corrigir esses erros e aprimorar suas escolhas. É fundamental, portanto, expor em detalhes como esses erros são formados e qual é a sua extensão, de acordo com a perspectiva teórica adotada na pesquisa.

Para isso, além de expor os caracteres gerais da obra de Kahneman e Tversky e do Programa de Heurísticas e Vieses, o Capítulo também apresentará as críticas formuladas por Gigerenzer e se posicionará minimamente em meio a essa controvérsia. Esta abordagem se justifica na medida em que, embora o trabalho de Kahneman e Tversky seja bem divulgado em língua portuguesa, a obra de Gigerenzer e de seus seguidores ainda é pouco divulgada[37],

[35] Essa divisão das críticas de Gigerenzer foi proposta originalmente por Peter B. M. Vranas (2000).
[36] Para uma visão geral dos diversos aspectos que compõem esse programa de pesquisas, Cf.: (GIGERENZER; SELTEN, 2002; GIGERENZER; TODD; ABC RESEARCH GROUP, 1999).
[37] Nesse ponto, não se pode deixar de mencionar a Dissertação de Doutoramento em Psicologia Cognitiva de Miguel Oliveira (2005), defendida na Universidade de Coimbra, sob a orientação de Luísa Morgado e do próprio Gerd Gigerenzer. Em caráter de divulgação, Cf., ainda, (TEIXEIRA, 2017).

especialmente no âmbito do direito. Evidentemente, não é pretensão deste trabalho tentar adotar uma posição taxativa em relação ao Grande Debate da Racionalidade, mas apenas apresentar o debate e deixar clara a opção teórica escolhida nesta dissertação.

1.1. O programa de heurísticas e vieses: o paradigma de Daniel Kahneman e Amos Tversky

Em 1974, Kahneman e Tversky publicaram na *Science* o seminal trabalho "*Judgment under Uncertainty: Heuristics and Biases"* (TVERSKY; KAHNEMAN, 1974), que registrava os resultados de suas pesquisas a respeito das percepções intuitivas das pessoas a respeito de problemas de probabilidade. Naquela oportunidade, Tversky e Kahneman (1974, p. 1124) demonstraram que muitas decisões a respeito da probabilidade de eventos incertos são tomadas com base em um número limitado de princípios heurísticos, que reduzem tarefas complexas de avaliação de probabilidades e previsão de valores a simples operações de julgamento.

Embora essas heurísticas sejam bastante úteis, Tversky e Kahneman constataram que elas às vezes levam a erros sistemáticos e severos, também denominados vieses cognitivos[38]. Esses vieses não podem ser atribuíveis a efeitos motivacionais como *wishful thinking* ou à distorção dos julgamentos em razão de recompensas ou punições (TVERSKY; KAHNEMAN, 1974, p. 1130); na verdade, alguns dos vieses descritos pelos autores aconteciam mesmo em face de incentivos e recompensas que direcionavam os participantes às respostas corretas. A utilização das heurísticas e a prevalência dos vieses, além disso, não se limitam às pessoas leigas. De fato, Tversky e Kahneman (1974, p. 1130) verificaram que mesmo *experts* tendem a cometer erros sistemáticos quando pensam intuitivamente.

Para entender como as heurísticas operam, uma primeira noção prévia importante é a ideia de que há diferentes conjuntos de processos cognitivos que podem ser diferenciados conforme a velocidade em que operam, a capacidade de controle a que estão submetidos e o

[38] Os vieses cognitivos descritos por Tversky e Kahneman são violações de regras de probabilidade subjetiva, que compõem o que os autores denominam a moderna teoria da decisão, segundo a qual a probabilidade subjetiva deve ser vista como a opinião quantificada de uma pessoa idealizada (TVERSKY; KAHNEMAN, 1974, p. 1130). Nessa abordagem, de maneira mais específica, a probabilidade subjetiva de determinado evento é vista como o conjunto de apostas que uma pessoa está disposta a aceitar sobre esse evento. Uma medida de probabilidade subjetiva internamente consistente ou coerente pode ser derivada para um indivíduo, se as suas escolhas entre essas apostas satisfizerem os axiomas dessa teoria. Por exemplo, em um problema de inferência bayesiana, em que a avaliação da probabilidade de se obter um resultado a partir de uma população específica deve ser realizada, o tamanho da população analisada, isto é, taxa-base, deve ser considerado. Se a taxa-base é negligenciada, há um viés cognitivo. Além da consistência interna, Tversky e Kahneman (1974, p. 1130) mencionam que, segundo a moderna teoria da decisão, julgamentos adequados ou *racionais* também devem ser compatíveis com a rede de crenças mantida pelo indivíduo.

conteúdo que processam (KAHNEMAN; FREDERICK, 2002). Adotando a terminologia cunhada por Keith E. Stanovich e Richard F. West (2000), Kahneman e Shane Frederick (2002) diferenciam esses processos entre o intuitivo Sistema 1 e o reflexivo Sistema 2[39]. Enquanto o primeiro é automático, associativo, rápido, atua com processos opacos, exige pouco esforço e é habilidoso, o último é controlado, dedutivo, lento, consciente, exige um considerável esforço e trabalha com base na aplicação de regras (KAHNEMAN; FREDERICK, 2002, p. 51).

Os julgamentos do Sistema 1 baseiam-se em impressões sensoriais a respeito do objeto que está sendo julgado que, no entanto, não são voluntárias ou explícitas verbalmente, apesar de os julgamentos do Sistema 1 serem marcados por intencionalidade e expressão verbal (KAHNEMAN, 2003, p. 699). Na visão de Kahneman e Frederick (2002, p. 51), a partir de tais impressões sensoriais, o Sistema 1 propõe rapidamente respostas intuitivas para problemas de julgamento, enquanto o Sistema 2 monitora a qualidade dessas propostas, podendo endossá-las, corrigi-las ou superá-las[40]. Um julgamento intuitivo, nessa abordagem, é expresso abertamente apenas se for endossado ou mantido com poucas modificações pelo Sistema 2. Esses diferentes conjuntos de processos cognitivos atuam ao mesmo tempo, de modo que as operações cognitivas automáticas e controladas competem pelo controle das respostas que são expressas pelo agente[41]. Em geral, o Sistema 2 parece endossar os julgamentos intuitivos com frequência, de maneira praticamente casual (KAHNEMAN; FREDERICK, 2002). Afinal, o monitoramento

[39] Existem inúmeras teorias que tentam explicar os diferentes processos da cognição com base em sistemas duais. Jonathan Haidt (2012), por exemplo, utiliza uma metáfora segundo a qual o sistema automático é um elefante, enquanto o sistema analítico é o condutor do elefante, cujo trabalho é servir ao elefante. Joshua Greene (2009, p. 581), por outro lado, apresentou uma teoria de processos duais nos julgamentos morais, em que julgamentos deontológicos são guiados por respostas emocionais automáticas, e julgamentos utilitários, por processos cognitivos controlados. Para uma revisão bibliográfica bastante interessante da utilização de teorias de processos duais na economia comportamental e na neuroeconomia, Cf.: (GRAYOT, 2020).

[40] Em linha com essa hipótese, Boaz Keysar, Sayuri L. Hayakawa e Sun Gyu An (2012) identificaram que utilizar uma língua estrangeira pode reduzir viés na tomada de decisão. Segundo os autores, é provável que o *efeito da língua estrangeira* seja provocado por um aumento da distância psicológica e da promoção da deliberação. Decisões tomadas em língua estrangeira podem reduzir a reação emocional dos agentes, o que, por sua vez, leva a decisões menos sistemáticas e menos enviesadas. Em outros termos, pode ser que, ao utilizar uma língua estrangeira, as pessoas acabem decidindo com base no Sistema 2, enquanto decisões em língua nativa são feitas a partir do Sistema 1.

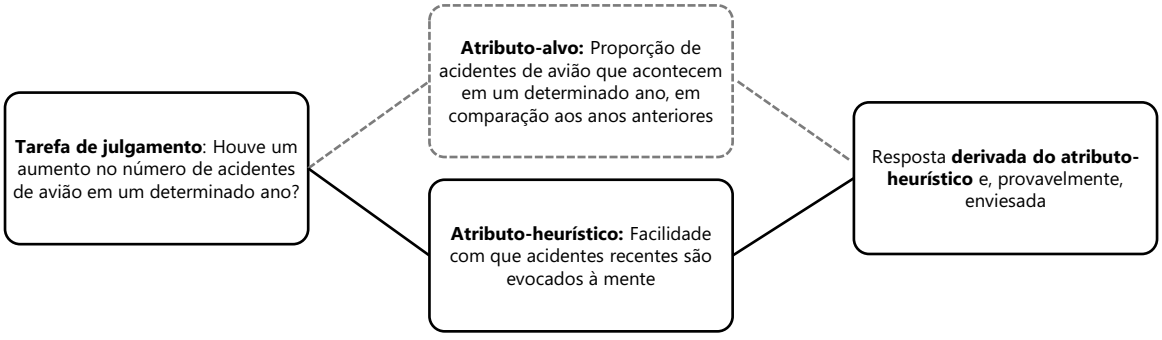
[41] Nesse ponto, é preciso ressaltar que o Sistema 1 e o Sistema 2 são ficções utilizadas para explicar diferentes tipos de processos cognitivos. Não se trata de uma divisão morfológica do cérebro ou de sistemas que efetivamente interagem entre si. Como explica Kahneman (2012, p. 40): "O Sistema 1 e o Sistema 2 são tão centrais à história que conto neste livro que devo deixar absolutamente claro que se tratam de personagens fictícios. Sistemas 1 e 2 não são sistemas no sentido clássico de entidades com aspectos ou partes que interagem. E não há nenhuma parte do cérebro que um ou outro sistema chamaria de lar. Você pode muito bem perguntar: Qual a finalidade de introduzir personagens fictícios com nomes horríveis num livro sério? A resposta é que os personagens são úteis devido a determinadas peculiaridades de nossas mentes, a sua e a minha. Uma sentença é compreendida mais facilmente se descreve o que um agente (o Sistema 2) faz do que se descreve o que determinada coisa é, quais propriedades ela tem. Em outras palavras, o 'Sistema 2' é um sujeito melhor para uma sentença do que a expressão 'aritmética mental'".

exercido por tais processos cognitivos é leve, já que a supressão de respostas intuitivas equivocadas é custosa e sua eficácia é reduzida pelo stress e por distrações.

A partir da interação entre as respostas intuitivas surgidas do Sistema 1 e o monitoramento intermitente realizado pelo Sistema 2, é possível explicar como as heurísticas surgem para mediar os julgamentos das pessoas. Como Kahneman e Frederick (2002) salientam, um julgamento é mediado por uma heurística quando um indivíduo avalia um atributo-alvo (*target attribute*) específico de um objeto que está sendo julgado e, na tentativa de obter a resposta procurada, substitui o atributo-alvo por um atributo-heurístico (*heuristic attribute*), isto é, por outra propriedade do objeto avaliado que é evocada à mente de maneira mais rápida. O julgamento mediado por heurística, nessa abordagem, é, portanto, um processo de substituição de atributos. Como o atributo-alvo é diferente do atributo-heurístico, a substituição de um pelo outro introduz vieses sistemáticos.

Um exemplo clássico do programa de pesquisas em heurísticas e vieses permite ilustrar o processo de substituição de atributos. Imagine que um sujeito quer avaliar a proporção de acidentes de avião que acontecem em um determinado ano, em comparação aos anos anteriores. O demônio *laplaceano*, com sua racionalidade ilimitada, buscaria integrar todos os dados disponíveis sobre voos e acidentes, discriminados anualmente, para tentar concretizar o cálculo mencionado. Humanos reais, na abordagem das heurísticas e dos vieses, fariam diferente. A resposta a esse problema costuma ser obtida pelo sujeito a partir da resposta a outra pergunta: a facilidade com que acidentes de avião vêm à mente. Dessa forma, após um acidente com ampla cobertura pela mídia, a resposta dada pelo sujeito acabará superestimada, já que acidentes passam a ser evocados mais rapidamente à mente.

Figura 1 - Heurísticas como substituição de atributos



Fonte: Elaboração do autor.

Nem sempre, entretanto, os julgamentos acontecerão por substituição de atributos. Em determinadas situações, a busca por um valor razoável para julgar um *atributo-alvo* pode se encerrar quase imediatamente, já que os valores podem estar na memória e ser facilmente acessados (KAHNEMAN; FREDERICK, 2002). Retomando o exemplo anterior, se o sujeito que faz o julgamento atua como controlador de voos, pode ser que acerte a proporção de acidentes em um determinado ano quase instantaneamente. Em outros casos, como explicam, mais uma vez, Kahneman e Frederick (2002), a busca por um valor acaba evocando o valor de outros atributos conceitualmente ou associativamente relacionados ao atributo-alvo. É o caso, por exemplo, de uma questão acerca da felicidade de um indivíduo de maneira geral, que pode ser respondida com base na medida em que ele está feliz num domínio particular de sua vida.

Quando, então, os julgamentos são mediados pela substituição de atributos? Na visão de Kahneman e Frederick (2002), esses julgamentos heurísticos acontecem quando o atributo-alvo é avaliado com base no valor de outro atributo numa escala alvo. No caso da avaliação dos acidentes de avião, a proporção de acidentes (atributo-alvo) é julgada a partir da facilidade com que acidentes de avião vêm à mente (atributo-heurístico). Na prática, julgamentos serão controlados por esse processo quando três condições forem satisfeitas: a) o atributo-alvo for razoavelmente inacessível; b) um atributo semântica e associativamente relacionado é altamente acessível e c) a substituição do atributo-heurístico no julgamento não é rejeitada pelas operações críticas do Sistema 2. Como o controle dos processos intuitivos pelo Sistema 2 é custoso, a substituição de atributos é confirmada com frequência. Afinal, as pessoas não estão acostumadas a pensar excessivamente e acabam se contentando com um julgamento intuitivo e plausível que venha rapidamente à mente.

Não é o objetivo deste trabalho pormenorizar as inúmeras heurísticas já catalogadas e os vários tipos de vieses que as acompanham[42]. De qualquer forma, para a compreensão da teoria e, em seguida, das críticas que lhe são direcionadas, é pertinente expor com maiores detalhes as heurísticas originalmente descritas por Tversky e Kahneman, assim como alguns outros aspectos relevantes de sua abordagem.

A primeira heurística a ser destacada é a representatividade. Como Tversky e Kahneman (1974, p. 1124) descreveram, muitas questões de probabilidade buscam aferir se um objeto A pertence a uma classe B. Ao responder perguntas dessa natureza, as pessoas normalmente baseiam-se na heurística da representatividade e avaliam a probabilidade a partir da medida em

[42] Para um panorama geral das heurísticas e dos vieses, Cf.: (ARIELY, 2008, 2010; CAMERER; LOEWENSTEIN; RABIN, 2004; DHAMI, 2016; GILOVICH; GRIFFIN; KAHNEMAN, 2002; KAHNEMAN, 2012; KAHNEMAN; SLOVIC; TVERSKY, 1982; THALER, 2015).

que A é representativo de B. Em outras palavras, a semelhança (atributo-heurístico) entre um objeto e a classe considerada é utilizada para avaliar a probabilidade (atributo-alvo) de que o objeto pertença à mencionada classe.

A dependência da representatividade, no entanto, leva a erros previsíveis de julgamento, já que a lógica da representatividade é diferente da lógica da probabilidade (TVERSKY; KAHNEMAN, 1982a, p. 89). O contraste mais marcante entre elas surge na avaliação de eventos compostos (*compound events*). De acordo com as leis da probabilidade, a especificação só poderia reduzir a probabilidade de um evento, isto é, a probabilidade de uma pessoa ser um Republicano e um artista deve ser menor do que a probabilidade dessa pessoa ser apenas um artista (TVERSKY; KAHNEMAN, 1982a, p. 90). Trata-se da chamada regra da conjunção[43]. Em julgamentos por representatividade, no entanto, essa regra é constantemente violada.

O exemplo clássico que Tversky e Kahneman (1982a) utilizam para ilustrar essa violação – é conhecido como "o problema de Linda". Em linhas gerais, os autores construíram traços de uma personalidade que representava uma mulher ativa no movimento feminista[44] e pediam que os participantes apontassem a probabilidade de oito resultados distintos sobre a vida dessa personagem, incluindo um cenário representativo – Linda ser do movimento feminista –, um cenário não representativo – Linda trabalhar como caixa de um banco – e um resultado composto pela união desses dois cenários. A maior parte dos participantes do experimento violava a regra da conjunção ao fazer tal avaliação: a probabilidade de Linda ser feminista e caixa bancária era considerada superior à probabilidade de ela apenas ser caixa bancária.

Também ilustra os julgamentos por representatividade a chamada insensibilidade ao tamanho da amostra (TVERSKY; KAHNEMAN, 1974, p. 1125). De acordo com os princípios da probabilidade, a taxa-base de uma população específica deve ser considerada para fazer julgamentos da probabilidade de um determinado resultado acontecer com ela. O problema clássico dos táxis, apresentado por Tversky e Kahneman (1982b), apresentava as seguintes informações para os participantes: a) houve uma colisão envolvendo um táxi, que fugiu do local do acidente; b) duas companhias de táxi, a Verde e a Azul, atuam na cidade; c) 85% dos táxis da cidade são Verdes e 15% são Azuis e d) uma testemunha identificou o táxi como Azul, sendo que o seu depoimento é confiável em 80% do tempo (KAHNEMAN, 2012, p. 211). Os autores,

[43] Em outros termos, de acordo com a regra da conjunção, $P\ (A\ \&\ B) \leq P\ (B)$.
[44] De maneira mais específica, Tversky e Kahneman (1982a) apontavam que Linda tinha 31 anos, era solteira, sincera e muito inteligente. Era formada em filosofia e, como estudante, preocupava-se com questões relacionadas à discriminação e à justiça social, além de ter participado de demonstrações antinucleares.

então, pediam que os participantes avaliassem a probabilidade de que o táxi envolvido no acidente tenha sido o Azul e não o Verde. Em geral, verificou-se que as taxas-base estatísticas costumam ser subestimadas, quando está disponível alguma informação específica sobre o caso (KAHNEMAN, 2012, p. 213). Em outros termos, em vez de combinar as duas fontes de informação apresentadas, as pessoas acabavam ignorando a taxa-base e respondendo que a probabilidade de que o táxi envolvido no acidente tenha sido o Azul seria de 80% – confiabilidade do depoimento da testemunha[45].

Outra heurística importante é a disponibilidade. Tversky e Kahneman (1974, p. 1127) esclarecem que há situações em que as pessoas avaliam a probabilidade de um evento pela facilidade com que ocorrências ou exemplos são trazidos à mente. A acessibilidade e a saliência[46] (*salience*) estão intimamente relacionadas com a disponibilidade (THALER; SUNSTEIN, 2009, p. 25). Eventos mais vívidos, recentes e impactantes, por serem mais acessíveis à memória e atraírem maior atenção, têm suas probabilidades superestimadas. O medo exacerbado que muitas pessoas têm de avião, quando acidentes em outros meios de transporte, como carros, são muito mais frequentes, é um exemplo claro dessa heurística. Nesse caso, a probabilidade de acidentes de avião ocorrerem é superestimada, na medida em que tais acidentes, por serem mais impactantes e atraírem maior cobertura midiática, são lembrados com maior facilidade.

A terceira heurística originalmente descrita por Kahneman e Tversky é a heurística de ancoragem e ajuste[47]. Em muitas situações, as pessoas fazem estimativas a partir de um valor inicial que é ajustado para produzir a resposta final (TVERSKY; KAHNEMAN, 1974, p. 1128). O ponto central da heurística é que os ajustes costumam ser insuficientes, de forma que pontos de partida diferentes produzem estimativas diferentes, enviesadas em direção aos valores iniciais. Mesmo âncoras obviamente irrelevantes interferem no processo decisório (THALER; SUNSTEIN, 2009, p. 23).

Em exemplo clássico de ancoragem e ajuste, Tversky e Kahneman (1974, p. 1128) pediram que grupos de participantes estimassem qual era o percentual de nações africanas entre os países membros da Organização das Nações Unidas – ONU. Antes disso, recebiam um valor

[45] Kahneman (2012, p. 212-213) ressalta, no entanto, que, quando a taxa-base é apresentada de maneira causal, as pessoas atribuem peso considerável a ela e fazem seu julgamento à maneira bayesiana. Por exemplo, quando se afirmava que táxis Verdes estavam envolvidos em 85% dos acidentes e ambas as empresas atuavam com o mesmo número de veículos, essa informação era combinada à confiabilidade da testemunha, de modo que as estimativas se aproximavam da estimativa bayesiana (41%).

[46] A saliência (*salience*) refere-se ao fato de que "estímulos coloridos, dinâmicos ou, de alguma maneira, distintivos geram atenção desproporcionalmente e, por consequência, diferentes julgamentos afetivos de maneira desproporcional" (TAYLOR, 1982, p. 192, Tradução livre).

[47] Para uma revisão bibliográfica sobre a ancoragem, Cf.: (FURNHAM; BOO, 2011).

em uma "roda da fortuna" que podia ser 10 ou 65. Em seguida, eram instruídos a dizer se o valor recebido era superior ou inferior ao percentual que deveriam estimar. Por fim, os participantes deveriam indicar qual era sua melhor estimativa sobre o percentual de nações africanas na ONU. A estimativa média dos sujeitos que receberam o valor 10 foi de 25%, ao passo que, para os sujeitos que receberam o valor 65, a estimativa média foi de 45%, ainda que os valores oferecidos não contribuíssem com qualquer informação útil para o julgamento a ser feito.

Embora a ancoragem e ajuste tenha sido incluída entre as três heurísticas gerais descritas originalmente por Tversky e Kahneman (1974), Kahneman e Frederick (2002) ponderam que a ancoragem não se enquadra propriamente na abordagem de substituição de atributos. Há, na verdade, um aumento de plausibilidade de um valor particular do atributo-alvo. Nesse contexto, Kahneman e Frederick afirmam que parece ser mais adequado incluir uma heurística do afeto no lugar da ancoragem na lista de heurísticas com propósitos gerais, baseadas na substituição de atributos. Há evidências robustas na literatura de que cada estímulo a que o sujeito é submetido evoca uma avaliação afetiva, que pode ocorrer inconscientemente[48]. Como a valência afetiva é uma avaliação natural, ela é uma candidata para substituição de atributos em várias situações, quando uma resposta afetivamente carregada é exigida. Kahneman, Ilana Ritov e David Schkade (1999, p. 207), por exemplo, demonstraram que a disposição para pagar por bens ambientais, como contribuir para a manutenção de uma espécie em extinção, reflete as atitudes[49] que as pessoas têm em relação a esses bens. Em outros termos, uma pergunta acerca da disposição para pagar para reparar um dano ambiental acabava sendo respondida a partir de uma indagação quanto às emoções que esse dano despertava nos participantes do experimento.

Expostas essas heurísticas, para encerrar esta subseção, convém expor um aperfeiçoamento conceitual que contribui para a compreensão do processo de substituição de atributos. Kahneman e Frederick (2002) afirmam que a heurística da representatividade, tal como originalmente descrita por Kahneman e Tversky (1974), invocava duas premissas em que a palavra representativo era utilizada de maneiras diversas. Por um lado, um protótipo, isto é, um exemplar representativo, é utilizado para representar categorias em uma tarefa preditiva. Linda, por exemplo, era utilizada para representar a categoria de caixas bancárias. Por outro

[48] Nesse sentido, o clássico trabalho de R. B. Zajonk (1980) apresenta resultados experimentais de que julgamentos afetivos podem ser independentes de operações perceptivas e cognitivas, além de precederem-nas em termos temporais. Afinal, as reações afetivas são frequentemente as primeiras reações dos organismos.

[49] A atitude, na perspectiva adotada por Kahneman, Ritov e Schkade (1999, p. 205), é a tendência psicológica que se expressa ao avaliar uma entidade a partir do grau em que se é favorável ou desfavorável a ela. O elemento central de uma atitude é a valoração (*valuation*), que atribui à entidade um valor afetivo que varia de extremamente positivo a extremamente negativo.

lado, a probabilidade de que o indivíduo pertença a uma determinada categoria é julgada pelo grau em que ele se assemelha - é representativo - ao estereótipo da categoria. Há, portanto, dois tipos de substituição.

Em expansão da teoria proposta originalmente, Kahneman e Frederick (2002) separam esses dois conceitos de representativo e passam a descrever heurísticas-protótipo (*prototype heuristics*), em que um protótipo substitui uma categoria, mas a representatividade não é necessariamente o atributo-heurístico. Nessa abordagem, os atributos-alvo aos quais as heurísticas-protótipo são aplicadas são extensionais, isto é, pertencem a uma propriedade agregada de um grupo ou de uma categoria cuja extensão é especificada. Por exemplo, quando pensamos na probabilidade de que um grupo de advogados inclua um sujeito específico, a probabilidade é o atributo-alvo extensional ao qual a heurística protótipo vai se aplicar (KAHNEMAN; FREDERICK, 2002).

Esse tipo de heurística poderia explicar vieses observados em previsões categóricas, na valoração econômica de bens públicos e em avaliações sumárias.

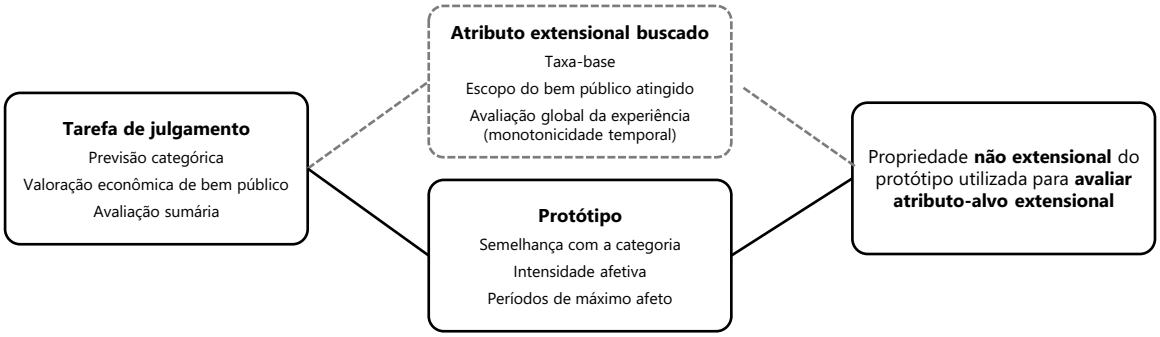
Como descrevem Kahneman e Frederick (2002), em termos normativos, em uma previsão categórica a probabilidade de um sujeito ser membro de determinada categoria deve variar de acordo com o tamanho relativo da população, isto é, com a taxa-base. Na previsão pela representatividade, no entanto, como já mencionado, a representação dos resultados por exemplares prototípicos acaba por descartar as taxas base, já que o protótipo da categoria não contém informações sobre o tamanho do seu grupo de membros. O já mencionado problema dos táxis ilustra essa afirmação.

Em outra perspectiva, Kahneman e Frederick (2002) apontam que a valoração contingente da disposição que um sujeito tem para contribuir com um bem público é insensitiva ao escopo do bem que está sendo avaliado. O exemplo dado é bastante ilustrativo. Em avaliação da disposição que as pessoas teriam para contribuir para salvar pássaros de afogamento em tanques de petróleo, identificou-se que os valores de contribuição entre diferentes grupos eram semelhantes, ainda que o número de pássaros afetados aumentasse consideravelmente de um grupo para o outro (de 2 mil para 20 mil e para 200 mil). Nesse caso, a morte de inúmeros pássaros acaba por ser representada por uma instância prototípica; por exemplo, a imagem de um pássaro coberto por óleo. O protótipo evoca uma resposta afetiva e a intensidade dessa emoção é mapeada em escala de dólar. Assim, o atributo-heurístico da intensidade afetiva acaba por substituir o atributo-alvo do valor econômico, ou seja, o real impacto do dano gerado acaba por não aumentar a disposição que o sujeito teria para contribuir com o bem público.

Finalmente, Kahneman e Frederick (2002) ponderam que, na avaliação sumária de experiências, a monotonicidade temporal é esperada dos indivíduos, em termos normativos. Em outras palavras, há uma intuição de que acrescentar um período extra de dor a um episódio de desconforto, por exemplo, só poderia piorar essa experiência. Barbara L. Fredrickson e Kahneman (1993), no entanto, identificaram que a avaliação retrospectiva das experiências também envolvia dois atos de substituição. Em primeiro lugar, o episódio seria representado por um momento prototípico - os períodos de máximo afeto (dor, por exemplo) registrados durante a experiência e o período final registrado quando a experiência estivesse terminando. Em seguida, o valor afetivo ligado a esses momentos representativos substituiria o atributo-alvo extensional da avaliação global da experiência. Nesse sentido, Kahneman e Fredrickson concluíram que a correlação entre as avaliações retrospectivas e a efetiva duração das experiências seria mínima. Embora o período de duração fosse importante, os momentos mais salientes - e, portanto, prototípicos - acabavam sendo evocados mais rapidamente à mente.

Em todos esses casos, Kahneman e Frederick (2002) afirmam que a negligência da extensão poderia ser explicada pelos processos duais que definem as heurísticas protótipo. Primeiro, uma categoria é representada por um exemplar prototípico; em seguida, uma propriedade não extensional do protótipo é utilizada como atributo-heurístico para avaliar um atributo-alvo extensional da categoria.

Figura 2 - Heurísticas-protótipo



Fonte: Elaboração do Autor.

Em linhas gerais, essas são as principais características do programa de heurísticas e vieses, desenvolvido inicialmente por Kahneman e Tversky. Apesar disso, como Kahneman (2012, p. 18) admite no livro "Rápido e Devagar: Duas Formas de Pensar", a reação ao trabalho que ele e Tversky desenvolveram não foi completamente positiva. De fato, o "foco em vieses

foi criticado no sentido de que sugeriam uma visão injustamente negativa da mente" (KAHNEMAN, 2012, p. 18). Kahneman (2012, p. 557) nota, ainda, que o crítico mais persistente à sua obra tem sido o renomado psicólogo alemão Gigerenzer. A próxima seção apresentará mais detalhes sobre essas críticas.

1.2. Racionalidade ecológica e heurísticas rápidas e frugais: uma alternativa à abordagem das heurísticas e dos vieses

Como mencionado no início deste Capítulo, Gigerenzer direciona suas críticas a aspectos diferentes do programa de heurísticas e vieses. Na linha de Vranas (2000), é possível afirmar que Gigerenzer critica a abordagem de Kahneman e Tversy segundo critérios empíricos, metodológicos e normativos. A ideia desta seção é apresentar minimamente os traços gerais da crítica de Gigerenzer, assim como de sua teoria, sem a pretensão de esgotar todos os detalhes de sua obra. Para facilitar a apresentação desses pontos, esta seção será dividida em diferentes subseções, nas quais cada um desses aspectos será examinado de maneira separada.

1.2.1. Normas rasas e cegas ao contexto: as críticas normativas de Gigerenzer

De acordo com Gigerenzer e Todd (1999), a abordagem de racionalidade limitada de Herbert Simon apresenta dois componentes: um cognitivo e outro ecológico. Em relação ao aspecto cognitivo, a partir da obra de Simon, considera-se que os modelos de julgamento e tomada de decisão devem ser construídos com base no que realmente se conhece a respeito da mente humana. Abandonam-se competências fictícias, já que, na realidade, estratégias de otimização são muitas vezes desconhecidas ou incapazes de ser conhecidas. Assim, para fazer um julgamento diante de incertezas, a busca de informações das pessoas é guiada por métodos mais simples, heurísticas que determinam quando elas devem parar e regras de decisão que tornem úteis as informações encontradas. É o aspecto cognitivo da obra de Simon que influenciou a abordagem de heurísticas e vieses, ainda que de forma indireta, como já mencionado acima[50]. O aspecto ambiental, entretanto, ganhou menos atenção.

Como salientam Gigerenzer e Todd (1999), na perspectiva da racionalidade limitada de Simon, a *estrutura ambiental* é de crucial importância, porque pode explicar quando e por que heurísticas simples funcionam bem, isto é, quando a heurística está adaptada à estrutura de

[50] Cf., nesse sentido, nota 322 *supra*.

informação do ambiente. A racionalidade é normalmente associada à manutenção da ordem interna de crenças e inferências. Não por outra razão, um critério utilizado para avaliar o desempenho de inferências humanas é a coerência (GIGERENZER; TODD, 1999). Busca-se, por meio das leis da lógica e da teoria da probabilidade, demonstrar que as preferências das pessoas são consistentes, transitivas e estáveis. A violação desses critérios é vista como sinal de irracionalidade.

Essa abordagem é adotada por Kahneman e Tversky em seu programa de heurísticas e vieses[51]. Nesse caso, embora o programa de heurísticas e vieses tenha se tornado célebre por descrever as falhas das abordagens que adotam a racionalidade ilimitada como pressuposto, a visão de Kahneman e Tversky não seria assim tão distante da abordagem que por eles é criticada (GIGERENZER; GOLDSTEIN, 1996). Afinal, se discrepâncias em relação a esse núcleo são consideradas vieses ou ilusões cognitivas, o núcleo normativo da racionalidade clássica seria mantido. A mudança se limitaria à descrição do comportamento humano.

De acordo com Gigerenzer e Todd (1999), no entanto, organismos reais teriam que lidar, ao longo da maior parte do tempo, com a desordem externa do seu ambiente, tentando tomar decisões que os tornem capazes de sobreviver e de se reproduzir. Para se adaptar a esses desafios, os indivíduos devem ser capazes de fazer inferências rápidas, frugais e acuradas; devem ser capazes de se beneficiar da forma como a informação está estruturada nos ambientes, ainda que essa informação não possa integrar um modelo estatístico que seja utilizado como norma para comportamentos racionais. Afinal, se explorar as dependências informacionais encontradas no ambiente fosse algo vantajoso evolutivamente para as pessoas, é provável que tal comportamento seria selecionado pela evolução.

Nessa perspectiva, para avaliar o desempenho de heurísticas, na perspectiva da racionalidade ecológica, é necessário comparar os julgamentos humanos à forma como eles se adaptam aos ambientes do mundo real (GIGERENZER; TODD, 1999). A função das heurísticas não é a coerência, mas a realização de inferências razoáveis e adaptativas sobre o mundo real e social, considerando as limitações de tempo e de conhecimento das pessoas; é a correspondência ao ambiente, em vez de a simples coerência.

Gigerenzer (1991) ilustra esse ponto com um exemplo de Richard E. Nisbett e Lee Ross (1980, p. 14-15). Imagine que você quer comprar um carro e tem a informação de que um Volvo, em uma amostra de algumas centenas de carros, tem maior vida útil do que um Saab[52].

---

[51] Cf., nesse sentido, nota 38 *supra*.

[52] Assim como a Volvo, a Saab é uma companhia sueca que, até 2011, fabricava carros. Nesse sentido, Cf.: https://www.saab.com/pt-br/markets/brasil/historias/2018/a-saab-ainda-faz-carros. Acesso em: 27 dez. 2021.

Ontem, o Volvo de seu vizinho quebrou. Que carro você compraria? Nisbett e Ross mencionam que a resposta lógica à informação do vizinho seria simplesmente a aquisição de mais um Volvo. Gigerenzer (1991), então, fornece o exemplo de outra situação, em que a resposta lógica parece ser diferente. Imagine que você vive na selva e deve decidir entre duas alternativas: deixar seu filho nadar em um rio ou deixá-lo se pendurar nas árvores. Ao longo dos últimos cem anos, houve apenas um acidente no rio, em que uma criança foi comida por um crocodilo. Ontem, seu vizinho disse que o seu próprio filho foi comido por um crocodilo. Aonde o seu filho iria? Nesse caso, não parece lógico aplicar o mesmo algoritmo utilizado no caso do Volvo; a mente, provavelmente, inseriria a nova informação no algoritmo e tomaria uma outra decisão, talvez suspeitando que o mundo daquele rio teria mudado.

Com base nesse exemplo, Gigerenzer (1991) argumenta que a racionalidade é ecológica, de forma que o conteúdo dos problemas tem importância vital para o entendimento dos julgamentos. Esse conteúdo incorpora conhecimentos implícitos sobre a estrutura do ambiente, que são utilizados pela mente para tomar decisões. Assim, para identificar as limitações que as heurísticas impõem ao raciocínio, seria necessário estudar a estrutura de cada ambiente em que elas são aplicadas. Em outras palavras, no caso da resolução de problemas estatísticos, a não aplicação de uma determinada fórmula – como o teorema de Bayes – não pode ser vista, por si só, como uma limitação ao raciocínio. Na verdade, seria necessário analisar a estrutura do ambiente e escolher se um algoritmo estatístico deve ser aplicado, além de escolher qual seria o modelo de algoritmo adotado.

Feitas essas considerações, Gigerenzer (1996) entende, então, que as abordagens adotadas como norma por Kahneman e Tversky são excessivamente rasas. Em geral, ao desenvolver um modelo, estatístico, é necessário, antes de mais nada, investigar o conteúdo de um problema e trabalhar em um conjunto de premissas que podem ser adotadas. Estabelecidas essas duas etapas, um modelo estatístico é finalmente construído com base em tais premissas. Para Gigerenzer, Kahneman e Tversky fazem o contrário: (i) partem de um princípio estatístico (como a regra da conjunção ou o teorema de Bayes), que é escolhido como norma; (ii) enquadram o conteúdo de algum problema nessa norma; e (iii) avaliam as respostas dadas ao problema de acordo com os parâmetros estabelecidos pela norma. Por essa razão, o conteúdo do problema deixa de ser analisado na construção do modelo normativo, e as premissas que as pessoas utilizam com base na situação descrita acabam sendo ignoradas.

Na abordagem de Gigerenzer (1996), porém, o contexto é fundamental para dizer se uma resposta é ou não adequada. Quando se pergunta, por exemplo, se é mais provável que alguém morra em uma semana ou em um ano – situação em que, para Kahneman e Tversky

(1996), a regra da conjunção deveria ser aplicada –, não seria possível concluir, imediatamente, que essa é a norma universalmente aplicável como raciocínio razoável para o caso. Se, por exemplo, o indivíduo analisado estiver padecendo de uma doença terminal, esse contexto poderia indicar que a probabilidade de morte na próxima semana é superior. A *falácia da conjunção*, nesse caso, não seria um viés cognitivo, e a decisão ecologicamente racional seria violar a mencionada regra.

Além disso, na visão de Gigerenzer (1991), Kahneman e Tversky presumem que os problemas narrados aos participantes de seus experimentos – como o clássico exemplo da Linda – teriam uma única resposta correta. Respostas distintas seriam, então, vieses comportamentais, isto é, violações da teoria da probabilidade. Gigerenzer afirma, no entanto, que os vieses extensamente descritos por Kahneman e Tversky partiriam de uma premissa equivocada: a adoção da teoria Bayesiana da probabilidade como padrão normativo a ser utilizado como critério para avaliar os julgamentos das pessoas. Essa visão representa apenas uma amostra pequena das teorias que são efetivamente utilizadas por estatísticos e probabilistas. Não é compartilhada, por exemplo, pelos proponentes da teoria frequentista[53], que seria a visão majoritária. Na interpretação frequentista, apenas as frequências relativas de eventos deveriam ser tratadas pela probabilidade. Assim, não faria sequer sentido perguntar a um sujeito qual é a probabilidade de um evento singular acontecer.

Os vieses surgem justamente a partir de interpretações que ofuscam a distinção entre eventos singulares e frequências (GIGERENZER, 1991). Por exemplo, no clássico exemplo de Linda, utilizado por Tversky e Kahneman (1982a) para ilustrar a heurística da representatividade e a falácia da conjunção, pede-se que o sujeito avalie a probabilidade de eventos singulares. Para um frequentista, esse problema não teria relação com a teoria da probabilidade. Logo, as respostas equivocadas dos participantes dos experimentos de Kahneman e Tversky não seriam vieses comportamentais, mas apenas um reflexo de perguntas que ignoram a distinção entre eventos singulares e frequências.

[53] Como explica Vranas (2000, p. 182, Tradução livre): "Um conceito frequentista de probabilidade se preocupa com medidas de probabilidade cujos domínios são conjuntos de coleções de eventos (ou conjuntos de proposições de múltiplos casos) e cujos valores são (limites de) frequências relativas em tais coleções. Uma concepção subjetivista de probabilidade se preocupa com medidas de probabilidade cujos domínios são conjuntos de proposições e cujos valores são graus de confiança em tais proposições".

1.2.2. Vieses cognitivos podem desaparecer: a crítica empírica de Gigerenzer

Do ponto de vista empírico, Gigerenzer (1991) relata que inúmeros experimentos[54] demonstram a possibilidade de eliminação ou de drástica redução dos vieses comportamentais, caso os problemas apresentados aos participantes desses experimentos sejam reestruturados em termos frequentistas.

Um exemplo clássico pode ser identificado com o problema de Linda. Como mencionado na Seção 2 acima, a maior parte dos participantes julgava que a probabilidade de Linda ser caixa bancária seria inferior à probabilidade de ela ser feminista e caixa bancária. Havia, em tal resultado, uma clara violação da regra da conjunção. Por isso, Tversy e Kahneman (1982a) identificaram esse viés como a falácia da conjunção, explicada pela heurística da representatividade. Apesar disso, Gigerenzer (1991, p. 92) demonstrou que o efeito da falácia da conjunção praticamente desapareceria se o problema fosse descrito em termos frequentistas, da seguinte forma: "Há 100 pessoas que se encaixam na descrição acima (i.e. a de Linda). Quantas delas são: (a) caixas bancárias; (b) caixas bancárias e ativas no movimento feminista"[55].

Em outra perspectiva, Gigerenzer (1991) demonstrou que a negligência da taxa-base pode desaparecer se a premissa de que a amostragem é aleatória consegue ser fixada nas pessoas que vão tomar as decisões. Um exemplo dado pelo autor é a utilização de uma urna, da qual os participantes poderiam retirar descrições de indivíduos e, em seguida, apresentar seus julgamentos de probabilidade. Nesse caso, a taxa-base passaria a ser levada em consideração e os julgamentos se aproximariam daqueles previstos pelo teorema de Bayes.

De acordo com Gigerenzer (1991), se as pessoas raciocinassem a partir da heurística da representatividade, os vieses identificados por Kahneman e Tversky não deveriam desaparecer com a simples reestruturação do problema apresentado. Nesse contexto, o psicólogo alemão afirma fazer mais sentido explicar a forma como as pessoas julgam questões de probabilidade a partir de modelos mentais probabilísticos, como será exposto na Subseção 1.2.4 abaixo. Nessa abordagem, o "estatístico intuitivo" da mente seria frequentista.

---

[54] Leda Cosmides e John Tooby (1996), por exemplo, demonstraram que a simples tradução de problemas de diagnósticos médicos para termos frequentistas – em oposição à versão que expunha as informações em termos percentuais – era capaz de estimular o raciocínio Bayesiano correto em 76% dos participantes. Quando, além dessa tradução, se exigia que os indivíduos criassem representações concretas e visuais em termos frequentistas de tais problemas, 92% dos participantes conseguiam empregar o raciocínio Bayesiano correto.

[55] Tradução livre. Texto original: "*There are 100 persons who fit the description above (i.e. Linda's). How many of them are: (a) bank tellers (b) bank tellers and active in the feminist movement".*

1.2.3. Heurísticas vagas e indefinidas: a crítica metodológica de Gigerenzer

Gigerenzer (1991) reconhece que Tversky e Kahneman (1974) desempenharam um papel importante ao introduzir as três heurísticas clássicas - representatividade, disponibilidade e ancoragem -, em seu seminal estudo de 1974, publicado na *Science*. Tais heurísticas foram um primeiro passo promissor para conectar a pesquisa Bayesiana com as teorias cognitivas. Para Gigerenzer, no entanto, a teoria das heurísticas e dos vieses de Tversky e Kahneman seria apenas um estágio de transição importante, que deveria ser transformado para que houvesse verdadeiro progresso na área. Em sua visão, em mais de 30 anos de pesquisas, o programa de heurísticas e vieses pouco teria avançado.

Nesse contexto, Gigerenzer (1996) aponta que as heurísticas definidas por Kahneman e Tversky – representatividade, disponibilidade e ancoragem – são noções vagas, indefinidas e pouco específicas. O psicólogo alemão afirma que Kahneman e Tversky não explicam as condições antecedentes que originam essas heurísticas ou as condições capazes de suprimi-las. Os processos cognitivos subjacentes às heurísticas tampouco são bem explicados. Em vez disso, Kahneman e Tversky utilizariam rótulos de uma palavra – como representatividade – e explicariam os fenômenos com base na redescrição desses próprios fenômenos. Por exemplo, em resposta às críticas empíricas elaboradas por Gigerenzer, Kahneman e Tversky (1996) firmaram ser da natureza das ilusões cognitivas o fato de existirem circunstâncias em que a resposta correta se torna transparente[56]. Nesse caso, expor os problemas em termos de frequências naturais tornaria a resposta correta mais saliente, de forma que as pessoas acabariam por escolhê-la. Para Gigerenzer (1996), porém, dizer que alguns sinais são mais *salientes* que outros não contribui para a adequada compreensão dos processos que levam ao desaparecimento de tais heurísticas. Em sua visão, trata-se apenas de uma redescrição do fenômeno, com outro conceito indeterminado.

Por essa razão, Gigerenzer (1996) afirma que, ao mesmo tempo em que as heurísticas de Kahneman e Tversky explicam muito pouco, elas acabam por explicar demais. De fato, se, por um lado, as circunstâncias em que as heurísticas funcionam e os processos cognitivos subjacentes a ela são desconhecidos, por outro lado, qualquer das heurísticas serve para explicar praticamente toda sorte de resultado experimental. Por exemplo, a negligência da taxa-base é

[56] Kahneman e Tversky (1996) chegaram a elaborar resposta às principais críticas feitas por Gigerenzer. O impasse, no entanto, continua, já que Gigerenzer (1996) ofereceu sua réplica a Kahneman e Tversky e demonstrou que os argumentos apresentados pelos autores israelenses não modificaram a sua forma de visualizar o programa das heurísticas e dos vieses.

compreendida a partir da heurística da representatividade, como demonstrado acima; ao mesmo tempo, seria possível explicar o resultado oposto – isto é, um excesso de peso à taxa-base – a partir da heurística da ancoragem (GIGERENZER, 1996, p. 592).

Feitas essas críticas, Gigerenzer (1996) argumenta que a maneira correta de explicar os julgamentos humanos não seria por meio das heurísticas descritas por Kahneman e Tversky e das ilusões cognitivas que elas originam, mas pela teoria dos modelos mentais probabilísticos. Essa teoria especifica as condições sob as quais julgamentos de frequência são válidos, além de especificar as razões que levam a esses julgamentos, como será demonstrado a seguir.

1.2.4. Heurísticas rápidas e frugais e modelos mentais probabilísticos: um novo programa de pesquisas

Ao descrever as teorias que adotam a racionalidade limitada como pressuposto, Gigerenzer e Todd (1999) descrevem, em primeiro lugar, a abordagem denominada *satisficiente* (*satisfacing*) - neologismo cunhado por Herbert Simon (1956) para ilustrar processos cognitivos que são, ao mesmo tempo, satisfatórios (*satisfatory*) e que são suficientes (*suffice*)[57]. Essa abordagem era adotada por Simon e descreve um método para fazer uma escolha em condições de incerteza, a partir de um rol de alternativas encontrado sequencialmente, quando não se sabe muito sobre as possibilidades futuras. Nessa perspectiva, as pessoas definiriam um nível de aspiração ajustável e encerrariam a busca por outras alternativas assim que esse nível fosse superado (GIGERENZER; TODD, 1999).

Por exemplo, se um sujeito pretende comprar um livro na *Amazon* e precisa estimar a probabilidade de que a opção escolhida lhe traga bem-estar, com processos *satisficientes* (*satisficing*), ele não tentará computar o bem-estar que cada uma das opções disponíveis lhe traria. Na verdade, o sujeito definiria um nível de aspiração, isto é, um limiar a partir do qual esse nível seria atingido. Ao avaliar as alternativas, uma obra de Paulo Coelho poderia estar abaixo do seu nível de aspiração e, por consequência, seria rejeitada; ao se deparar com um

[57] Nas palavras de Simon (1956, p. 1), "*A comparative examination of the models of adaptive behavior employed in psychology (e.g., learning theories), and of the models of rational behavior employed in economics, shows that in almost all respects the latter postulate a much greater complexity in the choice mechanisms, and a much larger capacity in the organism for obtaining information and performing computations, than do the former. […] Both from these scanty data and from an examination of the postulates of the economic models it appears probable that, however adaptative the behavior of organisms in learning and choice situations, this adaptiveness falls far short of the ideal of 'maximizing' postulated in economic theory. Evidently, organisms adapt well enough to 'satisfice'; they do not, in general, 'optimize'*".

clássico de Gabriel García Márquez, por outro lado, o nível de aspiração é atingido e a busca se encerra.

Para Gigerenzer e Todd (1999), no entanto, essa abordagem pode continuar a exigir um volume excessivo e irrealista de deliberação. Como seria possível determinar um nível de aspiração apropriado ou calcular como uma das opções se compara a esse nível? Por essa razão, Gigerenzer e Todd rejeitam a abordagem de Simon e propõem uma nova visão de racionalidade limitada: as heurísticas rápidas e frugais (*fast and frugal heuristics*). Essas heurísticas empregam uma quantidade mínima de tempo, conhecimento e capacidade computacional para fazer escolhas adaptativas em ambientes reais. A busca por informações ou objetos é limitada por heurísticas rápidas e frugais, com regras de parada facilmente computáveis. As escolhas são feitas com regras de decisão que também são simples.

Um exemplo simples é a heurística do reconhecimento[58], que explora a capacidade do reconhecimento de ser eficiente para fazer inferências sobre aspectos desconhecidos do mundo (GOLDSTEIN; GIGERENZER, 1999). Nessa perspectiva, uma comida cujo gosto é reconhecido poderá ser considerada mais segura do que outra que apresenta gosto desconhecido; uma universidade reconhecida será vista como melhor do que a desconhecida, e assim sucessivamente (GIGERENZER; TODD, 1999).

A tarefa normalmente utilizada para demonstrar a heurística do reconhecimento é a pergunta a respeito de qual dentre duas cidades tem a maior população. Para essas tarefas, a heurística do reconhecimento prescreve que: "Se um de dois objetos for reconhecido e o outro não for, então infira que o objeto reconhecido tem o maior valor" (GOLDSTEIN; GIGERENZER, 1999, p. 41)[59]. Além de demonstrar que a heurística do reconhecimento tem um desempenho satisfatório nos domínios em que é empregada, Goldstein e Gigerenzer (1999) demonstraram que, em algumas circunstâncias, essa heurística leva ao efeito menos-é-mais (*less-is-more effect*), em que a falta de conhecimento pode levar a inferências melhores, isto é, quando recebiam mais informações sobre o problema analisado, os indivíduos faziam inferências menos precisas.

A heurística do reconhecimento – assim como as demais heurísticas rápidas e frugais – integra os modelos mentais probabilísticos propostos por Gigerenzer (1991, p. 104 - 105). Como já mencionado, Gigerenzer argumenta que, como os vieses comportamentais

[58] Kahneman (2012, p. 566) chega a afirmar, inclusive, que a heurística do reconhecimento é a descoberta mais memorável do programa de pesquisas inaugurado por Gigerenzer e seus colaboradores.
[59] Tradução livre. Texto original: "*If one of two objects is recognized and the other is not, then infer that the recognized object has the higher value*".

desaparecem em julgamentos de frequência, o "estatístico intuitivo" da mente é frequentista. O processamento de frequências seria automático e ocorreria por meio dos modelos mentais probabilísticos. Para ilustrar essa teoria, Gigerenzer apresenta o seguinte problema: "Que cidade tem mais habitantes? (a) Heidelberg, ou (b) Bonn?", na linha das tarefas utilizadas para demonstrar a heurística do reconhecimento. De acordo com a sua teoria, a resposta a esse problema é construída, em primeiro lugar, por meio de um recorte inferencial, que é denominado modelo mental probabilístico. Esse modelo é uma generalização de um problema particular e consiste em uma classe de referência de objetos e uma rede de sinais de probabilidade (*probability cues*).

No caso do problema proposto acima, a classe de referência poderia ser algo como "todas as cidades da Alemanha", enquanto os sinais de probabilidade seriam, por exemplo, ter ou não um time de futebol na Bundesliga - 1ª divisão do campeonato nacional de futebol da Alemanha – ou ser ou não a capital do seu estado. A lista de sinais é, então, avaliada para aferir se um dos sinais pode ser ativado - por exemplo, se há um time da Bundesliga naquela cidade em especial. No caso apresentado por Gigerenzer (1991), como Bonn é uma capital, esse sinal poderia ser ativado, de modo a interromper a busca por novos sinais, com a subsequente conclusão do julgamento. Em outras palavras, no caso concreto, Bonn seria escolhida como a cidade com mais habitantes. A confiança do sujeito nesse julgamento particular dependeria da validade do sinal que foi ativado, isto é, da medida em que aquele sinal de probabilidade é, de fato, representativo da resposta que é buscada pelo sujeito[60].

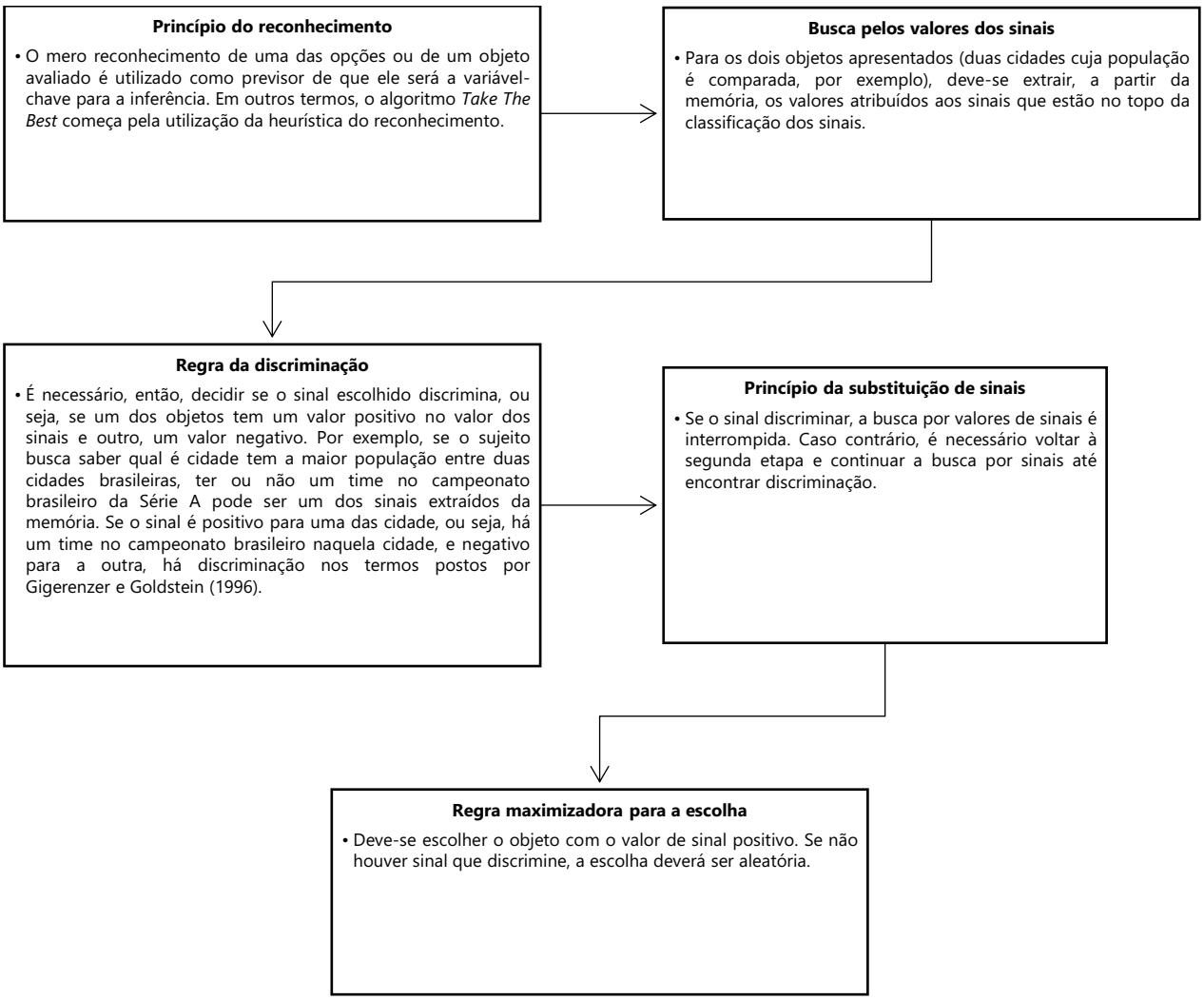
Além disso, de acordo com Gerd Gigerenzer e Daniel Goldstein (1996), a teoria dos modelos mentais probabilísticos se relaciona a três visões distintas. Em primeiro lugar, as inferências indutivas precisam ser estudadas com respeito ao ambiente natural e não em abstrato, com base na comparação a modelos normativos da estatística. O segundo ponto é que as inferências indutivas são levadas adiante por algoritmos satisficientes (*satisficing*), na visão adotada por Simon. Finalmente, Gigerenzer e Goldstein ressaltam que as inferências indutivas são baseadas em frequências de eventos em uma mesma classe de referência. Toda essa teoria substitui o cânone da racionalidade clássica por mecanismos psicológicos de inferências simples e plausíveis, que podem ter surgido por meio da evolução.

---

[60] Para Gigerenzer (1991), quando um problema é apresentado em termos de frequência, em vez de eventos singulares, os indivíduos também constroem um modelo mental para chegar a estimativas razoáveis. Nesse caso, porém, a classe de referências e a lista de pistas de probabilidade são distintas: como a variável que se busca, desta vez, é o número de respostas corretas em vez do número de habitantes, o sujeito deixa de utilizar pistas como times na Bundesliga; em vez disso, as taxas-base de eventos anteriores servem como pistas, enquanto situações de teste similares podem servir como classes de referência. Logo, os modelos mentais probabilísticos para julgamentos de eventos singulares são diferentes daqueles utilizados para julgamentos de frequência.

O algoritmo base da abordagem dos modelos mentais probabilísticos é o *Take The Best*. Como descrevem Gigerenzer e Goldstein (1996, p. 653) a política desse algoritmo é: "pegue a melhor opção e ignore o resto"[61]. O algoritmo presume uma classificação subjetiva de sinais, e o sinal mais bem colocado em tal classificação é considerado o melhor e funciona como uma série encadeada de passos:

Figura 3 - Algoritmo *Take the Best*



Fonte: Elaboração do Autor com base em Gigerenzer e Goldstein (1996)

Esse algoritmo apresenta algumas características marcantes, também descritas por Gigerenzer e Goldstein (1996). Em primeiro lugar, diferentemente de algoritmos utilizados por um agente racional no sentido clássico, o "*Take The Best*" estende a sua busca apenas por uma

[61] Tradução livre. Texto original: "[…] *take the best, ignore the rest.*"

porção do conhecimento total na memória e a interrompe imediatamente quando o sinal discriminador é encontrado. Em outra perspectiva, em vez de analisar toda a informação disponível e combiná-la em um único valor para obter a resposta do problema, o algoritmo "*Take The Best*" utiliza um sinal e substitui a resposta da pergunta buscada pela resposta indicada pelo sinal. Por fim, a quantidade de informação processada é contingente em cada tarefa - par de objetos avaliados, como duas cidades - e varia de maneira previsível entre indivíduos com diferentes graus de conhecimento.

É necessário ressaltar, ainda, que Gigerenzer e Goldstein (1996) demonstraram que o algoritmo "*Take The Best*" consegue fazer tantas inferências corretas sobre aspectos desconhecidos do ambiente no mundo real quanto qualquer algoritmo baseado na integração de informações - e mais até do que alguns desses algoritmos. Em outros termos, as normas clássicas da racionalidade podem ser suficientes, mas não são necessárias para boas inferências em ambientes reais. Ou, como questionam os autores, será que o raciocínio não poderia ser tanto racional quanto psicológico?

Estes são, em linhas gerais, os traços que marcam a abordagem de Gigerenzer e seus seguidores para explicar a tomada de decisões em condições de incerteza. Expostos a sua crítica e o seu modelo de comportamento, para encerrar este Capítulo, a próxima seção apresentará a posição adotada nesta dissertação em meio à controvérsia exposta.

1.3. Heurísticas são úteis em alguns ambientes, mas podem levar a erros que causam prejuízo em termos de bem-estar

Evidentemente, como um trabalho desenvolvido no Programa de Pós-Graduação em Direito, não é pretensão desta dissertação resolver o Grande Debate da Racionalidade. Apesar disso, posicionar-se em meio a tal controvérsia é fundamental, na medida em que o contraste entre essas abordagens distintas pode se refletir nos níveis de intervenção estatal que uma ou outra perspectiva teórica admitem como legítimos.

A distinção mais marcante entre o programa de heurísticas e vieses e a abordagem das heurísticas rápidas e frugais é a maneira pela qual essas teorias avaliam a qualidade dos julgamentos humanos. De acordo com o jurista Mark Kelman (2011, p. 235), o Programa de Heurísticas e Vieses considera que os julgamentos humanos, ainda que motivados por heurísticas, são bons para situações típicas, mas falíveis e sujeitos à manipulação por *outsiders* mal-intencionados. Os proponentes da abordagem das Heurísticas Rápidas e Frugais, em contrapartida, consideram os julgamentos humanos excelentes, ainda que elas não tenham a

capacidade sobrenatural que seria esperada de demônios *laplaceanos*, para retomar a expressão utilizada no início deste Capítulo.

Segundo Kelman (2011, p. 238), a incorporação dos *insights* do programa de Kahneman e Tversky, com uma visão negativa das capacidades decisórias das pessoas, torna mais provável a identificação de falhas de mercado oriundas de generalizações pouco precisas. A partir de tal identificação, intervenções mais paternalistas poderiam ser vistas como necessárias. Haveria, nessa abordagem, um certo ceticismo quanto à possibilidade de mais informações aprimorarem por si só a capacidade decisória de indivíduos e da coletividade. Na perspectiva das Heurísticas Rápidas e Frugais, por outro lado, como heurísticas são vistas como métodos excelentes de se fazer inferências, intervenções não seriam necessárias e técnicas baseadas na educação e no aprimoramento dos agentes decisores – os *boosts* que serão descritos no Capítulo 4 – deveriam ser promovidas e seriam suficientes.

Diferentemente do que Kelman faz parecer, no entanto, a visão de Kahneman e Tversky, não é necessariamente negativa; na verdade, essa abordagem é simplesmente mais *realista* e reflete o fato de que as pessoas frequentemente tomam decisões equivocadas, por agirem intuitivamente, com base em heurísticas *inadequadas para o seu ambiente*. A partir de tal premissa, esta dissertação se alinha ao posicionamento de Kahneman e Tversky e aos inúmeros experimentos elaborados a partir de seu programa de pesquisas. Apesar disso, pensar a racionalidade em termos *ecológicos* parece ser uma perspectiva promissora e alinhada à realidade; não há, porém, incompatibilidade entre essa visão e o alinhamento ao posicionamento de Kahneman e Tversky. O objetivo desta seção é apresentar as justificativas teóricas que embasam esta opção teórica.

1.3.1. A *calibração* como norma no julgamento de probabilidades: maximizar a precisão das crenças e evitar perdas de bem-estar

Em termos normativos, Gigerenzer sustenta a posição de que o Programa de Heurísticas e Vieses se baseia em normas rasas e cegas ao contexto com base em três argumentos principais: a) a probabilidade *bayesiana* não deve ser utilizada como norma de referência para avaliar a qualidade de julgamentos; b) em uma perspectiva frequentista, eventos singulares não deveriam ser tratados em termos de probabilidade, de modo que os vieses cognitivos não seriam erros em tal abordagem e c) Kahneman e Tversky não se preocupam com a estrutura do ambiente em que as heurísticas são aplicadas, isto é, com a racionalidade ecológica. Os dois primeiros pontos

serão enfrentados nesta subseção, enquanto a racionalidade ecológica será discutida na próxima subseção.

Para começar a enfrentar a crítica posta por Gigerenzer, convém destacar que, usualmente, em economia, uma decisão *racional* exige que os decisores sigam os axiomas da abordagem *bayesiana* (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 8). Kahneman e Tversky, no entanto, não *defendem* essa norma nas pesquisas que formaram o Programa de Heurísticas e Vieses. Por que, então, a racionalidade *bayesiana* é posta como referência nessas pesquisas?

A resposta é simples: com elas, Kahneman e Tversky almejavam testar se as pessoas faziam julgamentos em conformidade com os axiomas *bayesianos* (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 24). Kahneman e Tversky (1973, p. 237) chegaram a afirmar, por exemplo, que explorariam as regras que determinam as previsões intuitivas e julgamentos de confiança, em contraste com os princípios normativos da previsão estatística, que incluem alguns dos axiomas originados da abordagem *bayesiana*. Para fazer isso, evidentemente, o método cientificamente adequado seria testar diretamente as premissas individuais da racionalidade *bayesiana* ou testar um modelo que utilizasse essas premissas junto de outras premissas auxiliares (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 24). Afinal, Kahneman e Tversky não poderiam testar a racionalidade bayesiana, empregando princípios da abordagem frequentista como referência para esse teste.[62]

Além de cumprir o objetivo almejado – e demonstrar que o comportamento humano não era consistente com a racionalidade *bayesiana* –, Kahneman e Tversky foram além e identificaram várias classes de heurísticas que explicam o comportamento humano em diferentes contextos e enquadramentos (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 11). Essa compreensão mais ampla é o cerne do Programa de Heurísticas e Vieses.

Kahneman e Tversky, no entanto, não se preocupavam apenas com avaliações de probabilidade subjetiva, para as quais os princípios *bayesianos* seriam aplicados (KAHNEMAN; TVERSKY, 1996, p. 583). De fato, apenas dois dos vieses descritos originalmente por Tversky e Kahneman (1974) envolviam probabilidades subjetivas: a insensibilidade à probabilidade anterior dos resultados e a confiança excessiva em distribuições

[62] Sanjit Dhami, Ali al-Nowaihi e Cass R. Sunstein (2019, p. 40) afirmam, ainda, que os domínios de escolha com os quais os programas de Kahneman e Tversky e de Gigerenzer se preocupam são distintos. Enquanto o programa de heurísticas e vieses costuma avaliar situações de certeza, risco e incerteza subjetiva, para as quais a racionalidade *bayesiana* costuma ser aplicada, Gigerenzer geralmente se interessa por situações mais amplas, em que há verdadeira incerteza. Nesse caso, ninguém seria capaz de imaginar ou prever possíveis resultados – ou probabilidades subjetivas ou objetivas. As contribuições dos dois programas estariam, então, em domínios não sobrepostos, e o Grande Debate da Racionalidade, em vez de adversarial, seria equivocado e enganoso.

subjetivas de probabilidade. Os outros vieses descritos em tal trabalho – incluindo o efeito de âncoras, vieses de disponibilidade no julgamento de frequências, correlações ilusórias, previsões não regressivas e concepções equivocadas sobre aleatoriedade – não envolvem probabilidades subjetivas e não são afetados pela crítica de Gigerenzer.

Nesse contexto, Kahneman e Tversky (1996) afirmam que sua abordagem não é rasa e exclusivamente *bayesiana*, como Gigerenzer faz parecer. O programa de heurísticas e vieses, na verdade, adota padrões normativos ecléticos e, muitas vezes, objetivos. De fato, Kahneman e Tversky (1996, p. 583) demonstraram vieses variados, como: a) julgamentos de frequência e estimativas numéricas que desviam de valores medidos objetivamente; b) resultados de amostragem que se afastam de valores obtidos de acordo com princípios básicos da análise combinatória e da teoria da amostragem e c) previsões numéricas intuitivas que violam o princípio da regressão. Em outras palavras, não se trata de uma perspectiva *bayesiana* rasa.

Não por outra razão, Kahneman (2000, p. 682) afirmou expressamente que a intenção de sua obra não era demonstrar a irracionalidade humana, mas entender a psicologia dos julgamentos e escolhas intuitivos. Nessa perspectiva, desvios aos modelos tradicionais de racionalidade - como o raciocínio *bayesiano* - são estudados com o objetivo de diagnosticar processos psicológicos específicos. Em outras palavras, Kahneman e Tversky não defenderam a racionalidade *bayesiana* como norma, mas desenvolveram um modelo para compreender o raciocínio intuitivo a partir de escolhas que desviavam das previsões *bayesianas*.

Em outra perspectiva, a suposta inadequação da avaliação de probabilidades de eventos singulares, defendida por Gigerenzer, também não parece inviabilizar o alinhamento à visão de Kahneman e Tversky. Gigerenzer (1991, p. 104) sugere que "a mente distingue entre frequências e outros significados de probabilidade, assim como um estatístico da escola frequentista o faz".[63] Nesse caso, se as ilusões cognitivas tendem a desaparecer em julgamentos de frequência, Gigerenzer afirma ser tentador pensar que a estatística mental é frequentista.[64] Deixando de lado a questão empírica, que ainda será abordada nesta seção, em termos normativos, a abordagem de Gigerenzer não parece estar correta.

---

[63] Tradução livre. Texto original: "[...] *the mind distinguishes between frequencies and other meanings of probability, just as a statistician of the frequentist school does*."

[64] No mesmo sentido, Cosmides e Tooby (1996, p. 17, tradução livre) sustentam o seguinte: "[...] durante a evolução, humanos regularmente precisaram tomar decisões cujo sucesso poderia ser aprimorado se a natureza probabilística do mundo fosse levada em conta. Eles tiveram acesso a grandes quantidades de informação probabilística, mas principalmente ou talvez apenas em forma de frequências encontradas. Essa informação constituía uma rica fonte disponível a ser utilizada para melhorar a tomada de decisão, considerando procedimentos que poderiam tirar proveito disso. Consequentemente, eles desenvolveram mecanismos que tomavam frequências como inputs, mantinham essa informação como representações frequentistas, e utilizavam essas representações frequentistas como um banco de dados para o raciocínio indutivo efetivo".

Inicialmente, é essencial distinguir a definição matemática da medida de probabilidade, isto é, o cálculo da probabilidade, dos vários conceitos de probabilidade, ou seja, das interpretações daquele cálculo (VRANAS, 2000, p. 182), como a abordagem frequentista ou a subjetivista. Matematicamente, uma medida de probabilidade é simplesmente uma função de valor real, não negativa e aditiva, que tem como domínio um campo de subconjuntos dentro de um espaço amostral e atribui o valor 1 para esse espaço amostral como um todo (VRANAS, 2000, p. 182). Em termos conceituais, por mais que alguém defenda não existir probabilidade de eventos singulares, parece ser indiscutível que as pessoas têm graus de confiança distintos que um evento singular vá ou não acontecer.

Se um paciente será submetido a um procedimento cirúrgico, por exemplo, o médico pode lhe dar um prognóstico de recuperação, que nada mais é do que o grau de confiança que o médico atribui ao evento singular da recuperação de tal paciente. Mesmo um frequentista admitiria que pensar nesse grau de confiança é possível, ainda que não lhe atribua o caráter de probabilidade. Em termos de medida de probabilidade, no entanto, não parece haver nada que impeça a definição de uma função para atribuir um número entre 0 e 1 que represente o grau de confiança com que um indivíduo acredita que proposições diversas – incluindo eventos singulares - podem ser verdadeiras (VRANAS, 2000, p. 182). Mesmo que frequentistas, como Gigerenzer, afirmem que isso não seria probabilidade, a disputa aqui parece ser meramente terminológica; graus de confiança podem ser tratados como uma função de probabilidade, ainda que assim não sejam tratados em termos conceituais (VRANAS, 2000, p. 183).

Em outras palavras, ainda que esses julgamentos não sejam probabilísticos, as pessoas julgam o grau de confiança de eventos singulares a todo momento. Em tais julgamentos, atribuir 99% de chance para um evento e 99% para o seu complemento é um julgamento evidentemente equivocado (KAHNEMAN; TVERSKY, 1996). Apesar disso, na visão de Gigerenzer - denominada por Kahneman e Tversky como normativamente agnóstica[65] -, não haveria erro algum. Ainda assim, as pessoas tratam essa violação como uma falácia no dia a dia. Mais do que isso, violações como essa podem gerar efetivo prejuízo em termos de bem-estar. O sujeito que avalia a probabilidade de sofrer um acidente, por exemplo, pensa em termos de um evento singular; não busca avaliar a frequência com a qual acidentes acontecem dentro de um conjunto de eventos recorrentes. Nesse caso, pode acabar sendo excessivamente otimista, o que pode

[65] Segundo Kahneman e Tversky (1996, p. 585-586), Gigerenzer não considerava insignificantes as afirmações de probabilidade subjetiva, já que chegou até mesmo a coletar esses julgamentos de participantes de experimentos. Ao mesmo tempo, porém, ele invoca que probabilidades subjetivas não têm sentido, para argumentar que esses julgamentos não estão sujeitos a quaisquer padrões normativos. Por isso, a sua visão seria normativamente agnóstica.

levar a menor precaução e, por consequência, a mais acidentes; ou seja, ainda que não exista um erro em termos normativos, não seguir as regras da probabilidade para avaliar esse evento singular pode ser prejudicial ao próprio indivíduo.

Por mais que se tente desenvolver estratégias mais precisas de aprendizado, para que as pessoas aprendam a raciocinar em termos frequentistas, como Gigerenzer propõe, é inevitável que a probabilidade de eventos singulares tenha que ser avaliada em inúmeras situações. Afinal, ainda que as pessoas tenham evoluído para entender frequências naturais de forma mais precisa do que probabilidades, a maioria dos dados econômicos do mundo real são apresentados em termos percentuais (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 23). Nesses casos, a utilização de lógica e de probabilidade é essencial para que boas inferências sejam feitas. O sujeito que escolhe uma apólice de seguro para o seu veículo, por exemplo, tomará uma melhor decisão se conseguir ter uma percepção adequada dos riscos de se envolver em um acidente e dos custos correspondentes ao longo do tempo. Decidir com base em heurísticas, como a disponibilidade, em tais hipóteses, pode gerar prejuízos ao decisor em longo prazo, ainda que se reconheça que a abordagem *bayesiana* não é descritivamente adequada para representar o comportamento humano.

A partir do raciocínio exposto ao longo dos últimos parágrafos, é possível se chegar à conclusão de que existe uma norma adequada para o julgamento de eventos singulares. Essa norma, como defendido por Vranas (2000, p. 185), é a calibração - isto é, o grau de confiança que o sujeito tem na correção de sua resposta. Imagine, por exemplo, que um paciente é informado por seu cirurgião que ele tem 99% de confiança em sua completa recuperação; esse paciente não ficaria satisfeito em descobrir, em seguida, que o cirurgião apenas está correto em 75% do tempo (KAHNEMAN; TVERSKY, 1996, p. 588). Como Kahneman e Tversky ponderam, tanto para o cirurgião quanto para o paciente, essa falha de calibração é indesejada; ou, nos termos expostos por Vranas (2000, p. 185): "Gigerenzer seria indiferente entre obter um diagnóstico de quase certa recuperação de um cirurgião calibrado e de um cirurgião excessivamente confiante?"[66].

Normas de probabilidade, assim, deveriam ser utilizadas simplesmente para cumprir o objetivo de maximizar a precisão das crenças de um indivíduo. Nesse caso, é necessário se atentar para que, ao aplicar normas estatísticas a problemas do mundo real, o conteúdo dos problemas seja levado em conta (VRANAS, 2000, p. 185). Não há, assim, uma adoção indiscriminada da abordagem *bayesiana*; pode ser que, à luz do caso concreto, heurísticas

[66] Tradução livre. Texto original: "*Would Gigerenzer be indifferent between getting a diagnosis of almost certain recovery from a calibrated and from an overconfident surgeon?*"

rápidas e frugais levem a uma solução mais precisa. De todo modo, em inúmeras situações, a abordagem *bayesiana* pode ser recomendada, de modo que os vieses identificados por Kahneman e Tversky podem gerar efetivo prejuízo, ao levar a decisões menos calibradas.

1.3.2. A racionalidade ecológica moderada enquanto medida de racionalidade procedimental compatível com o programa de heurísticas e vieses

Como visto, Gigerenzer defende que a racionalidade seja avaliada em conformidade com o ambiente, de modo que estratégias heurísticas rápidas e frugais deveriam ser consideradas racionais, em termos ecológicos. Uma de suas críticas ao Programa de Heurísticas e Vieses envolve a suposta negligência do componente ecológico, por parte de Kahneman e Tversky. Em sua crítica, no entanto, a forma pela qual a racionalidade ecológica deveria ser levada em consideração não é muito clara.

No paradigma da economia comportamental, a relação entre o comportamento e o *ambiente* costuma ser enfrentada a partir da compreensão de que "*decisões humanas (incluindo heurísticas) são dependentes do contexto, da cultura, da história, do tempo e do enquadramento*" (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 16)[67]. Esse é, inclusive, o marco teórico desta dissertação, como exposto na Introdução. De acordo com Dhami, Al-Nowaihi e Sunstein (2019, p. 17), se a racionalidade ecológica significa que preferências, decisões e crenças dependem do contexto e do enquadramento, o Programa de Heurísticas e Vieses satisfaz plenamente esse tipo de racionalidade.

De fato, o reconhecimento de que diferentes enquadramentos (*framings*) para o mesmo problema de decisão ou julgamento podem levar a diferentes processos mentais é uma marca da abordagem de Kahneman e Tversky (1996, p. 583). Kahneman e Tversky (1982, p. 130), inclusive, já reconheceram que, aparentemente, o processo de raciocínio é vinculado a esquemas ou a conteúdos específicos, de modo que diferentes operações ou regras de inferência estão disponíveis em diferentes contextos. Nesse caso, "o raciocínio humano não pode ser adequadamente descrito em termos de regras formais independentes do conteúdo" (KAHNEMAN; TVERSKY, 1982, p. 130)[68].

---

[67] Tradução livre. Texto original: "[...] *human decisions (including heuristics) are context, culture, history, time and frame dependent* […]*".*

[68] Tradução livre. Texto original: "[...] *human reasoning cannot be adequately described in terms of content-independent formal rules*".

Como mencionado na última subseção, Kahneman (2000, p. 682) admite expressamente que o foco principal de sua obra não recai sobre a racionalidade. Apesar disso, Kahneman (2000, p. 682) reconhece que "[n]a teoria da decisão, a racionalidade é definida pela coerência de crenças e preferências".[69] Ainda que não seja psicologicamente plausível, a premissa de que há coerência entre crenças e preferências é essencial para o modelo teórico dominante nas ciências sociais. A coerência exige que as escolhas e crenças sejam imunes a variações no contexto e no enquadramento (*framing*). Por essa razão, julgamentos intuitivos de probabilidade ou frequência que quebram essa coerência foram vistos como vieses cognitivos. Apesar disso, o foco desse programa de pesquisas, como já exposto, é simplesmente entender a psicologia de julgamentos e escolhas intuitivas.

Para Gigerenzer e seus seguidores, todavia, como também já mencionado acima, a coerência não é considerada a medida adequada para a racionalidade; na verdade, defende-se que a racionalidade deve ser avaliada a partir da correspondência dos processos heurísticos ao ambiente em que os julgamentos são feitos. O exemplo de Gigerenzer e Todd (1999) é ilustrativo: para uma presa que foge de um predador, pode ser vantajoso exibir um comportamento inconsistente para escapar da captura. Tal comportamento seria ecologicamente racional, ainda que não fosse coerente.

Nem sempre, no entanto, as pessoas são capazes de avaliar corretamente os elementos do ambiente que as circunda. De fato, a racionalidade ecológica exige que exista correspondência entre os procedimentos de tomada de decisão de um sujeito, a sua estrutura psicológica e o seu ambiente de escolha (SCHMIDT, 2019, p. 528). Sem essa correspondência, decisões heurísticas podem ser irracionais em termos ecológicos, isto é, podem ser inadequadas se considerado o ambiente em que elas são feitas. Como mencionado na última subseção, em inúmeras situações, como a escolha de uma apólice de seguro para um veículo, a decisão que mais bem se amolda àquele ambiente específico exige uma adequada percepção dos riscos envolvidos no julgamento e a utilização dos axiomas da probabilidade. Decidir de outro modo, em tal contexto, pode ser ecologicamente irracional.

O ponto a que se quer chegar, aqui, é simples: nem sempre heurísticas rápidas e frugais são decisões ecologicamente racionais. Para esclarecer esse ponto em maiores detalhes, é necessário, a princípio, distinguir a racionalidade de conteúdo (*content rationality*) da racionalidade procedimental (*procedural rationality*). Enquanto a primeira diz respeito a escolher a opção racional em razão do seu conteúdo - por exemplo, a partir da escolha que

[69] Tradução livre. Texto original: "*In decision theory, rationality is defined by the coherence of beliefs and preferences* […]".

maximiza a sua utilidade -, a racionalidade procedimental se preocupa com a *forma* pela qual a escolha é feita (SCHMIDT, 2019, p. 517).

Como Andreas T. Schmidt (2019) exemplifica, imagine que você tenha que escolher entre dois copos, sendo que um deles contém água e o outro contém petróleo. Evidentemente, a opção racional em termos de conteúdo, considerando as preferências da maior parte dos indivíduos, é escolher o copo de água. Em termos procedimentais, no entanto, pode ser que a escolha seja feita de maneira irracional se, por exemplo, o sujeito escolhe um dos dois copos de olhos fechados, aleatoriamente, e chega ao copo de água de maneira fortuita.

Em termos filosóficos, considerando a racionalidade ecológica em uma perspectiva procedimental, as normas adequadas para a deliberação de um sujeito variam de acordo com o tipo de agente que está em análise e com o ambiente em que este agente está inserido (MORTON, 2011, p. 573). Afinal, o sucesso dessa deliberação em fazer com que o agente atinja os objetivos que ele almeja depende da interação entre as normas que guiam a sua deliberação e o seu contexto ecológico (MORTON, 2011, p. 574). Nessa perspectiva, não é possível pensar em normas de racionalidade em termos de conteúdo *a priori*; o ambiente e as disposições psicológicas do sujeito indicarão a norma adequada para guiar a sua deliberação.

Em outras palavras, é preciso distinguir claramente entre as normas que utilizamos para avaliar a deliberação de um agente das normas que guiam essa deliberação (MORTON, 2011, p. 569). Para exemplificar, Jennifer M. Morton (2011, p. 569) pede que se imagine um demônio maligno que cria um possível mundo em que deliberar sobre os meios para se atingir determinado fim distancia o sujeito da realização desse fim; em tal mundo ficcional, seguir os seus próprios instintos é uma atitude recompensada, que aproxima o sujeito de seus próprios objetivos. Nesse caso, em termos de racionalidade procedimental, seria recomendado que ele agisse com base nos seus instintos, em vez de tentar deliberar quanto aos meios que deve adotar. Essa é a forma de agir racionalmente em termos ecológicos.

Assim, a racionalidade ecológica avalia em que medida os processos de tomada de decisão levam a resultados acurados, sendo que o critério de acurácia é externo à racionalidade ecológica (SCHMIDT, 2019, p. 521). O foco nessa perspectiva permite que se identifique empiricamente quais normas de deliberação permitem a obtenção de um melhor resultado, dados os fins que devem ser buscados pelo agente decisor (MORTON, 2011, p. 580).[70] Também se abre espaço para que um rol mais amplo de procedimentos seja considerado racional, ainda que em contraposição aos axiomas clássicos da racionalidade, como exposto por Schmidt

[70] Como alerta Morton (2011, p. 580), no entanto, a definição desses fins não é feita de forma empírica; aqui reside a tarefa mais complexa para os filósofos.

(2019, p. 522). De fato, em vez de prescrever a otimização em todas as instâncias, a racionalidade ecológica também defende processos elaborados para serem apenas suficientemente bons (*satisficientes*). Ao mesmo tempo, nem sempre a decisão racional em termos ecológicos envolverá a integração do máximo de informação possível; ignorar deliberadamente parte da informação disponível pode ser uma decisão racional em tal perspectiva. Por outro lado, a racionalidade, aqui, nem sempre está ligada aos processos lentos e deliberativos do Sistema 2; em vez disso, heurísticas simples e regras de decisão formal podem ser consideradas processos decisórios ecologicamente racionais.

Além disso, essa é uma visão moderada de racionalidade ecológica em dois sentidos distintos (SCHMIDT, 2019, p. 523-525). O primeiro ponto é que a racionalidade ecológica é aqui entendida apenas em termos procedimentais. Não se prescreve exatamente quais normas são racionais em termos de conteúdo; em vez de defini-las *a priori*, essa avaliação deve ser feita à luz do ambiente e da disposição psicológica do indivíduo. O segundo ponto é que não se rejeita a possibilidade de existirem aspectos fixos da tomada de decisão racional entre ambientes e agentes; ainda que existam normas universais, elas não conseguem apresentar um panorama geral e adequado que guie ações racionais em termos procedimentais. De fato, ainda que se consiga demonstrar, por exemplo, que métodos de otimização sempre serão a estratégia que leva a melhores resultados - *i.e.* devem ser vistos como norma universal –, na prática, nem sempre esses métodos conseguirão ser implementados, uma vez que a racionalidade dos processos de tomada de decisão é contingente com relação ao contexto e dependente da estrutura psicológica do indivíduo e do ambiente em que ele está inserido.

Nesse ponto, é preciso deixar claro que, diferentemente do que argumentam Gigerenzer e seus seguidores, Kahneman e Tversky não têm uma visão negativa da utilização de heurísticas. Na verdade, eles apenas pretendiam demonstrar que as heurísticas eram importantes e não dizer se elas eram boas ou ruins (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 11). Ainda assim, em diversas oportunidades, eles deixaram claro que heurísticas funcionam bem em determinadas circunstâncias. Desde o início de sua abordagem, Tversky e Kahneman (1974, p. 1124) afirmaram que "[e]m geral, [...] heurísticas são bastante úteis, mas às vezes elas levam a erros severos e sistemáticos".[71] Em outra oportunidade, Kahneman (2000, p. 683) destaca que todas as heurísticas tornam os humanos mais inteligentes de maneira mais frequente do que o contrário.

[71] Tradução livre. Texto original: "*In general, [...] heuristics are quite useful, but sometimes they lead to severe and systematic errors*".

Em outras palavras, a abordagem de Kahneman e Tversky reconhece que heurísticas podem funcionar bem. Em tais situações, utilizar heurísticas para tomar decisões é uma medida ecologicamente racional. Um jogador de beisebol que deve decidir para onde correr diante de uma bola que está no ar não deveria tentar calcular a trajetória da bola a partir de um raciocínio lento e deliberativo, considerando a velocidade inicial, a distância, o ângulo da bola, a resistência do vento e outros fatores (SCHMIDT, 2019, p. 523); afinal, essa seria uma decisão obviamente irracional em termos ecológicos. Em vez disso, o jogador pode se valer de uma heurística simples: a heurística do olhar (*gaze heuristic*), em que o sujeito fixa o olhar na bola, começa a correr e, então, ajusta a velocidade da corrida para que o ângulo do olhar permaneça constante (TODD; GIGERENZER, 2012, p. 6).

Apesar disso, em outras situações, a racionalidade bayesiana ou normas de lógica ou probabilidade podem ser ecologicamente adequadas. Como mencionado acima, muitas decisões cotidianas são tomadas diante de dados percentuais e exigiriam um raciocínio estatístico adequado. Imagine, por exemplo, o sujeito que precisa escolher qual plano de previdência privada melhor se amolda às suas particularidades, balanceando as suas preferências do presente e do futuro. Se ele tem que tomar essa decisão ativamente, para ter *precisão*, regras de probabilidade e lógica seriam fundamentais; ignorá-las em um contexto em que o sujeito tem condições de aplicá-las seria uma escolha irracional em termos ecológicos.

Vale dizer, inclusive, que mesmo *experts* estão sujeitos a vieses que podem levar a decisões irracionais (DHAMI, 2016, p. 1449). Os erros, nesse caso, podem ser até mesmo erros de aplicação, uma vez que pode haver evidência de que as pessoas conhecem e aceitam a regra que deixam de aplicar (KAHNEMAN; TVERSKY, 1982, p. 125); não o fazem apenas pela predominância de processos intuitivos que as levam a não perceber aspectos importantes do ambiente em que decisões serão tomadas. Nesse caso, os próprios agentes decisores, se informados de seu erro, poderiam concordar que tomavam uma decisão inadequada. Pense, por exemplo, num estatístico que falha em aplicar a regra da conjunção a um problema trivial cuja resposta seria fornecida por essa regra.

Por fim, como defende Schmidt (2019, p. 526), para identificar eventuais descompassos entre os procedimentos utilizados para tomar decisões e o contexto ecológico em que essas decisões são tomadas, é essencial que o sujeito avalie os seus próprios processos decisórios. O objetivo, aqui, seria aferir em que medida os processos de tomada de decisão atingem os fins almejados pelo sujeito de maneira confiável em certos ambientes. Para tanto, inevitavelmente, o sujeito terá que se guiar pelas opções que são identificadas como dignas de escolha pela racionalidade de conteúdo. Um sujeito que está tomando decisões de investimento a partir de

processos satisficientes (*satisficing*), por exemplo, pode, diante de um prejuízo, optar por mudar seus procedimentos de escolha e tentar se envolver em processos de deliberação mais lentos, a fim de tentar, na medida do possível, guiar-se pelas normas que conduziriam à otimização de sua utilidade; ou, mesmo, pode delegar essa escolha para um terceiro.

Por todas as razões expostas nesta subseção, não há incompatibilidade entre o Programa de Heurísticas e Vieses e a racionalidade ecológica. Como sustenta Morton (2011, p. 573), as heurísticas descritas por Kahneman e Tversky, embora geralmente levem a decisões frutíferas, também podem conduzir ao erro. O que a pesquisa do Programa de Heurísticas e Vieses evidencia, em todo caso, são as múltiplas maneiras em que as disposições psicológicas de um agente podem variar. Ao mostrar que um agente pode estar sujeito a diferentes tipos de falhas de racionalidade, essa abordagem pode facilitar a identificação de normas que são mais ou menos adequadas para diferentes contextos ecológicos.

1.3.3. Modelos inerentemente experimentais: por que os modelos mentais probabilísticos não parecem ser uma alternativa adequada

Os modelos mentais probabilísticos expostos por Gigerenzer e seus seguidores são bastante controversos. Não é possível concluir imediatamente que esses modelos são utilizados em tarefas de julgamento reais, e, por diversas razões, eles não parecem ser uma alternativa mais promissora do que o modelo de julgamentos exposto pelo programa de heurísticas e vieses. O objetivo desta subseção é expor algumas dessas críticas.

Como exposto anteriormente, Gigerenzer argumenta que a mente seria um estatístico intuitivo, cujos julgamentos são feitos a partir de modelos mentais probabilísticos; o "*Take the Best*", exposto na Subseção 1.2.4 acima, é um desses modelos. A partir de heurísticas rápidas e frugais, os modelos mentais conseguiriam gerar boas inferências em situação de incerteza e, segundo Gigerenzer, chegariam a melhor resultado do que aqueles obtidos a partir de algoritmos baseados na integração de informações.[72] Em outras palavras, esses modelos levariam ao efeito menos-é-mais (*less-is-more effect*), já que, em certas situações, o fornecimento de mais informações para os indivíduos levava a menos precisão em suas inferências, se comparadas com aquelas feitas a partir das heurísticas dos modelos mentais.

Apesar disso, nos modelos mentais probabilísticos, as heurísticas - isto é, os sinais e classes de referência que o agente decisor utilizaria - costumam ser propostas pelos autores dos

[72] Cf., por exemplo, (GIGERENZER; GOLDSTEIN, 1996).

experimentos (DHAMI, 2016, p. 1426; DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 29). Essa não é uma abordagem realista. De um lado, as pessoas podem ter diferenças em suas crenças a respeito da classe de referência do objeto em análise que pode ser relevante para a comparação (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 29); isto é, pode ser que diferentes pessoas discordem da validade ecológica dos sinais analisados heuristicamente. Retomando o exemplo utilizado na Subseção 1.2.4 acima, pode ser que, para um determinado indivíduo, ter um time na *Bundesliga* seja um sinal ecologicamente válido para se julgar a medida em que uma cidade tem mais habitantes que outra; ao mesmo tempo, esse mesmo sinal pode ser irrelevante para um outro agente decisor.

Além disso, considerando a forma pela qual as heurísticas são apresentadas para as pessoas nos experimentos de Gigerenzer e de seus seguidores, é difícil definir como os indivíduos escolhem os sinais específicos em um ambiente específico. Como argumenta Dhami (2016, p. 1426), pode ser que os sinais sejam buscados a partir de outras heurísticas simples[73]; talvez, a partir do simples aprendizado associado a sinais do passado associados a resultados específicos; ou pode ser que os humanos tenham adquirido a habilidade de gerar simples sinais e heurísticas a partir da evolução. O exemplo da heurística do olhar (*gaze heuristic*) para pegar uma bola de beisebol é ilustrativo nesse sentido. No entanto, não há clareza a respeito de como heurísticas rápidas e frugais poderiam ser utilizadas para resolver problemas econômicos, como a elaboração de contratos, a escolha de planos de previdência ou de políticas monetárias. Nesses casos, a experiência evolutiva é de pouca utilidade e não é possível concluir com clareza quais heurísticas rápidas e frugais deveriam ser utilizadas.

Também não é claro se, em situações da realidade, as pessoas realmente usam heurísticas rápidas e frugais (DHAMI, 2016, p. 1426). Kahneman e Tversky identificaram as heurísticas descritas em seu programa a partir da observação de comportamentos de indivíduos em experimentos. Na abordagem de Gigerenzer, porém, o objetivo parece ser, antes de mais nada, demonstrar a possibilidade de heurísticas rápidas e frugais gerarem melhores inferências que algoritmos que integram informações. Nem sempre é claro, todavia, se, em tarefas reais de julgamento, as pessoas realmente utilizariam essas heurísticas. Goldstein e Gigerenzer (1999) chegaram a identificar que decisões se conformavam às previsões da heurística do reconhecimento, mas não há demonstrações suficientes da aplicação prática de outras heurísticas rápidas e frugais a ponto de se desenvolver uma teoria abrangente.

[73] Nesse caso, pode-se chegar a um problema de infinito regresso. Afinal, para definir as heurísticas utilizadas na busca por sinais, seria necessário definir como as pessoas chegariam a essa primeira heurística e assim sucessivamente, o que torna o modelo psicologicamente implausível. Nesse sentido, Cf.: (DHAMI, 2016, p. 1411).

Mesmo a heurística do reconhecimento é controversa. De fato, segundo Daniel Oppenheimer (2003), Goldstein e Gigerenzer (1999) confundem o reconhecimento de uma das opções com o conhecimento prévio dos participantes de que uma das cidades era maior do que a outra. Em outras palavras, a maioria das pessoas não só reconhece Berlim, como também sabe que ela é uma das maiores cidades da Alemanha. Não é possível determinar se os julgamentos foram feitos a partir do reconhecimento ou do conhecimento prévio de que as cidades eram maiores. Nessas condições, para Oppenheimer, qualquer modelo de julgamento - simples ou complexo - iria prever resultados semelhantes.

Por outro lado, de acordo com a heurística do reconhecimento, nenhuma informação além do mero reconhecimento seria utilizada; isto é, nenhuma informação seria capaz de reverter a escolha indicada pelo reconhecimento. Nessa perspectiva, uma cidade reconhecida seria julgada como maior do que a desconhecida, mesmo se a reconhecida fosse sabidamente menor. Oppenheimer desenvolveu um experimento para verificar se isso aconteceria e constatou que, em cinco de cada seis casos, os participantes apresentavam menor probabilidade de julgar a cidade reconhecida como maior; isto é, quando uma cidade reconhecida é pequena, as pessoas tendem a julgá-la como menor do que uma cidade não reconhecida.

Oppenheimer afirma, porém, que o resultado de seu experimento poderia ter sido motivado pelo fato de as pessoas reconhecerem que o questionário buscava cidades locais de maneira desproporcional. Sabendo que essas cidades são pequenas, os participantes podem ter decidido que a heurística do reconhecimento não estaria adaptada àquele ambiente e optado por utilizar uma heurística distinta. Para testar a prevalência da heurística do reconhecimento e, ao mesmo tempo, controlar esse fator de confusão, Oppenheimer elaborou um outro experimento; desta vez, as cidades escolhidas não eram conhecidas por serem grandes, mas por razões diversas. Chernobyl foi uma das cidades escolhidas, por exemplo. Nesse caso, não haveria muitos fatores além da heurística do reconhecimento para conduzir as decisões. Oppenheimer constatou, entretanto, que os participantes novamente apresentavam uma baixa probabilidade de utilizar o reconhecimento para guiar suas decisões. Com base nessas evidências, Oppenheimer conclui que o julgamento heurístico não seria tão rápido e frugal quanto esperado na visão de Gigerenzer e de outros autores dessa abordagem teórica, que dariam uma importância muito elevada ao reconhecimento na tomada de decisões.

Além disso, como já demonstrado acima, em modelos mentais como o algoritmo "*Take The Best*", quando não há sinais discriminantes, as pessoas teriam que escolher aleatoriamente. De fato, talvez, em muitos cenários, as pessoas não tenham sinais na memória capazes de levá-las a inferências adequadas. Esse tipo de escolha aleatória pode ser influenciado por aspectos

da arquitetura de escolhas, como a ordem em que as opções são apresentadas, e pode levar a resultados desastrosos. Decidir com base em heurísticas, nesses contextos, pode simplesmente ser uma estratégia irracional em termos ecológicos. Afinal, naquele ambiente específico, estratégias deliberativas, ao menos inspiradas pelos modelos propostos pela teoria da escolha racional, podem ser opções mais adequadas em termos de racionalidade ecológica.

Por fim, como expõe Dhami (2016, p. 1426), Gigerenzer e os demais proponentes do programa de heurísticas rápidas e frugais não estabelecem as condições em que essas heurísticas superam o desempenho de procedimentos estatísticos ou abordagens de otimização. Para Dhami, essas heurísticas podem até superar abordagens estatísticas para resolver problemas quando a estratégia ótima de solução não for conhecida. Se ela for conhecida, no entanto, nenhuma heurística seria capaz de superá-la. Por exemplo, se as curvas de demanda e de custos são conhecidas, a receita marginal será igual aos custos marginais. Nenhuma heurística chega a alguma solução melhor do que essa. Como a abordagem de Gigerenzer não é clara a respeito de quando as heurísticas rápidas e frugais levam a melhores decisões, torna-se difícil utilizá-las para descrever um padrão amplo de julgamentos com racionalidade limitada.

1.3.4. Validação empírica dos vieses comportamentais e fatores supostamente insignificantes que influenciam decisões e julgamentos

Um dos pontos mais reiterados por Gigerenzer - e, aqui, exposto na Subseção 1.2.2 - é a possibilidade de fazer com que os vieses expostos por Kahneman e Tversky desapareçam se apresentados em formato de frequências, em vez de probabilidades. Essa não parece ser uma crítica justa. Para além da questão normativa, exposta no item 1.3.1 acima, a crítica de Gigerenzer não se sustenta do ponto de vista empírico, além de não ser completamente fiel ao programa de heurísticas e vieses.

Como Kahneman e Tversky (1996) argumentaram, eles chegaram efetivamente a investigar o processo pelo qual diferentes representações e tarefas evocam diferentes heurísticas. Mais do que isso, Tversky e Kahneman (1983) chegaram a demonstrar que a exposição de problemas em termos de frequência relativa ajudava os participantes de um experimento a evitar a falácia da conjunção.[74] Na mesma oportunidade, Tversky e Kahneman

[74] O problema, nessa oportunidade, havia sido exposto inicialmente como uma questão a respeito do percentual de homens com mais de 55 anos que tenham tido um ou mais ataques cardíacos. Em termos frequentistas, a questão foi colocada nos seguintes termos: "Um questionário de saúde foi conduzido em uma amostra de 100 homens adultos na Colúmbia Britânica, com todas as idades e ocupações. Por favor, apresente a sua melhor estimativa dos

chegaram a mencionar que julgamentos de probabilidade variam de acordo com o grau em que eles seguem uma abordagem holística e com o grau em que a análise das probabilidades é analítica ou intuitiva. De um lado, existem questões que podem ser respondidas pelo cálculo da frequência relativa de resultados favoráveis. Aqui, a análise é frequentista, algorítmica e realizada em uma perspectiva decomposta. É o caso, por exemplo, da avaliação das chances de ganhar de uma determinada mão no pôquer. Por outro lado, há questões que normalmente são avaliadas de maneira holística e intuitiva; por exemplo, quando se pergunta qual a probabilidade de que uma testemunha esteja falando a verdade. A decomposição e os cálculos realizados de maneira algorítmica e frequentista oferecem alguma proteção contra vieses, mas, como o elemento intuitivo não pode ser eliminado, vieses continuam a acontecer em alguma medida.

Ao mesmo tempo, como sustentam Kahneman e Tversky (1996), a afirmação de que ilusões cognitivas desaparecem quando as pessoas avaliam frequências é baseada em uma leitura seletiva das evidências apresentadas pelo programa de heurísticas e vieses. De fato, Kahneman e Tversky chegam a apresentar dados em termos de frequências na identificação de heurísticas como a disponibilidade e a ancoragem. Outras heurísticas como a regressão à média, o viés de confirmação, a heurística do afeto e o viés de retrospectiva não são afetadas pela distinção entre frequências e probabilidades (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 21). Além disso, mesmo nos casos em que a apresentação em termos de frequência reduz a intensidade de vieses, como a falácia da conjunção, não se trata de uma panaceia: em grande medida, vieses significativos permanecem acontecendo na aplicação da regra de Bayes (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 21).

De qualquer maneira, como demonstrado na Subseção 1.3.1 acima, ainda que as pessoas tenham evoluído para compreender melhor dados expostos em termos de frequências, boa parte dos dados econômicos no mundo real são apresentados em termos percentuais (DHAMI, 2016, p. 1429; DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 23). Dessa forma, ainda que frequências possam reduzir vieses cognitivos, as pessoas costumam fazer estimativas de probabilidade sem ter acesso às frequências naturais. Assim, ainda que se admita como verdadeira a premissa de Gigerenzer de que os vieses podem desaparecer, as heurísticas e vieses descritos por Kahneman e Tversky são fundamentais para entender os julgamentos e a tomada de decisões em situações em que a informação é apresentada de maneira probabilística.

Por outro lado, outra crítica frequentemente reiterada por Gigerenzer é a suposta indefinição das heurísticas descritas por Kahneman e Tversky. Nessa perspectiva, como

---

seguintes valores: Quantos participantes tiveram um ou mais ataques cardíacos? Quantos participantes têm mais de anos e tiveram um ou mais ataques cardíacos" (TVERSKY; KAHNEMAN, 1983, p. 309, tradução livre).

demonstrado na Subseção 1.2.3, as heurísticas seriam vagas e indefinidas. Mais uma vez, porém, a crítica não parece ser bem direcionada. Como explicado por Dhami, al-Nowaihi e Sunstein (2019), uma marca da economia comportamental tem sido apresentar uma abordagem formal que define matematicamente as heurísticas e vieses em questão. Expor esses modelos formais ultrapassa os objetivos desta dissertação, embora a sua existência indique que as heurísticas de Kahneman e Tversky não são indefinidas como Gigerenzer faz parecer[75].

Feitas essas ressalvas, para concluir este Capítulo, o que se pretende argumentar é que, independentemente da posição que se adota na controvérsia exposta acima, há toda uma série de fatores supostamente irrelevantes que influenciam o comportamento humano (THALER, 2015, p. 9). Esses fatores - não capturados por modelos econômicos abstratos com base nos quais políticas públicas são formuladas - diminuem a precisão das descrições de comportamento de tais modelos. A consequência é menor precisão em suas previsões, de modo que incorporar aspectos mais próximos da realidade e rever esses modelos pode ser um meio de se obter resultados melhores.

Ao longo das últimas décadas e com o desenvolvimento da economia comportamental, uma torrente de estudos, com métricas diversas, tem enumerado diversos desses fatores supostamente irrelevantes: comportamento míope e excessivamente focado no presente[76]; influência significativa da opção padrão (*default*)[77], ignorância de atributos que não chamam a atenção, mas que são importantes para as escolhas[78]; procrastinação[79]; tendência a permanecer com a opção do *status quo*[80]; aversão à perda[81]; otimismo irrealista[82]; limitações de capacidade cognitiva e de controle executivo, que são potencializadas em situações de escassez[83]; dentre outros inúmeros elementos já catalogados na literatura.[84] Ainda que parte desses fatores decorram dos vieses cognitivos descritos pelo programa de heurísticas e vieses, claramente, os processos cognitivos dos quais eles se originam são diversos e ultrapassam a obra original de Kahneman e Tversky.

Gigerenzer (2016, p. 37-38) afirma que a economia comportamental apoiaria uma contrarrevolução neoclássica que ofusca a visão de Simon e faz com que ela seja ignorada. Os

---

[75] Para definições formais das heurísticas do programa de Kahneman e Tversky, Cf.: (DHAMI, 2016, p. 7; DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 17-20).
[76] Cf., *e.g.*, (MURAMATSU; FONSECA, 2008; RICK; LOEWENSTEIN, 2015).
[77] Cf., *e.g.*, (JACHIMOWICZ *et al.*, 2019; JOHNSON; GOLDSTEIN, 2003).
[78] Cf., *e.g.*, (BORDALO; GENNAIOLI; SHLEIFER, 2012; KAHNEMAN, 1973; SIMONS; CHABRIS, 1999).
[79] Cf., *e.g.*, (O'DONOGHUE; RABIN, 2001; STEEL, 2007).
[80] Cf., *e.g.*, (KAHNEMAN; KNETSCH; THALER, 1991; SAMUELSON; ZECKHAUSER, 1988).
[81] Cf., *e.g.*, (KAHNEMAN; KNETSCH; THALER, 1991; KAHNEMAN; TVERSKY, 1984).
[82] Cf., *e.g.*, (SHAROT, 2011).
[83] Cf., *e.g.*, (HAUSHOFER; FEHR, 2014; MULLAINATHAN; SHAFIR, 2016).
[84] Cf. nota 42 *supra*.

seus modelos, em vez de mais realistas, teriam se rendido à metodologia como-se (*as-if*)[85] e manteriam a estrutura da utilidade esperada, com acréscimos de novos parâmetros a partir de rótulos psicológicos. Para Gigerenzer, essa abordagem seria como acrescentar epiciclos de Ptolomeu[86] à astronomia e levaria a teorias mais irrealistas do que as teorias criticadas da economia neoclássica. Mais uma vez, essa não parece ser uma crítica justa.

Como expõem Dhami, al-Nowaihi e Sunstein (2019, p. 37–38), a otimização é simplesmente uma ferramenta utilizada pela economia comportamental, que, no entanto, não é favorecida de maneira *a priori*; e, enquanto ferramenta, a otimização pode ser útil. Não deve haver uma rejeição *per se* de sua utilização, mas a demonstração empírica das condições em que ela falha. A economia comportamental, de maneira geral, costuma estar aberta para essas demonstrações e não tem aversão a explorar alternativas ou levar a racionalidade procedimental em consideração.

De qualquer forma, a partir dos dados identificados nos inúmeros experimentos da economia comportamental - dos quais alguns exemplos foram mencionados acima -, verifica-se que teorias que são centrais para modelos econômicos baseados em otimização são inconsistentes com as evidências experimentais (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019, p. 37). Mais do que isso o que mais interessa para os objetivos desta dissertação é que, frequentemente, fatores supostamente irrelevantes levam a erros que podem gerar prejuízo em termos de bem-estar e podem impedir que as pessoas atinjam os seus próprios objetivos.

O uso abusivo de cigarros ou o consumo excessivo de alimentos gordurosos, por exemplo, podem ser comportamentos bastante difundidos, cujas raízes podem estar ligadas a um comportamento míope e excessivamente focado no presente. Tais comportamentos poderiam ser vistos como internalidades, que acontecem quando as pessoas tomam decisões que prejudicam seus *selves* futuros (SUNSTEIN, 2014a, p. 37-38). Outros exemplos de fatores supostamente irrelevantes que impactam as decisões das pessoas e originam erros que podem gerar prejuízos serão apresentados ao longo desta dissertação. De qualquer forma, considerando esses fatores, reconhecer que há um erro ou um comportamento inadequado independe de compromissos muito rígidos com um determinado modelo normativo; pode ser, inclusive, que

[85] Cf. nota 28 *supra*.

[86] No modelo astronômico de Ptolomeu, a Terra ficaria imóvel próxima ao centro do Universo e os planetas girariam ao seu redor, em órbitas circulares (epiciclos). O modelo foi superado pelo modelo heliocêntrico de Copérnico. Cf. em: http://astronomia.blog.br/epiciclos-de-ptolomeu-e-materia-escura/. Acesso em: 15 fev. 2022.

os próprios indivíduos reconheçam a inadequação de seu comportamento, quando o analisam de maneira reflexiva[87].

1.4. Principais conclusões do Capítulo 1

O objetivo deste Capítulo foi expor as premissas comportamentais adotadas nesta dissertação. Inicialmente, o Capítulo expôs as características do programa de heurísticas e vieses, desenvolvido por Kahneman e por Tversky e seus seguidores, e do programa de heurísticas rápidas e frugais, apresentado por Gigerenzer e seus colegas. Em seguida, foram expostos quatro argumentos principais, que podem ser resumidos da seguinte forma:

1. Seguir normas de probabilidade, como a abordagem *bayesiana*, pode ser necessário à luz do caso concreto, quando se tem como objetivo ter crenças *mais precisas e calibradas*. Não se espera, por exemplo, que um médico excessivamente confiante em seus prognósticos tome decisões *tão boas* quanto um médico calibrado.
2. O programa de heurísticas e vieses não é incompatível com uma visão moderada da racionalidade ecológica em termos procedimentais. Embora heurísticas levem a decisões adequadas em muitos contextos, *i.e.*, ecologicamente racionais, em outras ocasiões, as características de um ambiente podem indicar que procedimentos heurísticos conduzem a erro e podem exigir estratégias de otimização.
3. Os modelos mentais probabilísticos de Gigerenzer não parecem modelos efetivamente representativos da realidade; muitas de suas características são fornecidas apenas experimentalmente e não é claro se as heurísticas que o modelo propõe são realmente utilizadas na realidade. Por isso, tal modelo não parece ser um substituto adequado ao modelo proposto por Kahneman e Tversky.
4. Vieses cognitivos nem sempre desaparecem quando avaliados em termos de frequências naturais e essa avaliação nem sempre é possível. Mais do que isso, uma vasta quantidade de experimentos de economistas e cientistas comportamentais vêm catalogando, ao longo das últimas décadas, toda sorte de *fatores supostamente irrelevantes* que impactam a tomada de decisões e podem causar prejuízos.

[87] Nesse caso, como mencionado acima, os erros seriam *erros de aplicação*, uma vez que os próprios conhecem e aceitam a regra que deixam de aplicar ou a decisão mais adequada que deveriam tomar naquele contexto (KAHNEMAN; TVERSKY, 1982, p. 125).

Reconhecer que a tomada de decisões intuitivas leva a erros sistemáticos e a decisões de limitada racionalidade em termos ecológicos é um passo essencial para a análise que será empreendida nesta dissertação. A partir dessa constatação, nos Capítulos 2 e 3, pretende-se demonstrar como o reconhecimento empírico de erros cognitivos pode justificar um maior grau de intervenção estatal na vida privada e servir como base para o desenvolvimento de novas técnicas regulatórias.

2. FALHAS DE MERCADO COMPORTAMENTAIS: NOVAS JUSTIFICATIVAS PARA A INTERVENÇÃO ESTATAL NA VIDA PRIVADA

Uma célebre passagem do clássico *Sobre a Liberdade*, escrito pelo filósofo britânico John Stuart Mill (2017), introduz a discussão que será apresentada neste Capítulo. Naquela ocasião, Mill (2017, p. 81-82) pretendia defender um princípio que deveria condicionar todas as formas da sociedade de se relacionar com o indivíduo, incluindo compulsão e controle, penalidades legais impostas pela força física e coerção moral pela opinião pública. Esse princípio é o princípio do dano, segundo o qual "o único propósito para o qual o poder pode ser exercido com justiça sobre qualquer membro da comunidade civilizada, contra sua vontade, é o de evitar dano a outros" (MILL, 2017, p. 82).

O princípio do dano é justificado por Mill (2017, p. 169) a partir da visão de que cada indivíduo é a pessoa mais interessada em seu próprio bem-estar e é quem tem o conhecimento mais profundo sobre seus próprios sentimentos e suas próprias circunstâncias. A sociedade, em contrapartida, tem apenas interesse indireto no sujeito como indivíduo. Nesse caso, intervenções que restringissem a liberdade individual só seriam justificadas quando houvesse dano ou risco de dano, para um indivíduo ou para a sociedade, ou quando fosse desrespeitado um padrão de conduta social que consiste em não prejudicar os interesses de outras pessoas e suportar sua parcela dos sacrifícios necessários para defender a sociedade ou seus membros de prejuízos (MILL, 2017, p. 167-168).

Mill (2017, p. 175-176) até admite que indivíduos possam atuar de maneira contrária ao seu próprio bem-estar - por exemplo, consumindo bebidas excessivamente ou destruindo a sua própria propriedade. Segundo o filósofo britânico, porém, o meio adequado de fazer com que essas pessoas se conformassem ao padrão de conduta racional não seria esperar que elas fizessem algo racional e então puni-las, legal ou moralmente (MILL, 2017, p. 177-178). Na verdade, durante a infância, a sociedade teria absoluto poder sobre as pessoas para torná-las capazes de agir racionalmente na vida adulta. Em caso de insucesso, isto é, se qualquer número considerável de pessoas crescesse como crianças, incapazes de determinar suas ações por considerações racionais, a própria sociedade seria culpada pelas consequências correspondentes.

Essa visão, exemplificada na obra de Mill, é responsável por influenciar a certa aversão a qualquer tipo de paternalismo que é comum na tradição político-econômica liberal, marcada

por economias de mercado[88]. Mais do que isso, essa tradição influenciou a economia do bem-estar neoclássica, cujos modelos econômicos justificam a regulação do comportamento de indivíduos, quando constatada uma falha de mercado (DHAMI, 2016, p. 1579), isto é, diante da "incapacidade que alguns mercados não regulamentados têm de alocar recursos com eficiência" (MANKIW, 2013, p. 144). Nesse caso, considera-se que, em geral, o mercado livre seria capaz de resolver os problemas e que a soberania dos consumidores[89] deve ser respeitada. Em outras palavras, os consumidores são considerados os melhores juízes de seu próprio bem-estar (DHAMI, 2016, p. 1577).

A partir do desenvolvimento da economia comportamental e da identificação de inúmeros fatores supostamente irrelevantes que influenciam as decisões, conforme exposto no Capítulo 1, as premissas da posição antipaternalista passaram a ser questionadas. De fato, a partir de questionamentos relacionados à capacidade de indivíduos satisfazerem realmente suas preferências a partir de suas escolhas, não só Thaler e Sunstein como diversos outros pesquisadores[90] passaram a oferecer novos fundamentos paternalistas para a ação estatal, que não se confundem com os fundamentos clássicos do referencial neoclássico.

Em geral, boa parte dos autores que partem de tais premissas passaram a propor paternalismos suaves (*soft paternalisms*), em oposição aos paternalismos tradicionais, baseados em medidas de comando em controle.[91] Ao lado dos paternalismos suaves, porém, há ainda quem defenda que os fatores supostamente irrelevantes justificam medidas paternalistas coercitivas, para corrigir o rumo de ação das pessoas. Paternalismos comportamentais é a expressão utilizada nesta dissertação para designar tanto os paternalismos suaves, quanto o paternalismo coercitivo que também se origina de problemas na tomada de decisões em razão

[88] Como afirma Jean Tirole (2020, p. 166), "[a] maioria da população mundial vive numa economia de mercado mesclada com intervenção do Estado. Organização da sociedade que apreciamos, toleramos ou detestamos, embora nem sempre nos perguntemos se outros pontos de vista são possíveis. Desde o retumbante fracasso econômico, cultural, social e ambiental das economias planejadas, vemo-nos diante de uma espécie de fatalismo, temperado pela indignação de alguns". Esse é o caso do Brasil (1988), que define, em sua Constituição, que a ordem econômica será fundada em princípios como a propriedade privada, a função social da propriedade, a livre concorrência, a redução das desigualdades nacionais, a defesa do consumidor etc. Assegura, ainda, a todos o livre exercício de qualquer atividade econômica, independentemente de autorização de órgãos públicos, salvo nos casos previstos por lei.

[89] Thaler (2015, p. 268-269) afirma que a soberania do consumidor é o princípio basilar das crenças libertárias da Escola de Chicago. A ideia, como visto, é que as pessoas fazem boas escolhas, e, certamente, melhores escolhas que qualquer um poderia fazer por elas.

[90] Cf., por exemplo: (CAMERER *et al.*, 2003; CONLY, 2013; LOEWENSTEIN; CHATER, 2017; LOEWENSTEIN; HAISLEY, 2007).

[91] Como será demonstrado, esse não é o caso de Sarah Conly (2013). Diante de *fatores supostamente irrelevantes* que criam prejuízo de bem-estar, a autora defende medidas paternalistas coercitivas, já que a autonomia perderia seu fundamento sem a soberania dos consumidores.

de limitações dos agentes decisores. Em ambos os casos, falhas de mercado comportamentais fundamentam a regulação do comportamento privado.

O objetivo deste Capítulo é expor o desenvolvimento de paternalismos comportamentais, colocando-os lado a lado com os fundamentos utilizados pelo regulador tradicional. A ideia é demonstrar como fatores supostamente irrelevantes levaram ao desenvolvimento de novas razões para a regulação do comportamento privado, especialmente razões de caráter paternalista. Feita essa exposição, serão demonstrados os diferentes movimentos paternalistas comportamentais: o paternalismo libertário, o paternalismo assimétrico, o paternalismo leve pragmático e o paternalismo coercitivo comportamental. O foco da análise, naturalmente, será direcionado ao paternalismo libertário, por ser o mais célebre e o mais influente de tais movimentos. Qualquer discussão que envolve o paternalismo e os limites à regulação do comportamento privado, naturalmente, atrai preocupações ligadas à autonomia. Por essa razão, o Capítulo se encerra com a exposição do conceito de autonomia adotado neste trabalho.

2.1. As falhas de mercado: tradicionais motivos para a regulação do comportamento privado

A intenção de discutir as falhas de mercado tradicionais, já conhecidas e amplamente divulgadas em qualquer manual de introdução à economia, se justifica por um motivo muito simples: os paternalismos comportamentais, em geral, assim como a teoria que lhes é subjacente, surgem a partir da constatação da incompletude das razões tradicionais para a regulação do comportamento privado. Considerando essa intenção, esta subseção apresentará alguns exemplos clássicos, sem a pretensão de esgotá-los ou de replicar aqui o que seria o objeto de um livro-texto de economia. Do mesmo modo, também não se pretende defender ou prescrever o arcabouço regulatório neoclássico como caminho a ser seguido. O propósito desta seção - ressalta-se, mais uma vez - é preparar o terreno para as seções subsequentes, em que fundamentos comportamentais para a regulação estatal serão expostos. Feita essa ressalva, pode-se então passar à exposição das clássicas situações em que o mercado não é eficiente.

A teoria do mercado indica que o processo de compensação do mercado (*market-clearing*), isto é, o processo que leva a preços de equilíbrio, em que a oferta e a demanda são

idênticas, é o arranjo que obtém a alocação mais eficiente[92] dos recursos da sociedade (DUNNE, 2015, p. 7). Afinal, se os indivíduos tentam maximizar sua própria utilidade, como defendem os axiomas econômicos neoclássicos, no arranjo de mercado, a soma agregada de bem-estar total atinge seu máximo valor. Essa não é, no entanto, uma situação automática; algumas premissas devem ser cumpridas para que um arranjo de mercado leve a uma alocação de recursos eficiente. Becker (1976, p. 5), por exemplo, afirma que a combinação das premissas de um comportamento maximizador, do equilíbrio do mercado e de preferências estáveis forma o cerne de sua abordagem econômica – e são, portanto, as premissas que devem ser cumpridas para que um arranjo de mercado seja eficiente. Requisitos adicionais podem incluir informações suficientes para ter a capacidade de maximizar suas escolhas, a ausência de externalidades e mercados competitivos[93].

Na realidade, porém, mercados apresentam muitas falhas e nem sempre conseguem gerar uma alocação de recursos eficiente. Essas situações são falhas de mercado. Para ilustrá-las, a seguinte questão pode ser útil: se dois agentes fazem um negócio com mútuo consentimento, por que essa transação pode ser um problema para a sociedade, a ponto de dar origem a uma intervenção do Estado (TIROLE, 2016, p. 167)?

Uma primeira hipótese pode ser, simplesmente, a existência de incentivos para que esses agentes provoquem danos a terceiros, que, por definição, são contrários ao negócio. São as chamadas externalidades[94]. Como exemplifica Tirole (2020, p. 167), "[u]ma companhia de energia que produz eletricidade a partir do carvão emite poluentes (dióxido de enxofre, óxidos de nitrogênio) que geram chuvas ácidas ou gases de efeito estufa (dióxido de carbono)." Os

[92] Na economia do bem-estar, os critérios comumente utilizados para mensurar a eficiência – e, portanto, apurar o crescimento ou o desperdício de riqueza – são os critérios de Pareto, elaborado pelo economista e sociólogo italiano Vilfredo Pareto, e de Kaldor-Hicks, elaborado pelo economista húngaro Nicholas Kaldor e pelo economista inglês John Richard Hicks. No critério de Pareto, a situação ideal é aquela "em que os recursos de uma economia são alocados de tal maneira que nenhuma reordenação diferente possa melhorar a situação de qualquer pessoa (ou agente econômico) sem piorar a situação de qualquer outra" (SILVA, 2013, p. 91). Ou seja, em tais situações, a riqueza é maximizada quando se atinge um ponto em que é possível obter uma melhora de bem-estar sem que ninguém se prejudique com isso. No critério de Kaldor-Hicks, "como é improvável, na maior parte da dinâmica econômica, melhorar a situação de todos sem piorar a de alguém, a eficiência potencial se satisfaz se o incremento total de riqueza tiver saldo líquido, ou seja, se for positivo o bem-estar final" (SILVA, 2013, p. 91).

[93] Nesse sentido, Cf.: (DUNNE, 2015, p. 8; OGUS, 2004, p. 23–24).

[94] Como define Mankiw (2013, p. 184), "[u]ma externalidade surge quando uma pessoa se dedica a uma ação que provoca impacto no bem-estar de um terceiro que não participa dessa ação, sem pagar nem receber nenhuma compensação por esse impacto. Se o impacto sobre o terceiro é adverso, é denominado externalidade negativa. Se é benéfico, é chamado externalidade positiva. Quando há externalidades, o interesse da sociedade em um resultado de mercado vai além do bem-estar dos compradores e dos vendedores que participam do mercado; passa a incluir também o bem-estar de terceiros que são indiretamente afetados".

mecanismos de mercado, nesse caso, não conseguem proteger[95] a população afetada por conta própria, o que demanda a intervenção estatal. Esse é o principal fundamento econômico de políticas de proteção do meio ambiente (TIROLE, 2020, p. 167–168), cuja defesa é um dos princípios gerais da atividade econômica no Brasil[96].

Há externalidades negativas e positivas. Externalidades negativas são como o exemplo da poluição apresentado acima: ao maximizar suas preferências, o agente gera danos para terceiros, sem, necessariamente, gerar maior custo para si mesmo. Nesse cenário, o agente produzirá maior quantidade do bem do que seria desejável para a sociedade. A intervenção, nesse caso, faz com que o agente interiorize os custos gerados a partir da externalidade; pode-se pensar, por exemplo, em impostos sobre a poluição[97]. Externalidades positivas, por seu turno, são exatamente o contrário: o agente gera benefícios à sociedade sem que esses benefícios sejam integralmente incorporados por si próprio. Um sujeito que se vacina, por exemplo, gera mais benefícios para a sociedade como um todo do que para si mesmo. A consequência é um volume de vacinação subótima, o que leva o governo a tomar medidas para subsidiar a vacinação[98].

Por outro lado, pode ser que o negócio não possa ser executado de maneira suficientemente informada por alguma das partes. Se, por exemplo, o comprador não conhecer a periculosidade de um medicamento ou de um determinado produto, ele pode subestimar os seus riscos e tomar uma decisão ineficiente (TIROLE, 2020, p. 168). De fato, são inúmeras as situações em que há assimetrias de informações, de modo que um dos agentes tem acesso a um volume maior de informações do que o outro que faz parte da transação. A falta de consciência

---

[95] Uma solução *de* mercado até poderia ser especulada. De acordo com o Teorema de Coase, batizado em homenagem ao economista Ronald Coase (1960), independentemente da distribuição inicial de direitos, as partes interessadas sempre podem chegar a um acordo em que todos ficam numa situação melhor e em que o resultado é eficiente. No caso do exemplo mencionado, a empresa e a população afetada poderiam negociar, de modo que a empresa pagasse diretamente à população pelo direito de poluir, ou, por outro lado, a sociedade poderia pagar para que a empresa instalasse a sua fábrica em outro local ou deixasse de poluir. Na prática, porém, diferentemente da situação ideal imaginada por Coase, há custos de transação e um acordo nem sempre é possível. Quando o número de partes interessadas é grande, como no exemplo mencionado, o custo para coordenar todos os envolvidos será excessivamente dispendioso e pode tornar quase impossível uma negociação (MANKIW, 2013, p. 198).

[96] De acordo com o inciso VI do art. 170 da Constituição da República (BRASIL, 1988), a "defesa do meio ambiente, inclusive mediante tratamento diferenciado conforme o impacto ambiental dos produtos e serviços e de seus processos de elaboração e prestação, é um dos princípios que devem ser observados pela ordem econômica.

[97] Mankiw, nesse sentido, considerando o exemplo de uma fábrica de alumínio que gera poluição, afirma que uma maneira de atingir um resultado ótimo seria tributar os produtores de alumínio por tonelada vendida. Nesse caso, "[o] imposto deslocaria a curva de oferta de alumínio para cima no montante do imposto. Se o imposto refletisse exatamente o custo externo da fumaça lançada na atmosfera, a nova curva de oferta coincidiria com a curva de custo social. No novo equilíbrio de mercado, os produtores de alumínio produziriam a quantidade socialmente ótima de alumínio" (MANKIW, 2013, p. 186). Em outras palavras, a externalidade é *internalizada*, já que os agentes que a produzem passam a ter incentivos para levar os seus efeitos em consideração.

[98] A respeito das soluções para externalidades positivas, Mankiw (2013, p. 187) menciona que "o governo pode corrigir as falhas de mercado induzindo os participantes a internalizar a externalidade. A resposta apropriada, no caso de externalidades positivas, é exatamente oposta ao caso das externalidades negativas. Para deslocarem o equilíbrio de mercado social ótimo, as externalidades positivas requerem um subsídio".

adequada do risco de determinadas condutas é um dos fundamentos - além das já expostas externalidades - de regras como a obrigação de utilizar capacetes por motociclistas, cintos de segurança, seguros obrigatórios para danos causados por veículos automotores[99] e certificados de que veículos estão aptos a dirigir[100] (DHAMI, 2016, p. 1579).

Ao mesmo tempo, pode ser, por exemplo, que um sujeito – o agente – tome decisões em nome de outra pessoa - o principal. Se o principal não tiver meios de monitorar o comportamento do agente, haverá incentivos para que ele não se esforce o bastante para atingir os objetivos do principal. Há, nesse caso, uma situação de risco moral[101]. Inúmeras regras são desenvolvidas para incentivar o agente a atuar de maneira responsável. As regras que disciplinam a relação entre acionistas, administradores e outros envolvidos na vida de uma sociedade, por exemplo, são estruturadas para reduzir o risco moral inerente a essas situações[102].

Outro problema derivado da assimetria de informação é a seleção adversa, em que um comprador corre o risco de comprar um bem de baixa qualidade, já que o vendedor detém mais informações do que ele (MANKIW, 2013, p. 444). O exemplo clássico envolve o mercado de automóveis usados, em que os indivíduos não conseguem saber se estão adquirindo um veículo bom ou ruim (um limão, na expressão utilizada nos Estados Unidos) (AKERLOF, 1970). O proprietário, por outro lado, depois de possuir o automóvel por algum tempo, consegue estimar com mais precisão a probabilidade de seu veículo ser um limão. Em decorrência de tal assimetria de informações, carros usados acabam sendo vendidos por preços muito baixos, e os proprietários de bons veículos não conseguem receber o valor verdadeiro do seu carro nem o valor esperado de um carro novo. Regras como a certificação de produtos com o selo do Instituto Nacional de Metrologia, Qualidade e Tecnologia (Inmetro) ajudam a mitigar o problema da seleção adversa.

Além disso, pode ser que as empresas detenham poder de mercado e sejam capazes de "fazer os consumidores pagarem preços bem acima dos custos ou de oferecer produtos de

---

[99] É o caso, no Brasil, do Seguro Obrigatório de Danos Pessoais Causados por Veículos Automotores de Vias Terrestres, ou por sua Carga, a Pessoas Transportadas ou Não (Seguro DPVAT), criado pela Lei n.º 6.194, de 19 de dezembro de 1974 (BRASIL, 1974).

[100] Como o Certificado de Registro e Licenciamento de Veículo (CRLV), exigido no Brasil.

[101] Cf., nesse sentido, (MANKIW, 2013, p. 442–443).

[102] Reinier Kraakman, John Armour, Paul Davies e outros autores (2017), por exemplo, afirmam que uma das funções do direito societário é controlar os conflitos de agência que surgem entre (i) acionistas e administradores não sócios; (ii) sócios majoritários/controladores e sócios minoritários/não controladores; e (iii) a sociedade e as outras partes com quem ela celebra contratos, como credores, empregados e consumidores. Apenas para exemplificar, as regras previstas nos artigos 153 e seguintes da Lei n.º 6.404, de 15 de dezembro de 1976 (BRASIL, 1976), que estabelecem os deveres e responsabilidades do administrador de uma sociedade anônima, têm nítida função de reduzir o risco moral. Afinal, diante de tais regras, administradores têm claros incentivos para atuar de forma responsável.

qualidade medíocre" (TIROLE, 2020, p. 169). De fato, se não houver pressão competitiva, pode ser que os consumidores simplesmente tenham que se sujeitar às condições oferecidas pelo agente com poder de mercado. Por essa razão, desenvolvem-se leis de defesa da concorrência[103] e regulação setorial[104], a fim de coibir infrações contra a ordem econômica, em prol da manutenção da livre concorrência.

Ao lado da regulação setorial ligada à livre concorrência, em determinados mercados, "[o] Estado reconhece que [...] é indispensável uma solução regulatória mais forte, mais frequente, que garanta higidez (bom estado de saúde econômica), segurança e estabilidade" (SILVA, 2003, p. 173). Exemplos são os mercados bancário, aéreo e de seguros. São casos em que uma falência pode representar um custo social excessivo e indesejado e fazer com que a realização de negócios ultrapasse a capacidade do indivíduo. Afinal, se você coloca o seu dinheiro no banco ou contrata uma apólice de seguros, você pode passar necessidades financeiras se o banco ou a companhia de seguros tiver falido (TIROLE, 2020, p. 169). Adota-se, por essa razão, regras de regulação prudencial, voltadas às condições de acesso ao mercado e de exercício posterior da atividade econômica (SILVA, 2003, p. 173).[105]

Em outra perspectiva, em situações em que existem bens públicos, isto é, bens não excludentes[106] e não rivais[107], uma falha de mercado que comumente se origina é o incentivo para que as pessoas sejam caronas (*free riders*). "Um carona é uma pessoa que recebe o benefício de um bem, mas evita pagar por ele" (MANKIW, 2013, p. 206). Blocos de carnaval, por exemplo, são, em tese, não excludentes, já que acontecem em vias públicas, e não rivais – considerando, hipoteticamente, um cenário em que as vias têm espaço suficiente para comportar toda a população. Se um sujeito resolve organizar um bloco em Belo Horizonte, como as pessoas poderiam participar do evento sem pagar ingresso, seria difícil arrecadar valor suficiente para compensar os custos da organização. Diante da ausência de resultado eficiente

---

[103] A livre concorrência é um dos princípios que orientam a ordem econômica no Brasil, como previsto no inciso IV do art. 170 da Constituição (BRASIL, 1988). Do mesmo modo, o § 3º do art. 173, prevê que lei reprimirá o "abuso do poder econômico que vise à dominação dos mercados, à eliminação da concorrência e ao aumento arbitrário dos lucros". Essa função é exercida pela Lei n.º 12.529, de 30 de novembro de 2011 (BRASIL, 2011), que estrutura o Sistema Brasileiro de Defesa da Concorrência.

[104] Em termos concorrenciais, a regulação setorial é uma solução comum no caso de monopólios naturais, como empresas de água e energia elétrica (MANKIW, 2013, p. 301) e infraestruturas essenciais, cujo acesso é necessário para que haja concorrência. Por serem monopólios, deixar essas empresas atuarem livremente submeteria os consumidores a preços abusivos ou excluiria o acesso de concorrentes às infraestruturas essenciais. Por isso, existem agências reguladoras como, por exemplo, a Agência Nacional de Energia Elétrica (ANEEL), a Agência Nacional de Telecomunicações (ANATEL) e a Agência Nacional de Transportes Terrestres (ANTT).

[105] No Brasil, no caso dos setores citados, essas funções são desempenhadas pelo Banco Central do Brasil, pela Agência Nacional de Avião Civil (ANAC) e pela Superintendência de Seguros Privados (SUSEP).

[106] Se um bem é não excludente, as pessoas não podem impedir que outras pessoas o utilizem.

[107] Bens não rivais são bens cuja disponibilidade não é reduzida após a utilização por alguma pessoa. Desse modo, múltiplos agentes poderão utilizá-lo sem que o benefício proporcionado aos demais seja reduzido.

a partir do mercado, o Estado oferece patrocínios aos blocos de carnaval, a fim de garantir que eles sejam oferecidos[108].

Diferentemente de bens públicos, recursos comuns não são excludentes – todos podem utilizá-los –, mas são rivais, isto é, "o uso de um recurso comum por uma pessoa reduz a possibilidade que outras pessoas têm de usá-lo" (MANKIW, 2013, p. 210). Considere – como no clássico exemplo do ecologista Garrett Hardin (1968) – um pasto compartilhado por diversos pastores, que criam suas ovelhas. Enquanto a terra é abundante – e, portanto, não se trata de um bem rival –, não há problema algum; com o aumento da população e a diminuição do espaço comum, os pastores passam a ter incentivos para criar mais ovelhas, o que é replicado por seus pares e leva ao esgotamento da terra. Na conclusão pessimista de Hardin (1968, p. 1244), a "[r]uína é o destino para o qual todos os homens se dirigem, cada um perseguindo seus próprios interesses em uma sociedade que acredita na liberdade dos comuns"[109]. Diante de tal problema - a chamada Tragédia dos Comuns -, intervenções estatais podem ser exigidas; por exemplo, para definir um limite de ovelhas por família ou para simplesmente dividir as terras comuns em propriedades privadas (MANKIW, 2013, p. 211)[110].

Finalmente, pode ser que, embora produza eficiência, não existam incentivos no mercado para produzir equidade ou uma adequada distribuição de renda (TIROLE, 2020, p. 169-170). No primeiro caso, o mercado pode incentivar a adoção de comportamentos injustos que não são desejados pela sociedade. O exemplo de Tirole é ilustrativo: em contratos de seguro-saúde, se as seguradoras ou a previdência social pudessem discriminar de acordo com o estado de saúde atual ou com dados genéticos, um sujeito que tem câncer ou predisposição genética para ter problemas de saúde pode jamais conseguir um plano de saúde a um preço acessível. Já quanto à desigualdade, *a priori*, o mercado simplesmente não terá incentivos para produzir distribuição de renda. Em um primeiro aspecto, como pontua Tirole, a desigualdade pode ser vista como uma falta de seguro, já que, antes de nascermos, não conhecemos nosso

[108] A rigor, o efeito carona surge de uma externalidade positiva gerada pelo comportamento de um agente que investe em um bem público. Diante de tal externalidade, o valor social do bem é superior ao valor que o investidor auferiria, de modo que o mercado tende a produzir uma quantidade subótima do bem.

[109] Tradução livre. Texto original: "*Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons*".

[110] A questão, mais uma vez, decorre de externalidades. Se o rebanho de uma família pasta em terras comuns, reduz-se a qualidade de terra disponível para a outra família, o que lhe causa prejuízos (MANKIW, 2013, p. 210-211). Convém ressaltar, ainda, que a cientista política Elinor Ostrom (2015), laureada com o prêmio Nobel de economia em 2009, demonstrou que, em determinadas condições, pequenas comunidades conseguem gerir seus recursos comuns sem recair na ruína descrita por Hardin. Para isso, valem-se de mecanismos informais de incentivos, como a definição clara do conteúdo dos recursos comuns, exclusão de não membros, monitoramento por membros do grupo, sanções para indivíduos que violam as regras, mecanismos de resolução de conflitos baratos e de fácil acesso etc.

lugar no futuro – estamos "sob o véu da ignorância" – e gostaríamos de recompensar o esforço pessoal, para incentivar a criação de riqueza na sociedade, além de garantir que, se, por acaso, figurássemos entre as pessoas mais pobres, teríamos condições de vida dignas. Ao mesmo tempo, a desigualdade gera efeitos negativos por poder distorcer o laço social e criar externalidades para a população como um todo. Nesse caso, medidas por parte do Estado são exigidas para buscar a redução das desigualdades. Diante do crescimento galopante das desigualdades em vários países desenvolvidos nas últimas décadas[111], este é um tema que está na ordem do dia de boa parte dos governos.

Tabela 1 - Um panorama geral das tradicionais falhas de mercado[112]

Falha de Mercado	Exemplo	Possível Intervenção
Externalidade	Poluição	Política de proteção ao meio-ambiente
Risco moral (assimetria informacional)	Dificuldade para uma sociedade (principal) monitorar o comportamento de seu administrador (agente).	Normas cogentes de direito societário quanto ao comportamento do administrador
Seleção adversa (assimetria informacional)	Dificuldade para consumidores atestarem a real qualidade de um produto	Certificação de qualidade por órgãos independentes, como o Inmetro
Poder de mercado	Abuso de posição dominante por parte de um agente com alta participação de mercado	Política de defesa da concorrência
Proteção contra risco sistêmico	Potencial falência de uma instituição financeira	Regulação prudencial (condições de acesso e exercício da atividade econômica)
Carona	Organização de um bloco de carnaval em via pública (externalidade positiva)	Subsídio estatal à organização do bloco
Esgotamento de recursos comuns	Criação de animais em um pasto comum com tamanho e recursos limitados	Definição de limites de animais por famílias ou divisão das terras comuns
Comportamentos injustos	Cobrança de preços impraticáveis para o plano de saúde de um sujeito que tem câncer	Proibir o condicionamento de seguros-saúde a certos dados dos indivíduos
Desigualdade	Enorme diferença de renda entre a população mais rica e a população mais pobre de um país	Políticas de redistribuição de renda, prestação de serviços públicos universais, etc.

Fonte: Elaboração do Autor.

Diante dessas e de outras falhas de mercado, o Estado cria mecanismos para regular o comportamento dos agentes econômicos e minimizar as falhas de mercado. Por exemplo, o Estado faz com que eles introduzam as externalidades em suas funções de custo - ou de receita -

[111] Esse crescimento foi evidenciado de maneira notável na já clássica obra do economista francês Thomaz Piketty (2014). Além de também evidenciar o crescimento da desigualdade em níveis locais em países como os Estados Unidos e o Reino Unido, assim como fez Piketty, o economista sérvio Branko Milanovic (2017, 2020) demonstrou que, em níveis mundiais, a desigualdade tem reduzido, muito por conta de um processo de convergência entre as rendas ocidentais e orientais, movido não só pela estagnação e aumento das desigualdades nos Estados Unidos, como também pelo crescimento chinês e de outros países como a Índia e o Vietnã.
[112] Elaboração do autor.

ou tenham incentivos para reduzir assimetrias informacionais. Do mesmo modo, o Estado toma medidas ativas para coibir o exercício de poder de mercado por agentes incumbentes, regula ativamente mercados estratégicos e trabalha para reduzir as desigualdades.

2.2. Prejuízos aos próprios indivíduos em curto e longo prazo: mais motivos para regular o comportamento privado

Na abordagem das falhas de mercado tradicionais, expostas na última subseção, adota-se como parâmetro, como já ressaltado, a soberania do consumidor. Nesse modelo, os agentes econômicos – inspirados pelo arquétipo do *Homo economicus* – tomam decisões ponderadas e são os melhores árbitros para as suas próprias escolhas. Suas preferências, afinal, são reveladas pelas suas próprias escolhas, de modo que opções alternativas eventualmente desprezadas não seriam capazes de lhes trazer mais bem-estar. Intervenções, nesse caso, são bem-vindas quando constatadas falhas de mercado; caso contrário, o caminho seria não regular o comportamento privado de qualquer maneira.

Apesar disso, o alvorecer das ciências comportamentais e o mapeamento de inúmeros fatores supostamente irrelevantes que impactam a tomada de decisões apresentam um cenário mais complexo, que lança dúvidas sobre a teoria das preferências reveladas como parâmetro de bem-estar. Ao fim e ao cabo, as pesquisas das ciências comportamentais parecem apresentar novos fundamentos para a intervenção estatal na vida privada, ou, em outros termos, falhas de mercado comportamentais[113]. Esta subseção se dedica à exposição de tais fundamentos.

2.2.1. Os problemas do modelo de preferências reveladas e as falhas de mercado comportamentais

Como já mencionado, tradicionalmente, considera-se, em economia, que as preferências de um indivíduo, isto é, as opções que maximizam o seu bem-estar, são reveladas por suas próprias escolhas. Em outras palavras, se vemos um agente adquirir o disco Clube da Esquina, de Milton Nascimento e Lô Borges, em vez de adquirir o disco *Debut*, da islandesa Björk, presume-se que a opção escolhida é aquela que lhe traz maior bem-estar.

Os axiomas da teoria da preferência revelada foram elaborados pelo economista americano Paul A. Samuelson (1937, 1948). Segundo Samuelson (1948, p. 243), ao se comparar

[113] Vide: (SUNSTEIN, 2013a, 2014a).

os custos de diferentes combinações de bens em diferentes situações de preço relativo, é possível inferir quando determinado conjunto de bens seria preferido a outro. Amartya Sen (1973, p. 241) acrescenta que, nessa abordagem teórica, quando se observa alguém adquirir o grupo de bens *x*, quando *y* estava disponível e dentro de seus limites orçamentários, presume-se que esse indivíduo prefere *x* a *y*. Do ponto de vista da introspecção do decisor, o processo decisório começaria por suas preferências e, em seguida, a escolha seria feita. Por outro lado, do ponto de vista de um observador científico, o processo seria observado na direção contrária: primeiro, as escolhas seriam observadas, e, a seguir, as preferências seriam presumidas a partir de tais observações. A partir dessa ideia, tem-se que o bem-estar dos agentes econômicos poderia ser traduzido pela satisfação das preferências reveladas pelas escolhas. Daí a ideia já exposta de soberania dos consumidores[114].

Na perspectiva da economia comportamental, por outro lado, fatores supostamente irrelevantes influenciam a opção escolhida; ou seja, ainda que, em termos normativos, esses fatores não acrescentem nada em termos de bem-estar ao indivíduo, eles podem direcionar a opção que será efetivamente escolhida. Retomando o exemplo anterior, pode ser que o disco Clube da Esquina seja adquirido simplesmente por estar em posição de maior destaque no *Spotify*; o agente econômico, na verdade, preferiria adquirir o disco da Björk, mas, em razão de fatores supostamente irrelevantes acaba adquirindo o Clube da Esquina. Nessa mesma perspectiva, pode ser que as escolhas de tais indivíduos acabem sendo prejudiciais ao seu próprio bem-estar. Se, por exemplo, por excesso de otimismo, um indivíduo desembolsa uma quantia excessiva para jogar na loteria, pode ser que esta escolha não refletisse efetivamente a opção que maximizaria o seu bem-estar e, ao invés disso, simplesmente lhe cause prejuízo; há que se reiterar, como exposto no Capítulo 1, que inúmeras decisões são tomadas de maneira intuitiva, com base no Sistema 1.

Nesse contexto, começa-se a considerar que a preferência revelada pela escolha do indivíduo nem sempre revelará a opção que maximiza o seu bem-estar (THALER; SUNSTEIN, 2003a, p. 175). Quando se abandona a ideia de preferências reveladas, como fazem Thaler e Sunstein (2003a) ao considerar que as preferências são instáveis, transitórias e manipuláveis, é possível identificar mais situações em que as pessoas se comportam de maneira subótima e o

[114] Em termos técnicos, como exposto por Dhami (2016, p. 1580), indivíduos derivam a utilidade de ações e de estados psicológicos. Há uma função de *feedback* que liga ações a estados psicológicos. Os decisores clássicos, inspirados no arquétipo do *Homo economicus*, sem vieses ou incompreensões, levam o *feedback* em consideração para a tomada de decisões subsequente. Decisores comportamentais, no entanto, ignoram o *feedback*, e, por essa razão, as escolhas reveladas são diferentes das escolhas normativas, o que leva a um nível mais baixo de bem-estar.

mercado é incapaz de mudar suas decisões por conta própria. Em outros termos, é possível descrever novas falhas de mercado, desta vez de caráter quase puramente comportamental.

Em linha com o exposto por Sunstein (2013a, 2014a), serão apresentados quatro grupos diferentes de falhas de mercado comportamentais. A ideia é retomar alguns dos fatores supostamente irrelevantes expostos no Capítulo 1, para demonstrar como eles servem de fundamento para a regulação do comportamento privado. Evidentemente, mais uma vez, não se pretende esgotar o assunto, mas, sobretudo, indicar exemplos que permitam a adequada compreensão de como a descrição de problemas na tomada de decisões leva a novos fundamentos para a criação de normas para regular o comportamento privado.

Inicialmente, uma primeira falha de mercado comportamental se relacionada ao viés para o presente, à inconsistência temporal e a problemas de autocontrole.

Em escolhas intertemporais, os custos e as consequências de uma determinada opção se diluem no tempo. Na perspectiva da economia tradicional, as pessoas descontariam o futuro em uma taxa constante e maximizariam a utilidade descontada de suas escolhas.[115] Se o consumo imediato de um determinado bem no presente gera uma utilidade de 100 *utils*[116] ao indivíduo, com uma taxa constante de desconto de, por exemplo, 10% ao ano, o consumo do mesmo bem geraria 90 *utils* para o indivíduo no ano subsequente, 81 *utils* em dois anos, e assim sucessivamente (THALER, 2015, p. 92).

Na realidade, porém, a maioria das escolhas intertemporais apresenta, ao mesmo tempo, resultados imediatos concretos e tangíveis e consequências futuras intangíveis ou indefinidas (RICK; LOEWENSTEIN, 2015, p. 78). Quando se consome compulsivamente alimentos gordurosos, por exemplo, o prazer do consumo imediato é concreto, enquanto eventuais problemas de saúde no futuro são incertos e intangíveis quando a escolha é realizada. O mesmo raciocínio pode ser utilizado para o consumo de drogas ilícitas e para hábitos tabagistas. Nesse contexto, é improvável que o decisor desconte o futuro em taxas constantes e consiga maximizar a utilidade descontada. Inúmeros experimentos vêm demonstrando, inclusive, um comportamento notadamente marcado por um viés para o presente, em que quaisquer opções

---

[115] O modelo tradicional econômico que descreve o comportamento de um agente racional diante de escolhas intertemporais foi elaborado por Samuelson (1937) e foi denominado modelo da utilidade descontada constante. De acordo com essa proposta, o consumo atual teria mais valor para o agente do que o consumo futuro, mas a utilidade do consumo futuro seria descontada em um percentual constante, em razão do tempo de espera que deveria ser superado pelo agente. Nessa perspectiva, por serem constantes a depender do tempo de espera, os descontos seriam descritos por uma função exponencial e haveria, então, "trade-offs explícitos entre os custos e os benefícios que ocorrem em momentos distintos" (RICK; LOEWENSTEIN, 2015, p. 76). Nesse caso, com taxas de desconto sempre constantes, os agentes maximizariam a utilidade descontada de suas opções e descontariam o futuro de acordo com uma função exponencial (THALER, 2015, p. 90).

[116] *Utils* são medidas arbitrárias utilizadas por economistas para descrever níveis de utilidade ou felicidade (THALER, 2015, p. 90).

futuras são mais custosas, em comparação às opções do presente, e por inconsistência intertemporal, já que há reversão de preferências de modo que as preferências de um indivíduo num determinado tempo *t* são diferentes de suas preferências em momentos futuros.[117,118] Favorecem ainda mais esse cenário a importância do *status quo* à definição das escolhas que realmente são feitas pelos agentes[119], assim como um comportamento marcado pela procrastinação[120].

Todos esses problemas de autocontrole e erros de julgamento derivados de decisões indevidamente míopes podem ser considerados internalidades, já que as pessoas tomam decisões que causam danos aos seus *selves* futuros (SUNSTEIN, 2014a, p. 37-38). Embora esse seja um claro cenário de decisões ineficientes para o próprio indivíduo, pode ser que não existam incentivos para que os agentes do mercado tentem contrapor esses fatores supostamente irrelevantes que levam ao viés para o presente. Nesse cenário, em que os sujeitos prejudicam o seu próprio futuro e, por vezes, até mesmo gostariam de agir de forma diversa, uma intervenção estatal pode ser recomendada para corrigir esse rumo de ação. Se o consumo de cigarro é regulado, em termos tradicionais, por assimetrias de informações e potenciais externalidades, vieses para o presente apresentam motivos possivelmente até mesmo mais fortes para uma regulação do comportamento privado.

Por outro lado, outra falha de mercado comportamental se relaciona aos atributos ocultos (*shrouded attributes*) que impactam a tomada de decisões. Atributos ocultos são características que não são diretamente observáveis em bens adquiridos por consumidores

---

[117] No clássico modelo de David Laibson (1997), os descontos intertemporais seriam descritos por funções hiperbólicas, em que há altas taxas de desconto em curtos horizontes temporais e taxas de desconto menores quando períodos distantes são postos em evidências. Há, nesse cenário, um conflito entre as preferências atuais e as preferências que serão mantidas no futuro. Jess Benhabib, Alberto Bisin e Andrew Schotter (2010), por outro lado, embora não tenham identificado evidências de um desconto hiperbólico ou quase hiperbólico, identificaram um viés para o presente na forma de um custo fixo atrelado a decisões futuras, em comparação às decisões do presente.

[118] Como explicam Colin Camerer e George Loewenstein (2004, p. 23), na perspectiva comportamental, prevê-se que os agentes farão escolhas prudentes quando planejam com antecedência – custos e benefícios futuros –, mas, quando alguns custos e benefícios são imediatos, esses mesmos agentes tendem a apresentar uma miopia para o presente. O indivíduo, então, faz previsões de que, no futuro, atuará de maneira comedida, mas como apresenta inconsistências temporais, atua de maneira contrária aos seus desejos prévios e busca gratificações imediatas quando o futuro chega.

[119] William Samuelson e Richard Zeckhauser (1988), por exemplo, identificaram uma série de experimentos demonstrando que os agentes permanecem com a opção do *status quo* de maneira desproporcional. Kahneman, Jack L. Knetsch e Thaler (1991), no mesmo sentido, afirmam que uma das implicações da aversão à perda é uma forte tendência de permanecer no *status quo*, já que as desvantagens de mudar de opção têm mais peso do que as vantagens correspondentes.

[120] Ted O'Donoghue e Matthew Rabin (2001), por exemplo, elaboraram um modelo formal de escolha intertemporal e identificaram que: a) fornecer opções adicionais pode induzir a procrastinação, já que, se o sujeito tiver suficientes benefícios em longo prazo com a opção nova, é possível que ele opte por começar a seguir essa nova opção; e que b) as pessoas podem procrastinar mais na busca por objetivos importantes do que em escolhas insignificantes, já que os planos mais importantes tendem a ser mais ambiciosos e exigir um maior esforço.

(DHAMI, 2016, p. 1431). Na abordagem econômica tradicional, considera-se que os custos para tornar claros os atributos ocultos são tão pequenos que se presume que tais atributos são visíveis aos consumidores, ou que, em outra perspectiva, a concorrência entre empresas faz com que elas tornem esses atributos claros (DHAMI, 2016, p. 1431). Afinal, consumidores racionais considerariam atributos ocultos um mau sinal e tenderiam a fazer negócios com outras empresas, de modo que todos os agentes do mercado tentariam revelar tais atributos.

A realidade, porém, não apresenta esses incentivos. No clássico exemplo de Daniel J. Simons e Christopher F. Chabris (1999)[121], os participantes do experimento contam o número de vezes em que uma bola de basquete é passada entre os jogadores do time de branco; ao mesmo tempo, um sujeito fantasiado de gorila passa pela imagem e não é identificado por boa parte dos participantes que estão contando o número de passes. Do mesmo modo, os agentes frequentemente não têm uma percepção adequada de atributos ocultos e são menos responsivos a alterações nesses atributos, em comparação com mudanças em atributos visíveis.[122]

A consequência é que as empresas terão incentivos para manter tais elementos ocultos. Pense, por exemplo, na economia de combustível proporcionada por um determinado veículo; em geral, pode não haver incentivos para que a montadora deixe esse elemento claro na arquitetura de escolhas dos consumidores; afinal, esse atributo oculto pode não ser considerado pelos consumidores. Diante da ausência de uma solução de mercado, pode ser necessária a regulação estatal, para exigir, por exemplo, que as empresas sejam transparentes e deixem *salientes* atributos que poderiam permanecer ocultos sem a presença da intervenção[123].

Além disso, outra falha de mercado comportamental decorre do comportamento excessivamente otimista que as pessoas apresentam. Um agente racional, a partir de dados sobre os riscos de determinados eventos e de sua própria experiência positiva e negativa ao longo da vida, conseguiria atualizar suas crenças e ter expectativas mais ou menos corretas. Na realidade, porém, as pessoas acreditam que têm menor probabilidade de sofrer infortúnios, como ser demitido, diagnosticado com câncer ou sofrer de problemas com álcool, do que a média

---

[121] Recomenda-se que o vídeo indicado no *link* a seguir seja visualizado, antes que se continue a ler este parágrafo: https://www.youtube.com/watch?v=IGQmdoK_ZfY. Acesso em: 25 fev. 2022.

[122] Nesse sentido, Tanjim Hossain e John Morgan (2006) elaboraram um experimento em que pares de CDs e jogos de Xbox eram leiloados no eBay; o objetivo era avaliar se o preço final dos leilões pelos produtos seria inversamente proporcional ao valor cobrado a título de frete (*atributo oculto*). Os autores verificaram que cobrar preços mais altos de frete e atribuir um baixo preço inicial ao produto levava a mais participantes nos leilões e maior receita para a empresa, desde que o preço do frete não fosse excessivo.

[123] A Resolução n.º 3.919, de 25 de novembro de 2010, do Banco Central do Brasil (2010), por exemplo, determina, em seu art. 13, que faturas de cartão de crédito explicitem informações a respeito de elementos como o limite de crédito, gastos realizados com o cartão, o valor dos encargos a serem cobrados etc. Sem essa determinação, instituições financeiras teriam incentivos para manter alguns desses elementos ocultos.

(SHAROT, 2011)[124]. Essa excessiva confiança é uma das implicações do viés de confirmação - a tendência exibida pelos agentes de interpretarem evidências futuras de maneira que reforce as suas crenças iniciais (DHAMI, 2016, p. 1389)[125,126].

O problema do otimismo excessivo é que ele leva as pessoas a correrem riscos desnecessários e a tomarem decisões que lhes são prejudiciais. O exemplo de Tali Sharot (2011) é ilustrativo. Embora tenha sido alertado pelo espião soviético Leopold Trepper de que a Alemanha nazista preparava uma invasão ao território soviético, Joseph Stalin instruiu que o referido espião fosse punido por contar mentiras; pouco tempo depois, a União Soviética foi realmente invadida por 4,5 milhões de tropas. Excessivamente otimista, Stalin não acreditava que os alemães poderiam atacar seu país. Do mesmo modo, pessoas que subestimam as chances de se endividarem podem se comportar de maneira imprudente. O mercado, nesse cenário, mais do que buscar meios de estimular um comportamento realista, pode acabar explorando o otimismo excessivo. Bancos podem, por exemplo, conceder um limite de crédito alto o suficiente para que indivíduos se endividem, resultado cuja probabilidade é subestimada por tais agentes. Pode ser, então, que o viés do otimismo seja mais um fundamento para que o Estado elabore medidas regulatórias[127].

Por fim, problemas com raciocínio estatístico também podem ser falhas de mercado comportamentais. Como amplamente demonstrado e discutido no Capítulo 1, os agentes frequentemente deixam de seguir as normas que seriam recomendadas pela teoria da probabilidade. Ignoram, por exemplo, a taxa-base ao estimar a probabilidade de um evento

---

[124] Neil D. Weinstein (1980) cunhou a expressão *viés do otimismo* (*optimism bias*) e demonstrou experimentalmente que estudantes tendiam a acreditar ter maior probabilidade do que seus pares de ter experiências positivas e menor probabilidade de passar por eventos negativos.

[125] Rabin e Joel Schrag (1999) apresentaram três conclusões a respeito do viés de confirmação, que poderiam ser extraídas da literatura da psicologia: a) o viés de confirmação e a confiança excessiva surgem quando as pessoas têm que interpretar evidências ambíguas; b) o viés de confirmação pode surgir quando as pessoas são obrigadas a interpretar evidências estatísticas para avaliar a correlação entre fenômenos separados no tempo; c) o viés de confirmação ocorre quando as pessoas coletam ou avaliam evidências de maneira seletiva. Seu modelo demonstrou, além disso, que o viés de confirmação leva à confiança excessiva.

[126] O modelo social-intuicionista de julgamentos elaborado por Haidt (2001, 2012, p. 55), em outra perspectiva, também auxilia a compreender o viés de confirmação. Nesse modelo, as intuições a respeito de um evento surgem primeiro e o raciocínio é usualmente feito após a realização de um julgamento, a fim de influenciar outras pessoas. Como o julgamento é feito prévia e intuitivamente, o raciocínio busca, em geral, reforçar essas intuições, e, por isso, as evidências tendem a ser interpretadas de maneira enviesada. A mudança dos julgamentos intuitivos pode até surgir a partir do raciocínio e da reflexão privada, mas esse é um evento mais raro; para Haidt, há maior probabilidade de mudança nos julgamentos intuitivos a partir da influência social e do contato com intuições de terceiros. Daí o modelo ser denominado *social-intuicionista*.

[127] A Resolução n.º 4.549, de 26 de janeiro de 2017, do Banco Central do Brasil (2017), por exemplo, proibiu a permanência em uma linha de crédito rotativo por mais do que um mês, já que essa linha emergencial de crédito acabava por ser excessivamente utilizada por indivíduos *excessivamente otimistas*, o que levava a um cenário de superendividamento. A respeito de tal alteração regulatória, Cf.: (SILVA *et al.*, 2019).

acontecer; ou, por outro lado, superestimam a probabilidade de que um evento vividamente marcante aconteça em um determinado período.

Esses fatores supostamente irrelevantes podem levar a erros que podem causar prejuízos. Pode ser, por exemplo, que a probabilidade de que ocorra um desastre ambiental seja subestimada e os agentes tomem menos cuidados com a preservação do meio ambiente.[128] Por outro lado, se o sujeito superestima a probabilidade de sofrer um acidente de avião, pode ser que ele tome medidas inadequadas, como a contratação de seguros a preços exorbitantes ou, até mesmo, que deixe de utilizar tal meio de transporte. A partir de tais constatações, é possível pensar em medidas regulatórias que tentem aprimorar esse comportamento e fazer com que as pessoas decidam de maneira minimamente informada.

Tabela 2 - Um panorama geral das falhas de mercado comportamentais

Falha de Mercado	**Exemplo**	**Possível Intervenção**
Internalidades	Consumo excessivo de cigarros	Elevação dos tributos do cigarro; avisos chamativos a respeito dos riscos associados ao uso; proibição de propagandas.
Ignorância de atributos ocultos importantes	Baixa responsividade ao aumento de encargos bancários em cartões de crédito	Exigência regulatória de que os encargos sejam explicitados separadamente na fatura.
Otimismo excessivo	Aceitação de limites de crédito superiores à capacidade patrimonial do agente e superendividamento	Limitação temporal na utilização da linha de crédito rotativo; divulgação mais *saliente* da evolução de eventuais dívidas.
Vieses na avaliação de probabilidades	Contratação de seguros excessivamente caros, para cobertura de eventos raros	Exigência de divulgação clara e saliente dos riscos envolvidos no evento.

Fonte: Elaboração do Autor.

[128] Como mencionam Christine Jolls, Sunstein e Thaler (1998, p. 1519), a heurística da disponibilidade pode levar à regulação insuficiente e à regulação excessiva ao mesmo tempo. As pessoas podem subestimar a probabilidade de eventos raros ou pouco salientes, simplesmente porque essas ameaças não parecem estar nos "*radares*" das pessoas. Riscos ambientais e sanitários, como problemas de saúde originados de dietas inadequadas e da falta de exercício físico, podem se enquadrar nessa descrição para parte da população. No entanto, se um determinado evento se torna disponível, com um risco recentemente materializado, a regulação passa a ser demandada. A aversão à utilização de energia nuclear, ainda que ela possa ser menos prejudicial do que outras matrizes energéticas, por exemplo, pode eventualmente ser explicada pela disponibilidade de grandes acidentes como aqueles ocorridos em Chernobyl, na Ucrânia, e em Fukushima, no Japão.

Exposto esse conjunto de falhas de mercado comportamentais - que, repita-se, não é exaustivo nem pretende esgotar o tema -, um último questionamento é válido para encerrar este item. Por que os fatores supostamente irrelevantes expostos acima são denominados falhas de mercado?

A resposta passa pela análise do equilíbrio de mercado diante de tais vieses comportamentais. Sem intervenções, em equilíbrio, há uma série de interesses econômicos alinhados contra os consumidores, na medida em que há incentivos para que as empresas explorem os fatores supostamente irrelevantes, a fim de maximizar sua receita. Mais do que isso, se uma empresa deixar de explorar esses *vieses* e os seus concorrentes o fizerem, pode ser que ela simplesmente perca participação de mercado (LOEWENSTEIN; HAISLEY, 2008, p. 234). Como George Akerlof e Robert Shiller (2015) exploram extensamente, o mercado fornece incentivos à pescaria de tolos (*phishing for phools*), isto é, à exploração de fatores supostamente irrelevantes com o objetivo de auferir lucros. No equilíbrio da pescaria (*phishing equilibria*), os agentes do mercado vão tirar proveito de tais aspectos comportamentais[129].

Não por outra razão, empresas exploram o consumo de cigarro e de álcool ou bancos podem estimular um consumo desenfreado, muitas vezes disponibilizando limites de crédito superiores às capacidades patrimoniais dos clientes. Embora existam forças econômicas que favorecem os consumidores - afinal, empresas lucram, por exemplo, com dietas, adesivos de nicotina e o acúmulo de ativos financeiros -, as forças no sentido contrário são muito mais fortes (LOEWENSTEIN; HAISLEY, 2008, p. 234). De fato, consumidores que, por exemplo, precisam de adesivos de nicotina são acometidos pelo vício em cigarro em momento prévio; ou seja, o produtor de tais adesivos tem incentivos ambivalentes: uma população dependente de cigarros é necessária para que ele possa vender seu produto, o que gera um interesse em promover o consumo de cigarros em um primeiro momento.

---

[129] As conclusões de Akerlof e Shiller (2015) estão diretamente relacionada às falhas de mercado comportamentais, como o trecho que inicia o livro evidencia: "*The psychologists have taught us over the course of more than a century [...] that people frequently make decisions that are not in their best interest. [...] Such bad decisions make it possible for them to be phished for phools. [...] The fundamental concept of economics is quite different: it is the notion of market equilibrium. [...] As businesspeople choose what line of business to undertake [...] they [...] pick off the best opportunities. This too creates an equilibrium. Any opportunities for unusual profits are quickly taken off table, leading to a situation where such opportunities are hard to find. [...] The principle also applies to phishing for phools. That means that if we have some weakness or other – some way in which we can be phished for phools for more than the usual profit – in the phishing equilibrium someone will take advantage of it. Among all those business persons figuratively arriving at the checkout counter, looking around, and deciding where to spend their investment dollars, some will look to see if there are unusual profits from phishing us for phools. And if they see such an opportunity for profits, that will (again figuratively) be the "checkout lane" they choose. And economics will have a phishing equilibrium in which every chance for profit more than ordinary will be taken up*" (AKERLOF; SHILLER, 2015, p. 1-2)

Em todas essas situações, seria possível se atingir um resultado melhor para o próprio indivíduo, se ele conseguisse agir de outro modo; pode ser, inclusive, que se atingissem resultados mais eficientes, à sociedade como um todo. Na prática, no entanto, agir de outro modo nem sempre é possível sem uma intervenção estatal. O que se constata, então, é que o alvorecer das ciências comportamentais, além de possibilitar o reconhecimento, por métodos de economia positiva, das limitações do agente na busca pelo seu autointeresse – retomando ideais clássicos já expostos, por exemplo, por Adam Smith –, deu lugar a uma nova área de pesquisas em economia normativa, com enfoque em dois elementos: como medir adequadamente o bem-estar de agentes com preferências inconsistentes e como elaborar arranjos econômicos e sociais para maximizar esse bem-estar (LOEWENSTEIN; HAISLEY, 2008, p. 238).

2.2.2. Existem falhas de mercado que não são comportamentais?[130]

A exposição de falhas de mercado tradicionais e falhas de mercado comportamentais pode suscitar facilmente o seguinte questionamento: até que ponto existem falhas de mercado que não são comportamentais?

Considere, por exemplo, a vacinação. Como exposto na Subseção 2.1.1, a vacinação é um clássico caso de externalidade positiva. Os benefícios sociais gerados pela vacinação individual são superiores aos benefícios individuais. Ao mesmo tempo, o sujeito pode ser excessivamente otimista e subestimar os riscos de contrair uma doença, sobretudo diante de uma doença pouco usual ou já quase erradicada – afinal, as consequências de tal doença não estarão facilmente disponíveis ao indivíduo. Nesse cenário, tanto a partir de externalidades quanto do viés do otimismo e da heurística da disponibilidade, é possível inferir que haverá um consumo subótimo de vacinas, o que justifica uma intervenção estatal.

A mesma afirmação vale para outros exemplos de falhas de mercado tradicionais. A tragédia dos comuns, por exemplo, pode ser explicada por problemas de coordenação que decorrem de externalidades geradas pela natureza dos recursos comuns – rivais, mas não excludentes. Esse mesmo evento, contudo, pode ser explicado por internalidades: enviesados para o presente, os agentes acabam priorizando benefícios atuais em detrimento de benefícios maiores no futuro. A consequência, nesse caso, poderá ser o esgotamento dos recursos comuns, o que, mais uma vez, exige a regulação do comportamento privado.

[130] As provocações feitas pela Prof. Dra. Rute Saraiva originaram a discussão apresentada neste tópico.

Qual é, então, a distinção entre falhas de mercado tradicionais e falhas de mercado comportamentais?

Em um primeiro aspecto, falhas de mercado tradicionais são explicações para cenários em que o mercado, por si só, produzirá ineficiências; a solução ótima não passa pelo mercado. Todos esses problemas existem em modelos de tomada de decisão baseados no *Homo economicus*; isto é, falhas de mercado tradicionais surgem ainda que existisse um agente racional maximizador tomando decisões sem a influência de fatores supostamente irrelevantes. São problemas inerentemente econômicos. Em outras palavras, mesmo no paradigma da economia neoclássica, as intervenções governamentais seriam bem-recebidas diante de tais falhas de mercado.

Falhas de mercado comportamentais, por outro lado, não existiriam em modelos de tomada de decisão racional, baseados em agentes racionais maximizadores. Na verdade, a própria constatação de tais falhas de mercado decorre da incorporação de *insights* comportamentais ao pensamento econômico. Embora o *Homo economicus* não tenha problemas de autocontrole e consiga maximizar a utilidade descontada de suas preferências, o *Homo sapiens* acaba apresentando esse comportamento. Os problemas, aqui, são comportamentais. De todo modo, as falhas de mercado comportamentais também podem gerar situações subótimas do ponto de vista social e podem, ainda, criar prejuízos aos próprios agentes decisores.

Ressalvadas essas distinções, na perspectiva adotada nesta dissertação, a divisão entre falhas de mercado tradicionais e comportamentais é artificial e meramente didática.

A abordagem da economia comportamental - que compõe o marco teórico desta pesquisa - não rejeita a noção de que incentivos importam nem de que indivíduos *podem* agir de forma a maximizar o seu bem-estar em algumas situações[131]. A economia comportamental simplesmente expande a teoria econômica, de maneira similar a outras extensões que já foram experimentadas em modelos tradicionais da economia no passado. Como expõem Colin Camerer e outros autores (2003, p. 1218), os modelos econômicos mais simples presumem que há concorrência perfeita, informação perfeita e racionalidade perfeita. São modelos obviamente reducionistas e distantes da realidade, que permitem, porém, a obtenção de insights úteis. De todo modo, o relaxamento gradual de tais premissas estritas tem se mostrado produtivo. A

[131] Thaler e Sunstein (2021), por exemplo, admitem esse ponto expressamente e afirmam que, embora apresentem fatores negligenciados pela teoria econômica tradicional, eles não pretendem sugerir que as forças usuais da economia não são importantes. Eles acreditam em oferta e demanda e ressaltam que incentivos devem ser levados em consideração na elaboração de políticas informadas por insights comportamentais.

concorrência perfeita entre agentes e empresas, por exemplo, começou a ser flexibilizada a partir dos anos 1930, o que permitiu o desenvolvimento de pesquisas inovadoras em organização industrial. A partir dos anos 1970, por outro lado, a premissa de informação perfeita foi flexibilizada em modelos de teoria dos jogos e interações estratégicas com informação imperfeita que obtiveram enorme sucesso.

Nesse contexto, flexibilizar as premissas que seriam exigidas por uma visão de perfeita racionalidade, como faz a economia comportamental, parece ser apenas um próximo passo lógico na progressão apresentada. Afinal, utilizar um modelo de perfeita racionalidade envolve um deslocamento tão grande da realidade que as previsões decorrentes do modelo podem acabar sendo imprecisas. De todo modo, isso não quer dizer que o paradigma comportamental deve desconsiderar problemas puramente econômicos ou simplesmente jogar fora tudo o que se produziu em economia até hoje.

Na prática, para inúmeros problemas, as razões para se intervir podem ser apresentadas pela economia tradicional, pela economia comportamental, ou por uma análise híbrida entre os dois paradigmas (LOEWENSTEIN; CHATER, 2017, p. 29). Intervenções para combater o consumo de cigarro, por exemplo, podem ser explicadas por: a) razões tradicionais como externalidades, com relação aos fumantes passivos e a eventuais custos à saúde pública, e assimetrias informacionais, já que os consumidores não têm informações suficientes para avaliar os riscos do consumo de cigarro; b) razões comportamentais, ligadas a internalidades, já que os custos da ação estão no futuro, enquanto os benefícios estão no presente, o que gera um comportamento enviesado para o presente, e ao viés do otimismo; e c) razões *híbridas*, uma vez que as empresas podem maximizar a sua utilidade ao explorar vieses comportamentais dos consumidores. Nesse cenário, como ponderam Loewenstein e Nick Chater (LOEWENSTEIN; CHATER, 2017, p. 48), o desenvolvimento da economia comportamental deve ser visto como suplementar aos métodos tradicionais de análise da economia e de políticas públicas. *Políticas* efetivas exigem que os problemas sejam enfrentados com uma combinação *fluída* e *flexível* de insights comportamentais e econômicos.

Não por outra razão, Tirole (2020) não separa falhas de mercado tradicionais e falhas de mercado comportamentais ao discutir o arcabouço institucional da economia e apresentar as razões que levam o Estado a regular o comportamento privado. De fato, em meio às falhas de mercado, Tirole (2020, p. 168) menciona que "[é] possível que o comprador seja sua própria vítima, careça de autocontrole e adote um comportamento impulsivo" e, nesse caso, podem ser necessárias medidas para "proteger os consumidores contra si mesmos". Em outras palavras, Tirole reconhece a possibilidade de simplesmente mencionarmos falhas de mercado como

razões econômicas ou comportamentais para justificar a regulação do comportamento privado. A separação, como mencionado, é meramente didática.

2.3. Diferentes espécies de paternalismos comportamentais: entre paternalismos suaves e o paternalismo coercitivo

A identificação de falhas de mercado comportamentais mostra que nem sempre as escolhas dos agentes refletem as suas verdadeiras preferências. Diante de prejuízos ao próprio bem-estar em curto e, especialmente, em longo prazo, pode ser que os próprios agentes demandassem intervenções estatais para corrigir ou aprimorar suas escolhas. Em outras palavras, suas escolhas podem ser erros de aplicação, tal como exposto no Capítulo 1[132]. A partir dessa constatação, como mencionado no início deste Capítulo, diversos autores passaram a defender intervenções estatais em situações nas quais nenhuma intervenção era sugerida no arcabouço tradicional da economia neoclássica. Expor essas propostas é o objetivo desta seção.

Na abordagem das tradicionais falhas de mercado, expostas na Seção 2.1 acima, há pouco espaço para medidas paternalistas. Em geral, as intervenções que corrigem falhas de mercado almejam simplesmente obter mais eficiência e um melhor resultado em termos sociais. Quase todas elas surgem a partir da constatação de alguma externalidade, de modo que as intervenções costumam estar alinhadas ao princípio do dano elaborado por Mill. Em outros termos, o fundamento para se intervir é evitar o dano provocado a um terceiro – seja ele um dano direto, originado da produção ou do consumo excessivo de bens que geram prejuízo para terceiros, sem que esse prejuízo seja internalizado pelo produtor ou pelo consumidor, seja um dano indireto gerado pela produção ou pelo consumo subótimo de um bem que gera externalidades positivas.

Em alguns casos, no entanto, existem intervenções propostas a partir das tradicionais falhas de mercado que têm nítido caráter paternalista. A obrigação de utilizar cinto de segurança em veículos ou capacetes em motocicletas, por exemplo, parece se pautar na falta de adequada internalização dos riscos originados de condutas como dirigir sem cinto de segurança ou pilotar uma motocicleta sem capacete (DHAMI, 2016, p. 1579), ainda que existam potenciais externalidades derivadas de tais condutas[133]. As consequências de tais ações podem ser

[132] Como mencionado no Capítulo 1, em erros de aplicação, há evidências de que as pessoas conhecem e aceitam a regra que deixam de aplicar (KAHNEMAN; TVERSKY, 1982, p. 125). Cf., ainda, a nota 87 *supra*.

[133] Afinal, acidentes de passageiros sem cintos de segurança ou de motociclistas sem capacete podem ser mais graves e gerar maiores problemas para a saúde dos envolvidos. No limite, se a maior parte das pessoas andassem

definitivas, e, nesse caso, assimetrias de informação podem gerar prejuízos duradouros. De todo modo, como mencionado, à exceção de tais hipóteses, as falhas de mercado tradicionais levam a poucas intervenções paternalistas.

As falhas de mercado comportamentais, por outro lado, em geral conduzem a políticas paternalistas. Grosso modo, essas políticas extrapolam o princípio do dano de Mill e não necessariamente respeitam a soberania do consumidor. Em outras palavras, haveria fundamentos para a regulação do comportamento privado, ainda que esse comportamento não gerasse danos a terceiros e se originasse de uma escolha livre e consentida do próprio agente.

2.3.1. Paternalismo: conceito e distinções importantes

Uma das ideias prévias essenciais ao desenvolvimento desta dissertação é o conceito de paternalismo. Afinal, para apreciar a legitimidade normativa de intervenções paternalistas que buscam coibir falhas de mercado comportamentais, faz-se necessário, em primeiro lugar definir o que se entende por paternalismo.

De maneira geral, dizer que uma medida é paternalista significa dizer que ela envolve algum tipo de limitação na liberdade ou na autonomia dos agentes e que isso é feito por um conjunto particular de motivos (DWORKIN, 2020). Não há consenso, porém, a respeito dos exatos contornos do conceito, isto é, das condições necessárias para que se tenha paternalismo no caso concreto.

Como Gerald Dworkin (2020) menciona, em algumas acepções, o termo é utilizado de maneira avaliativa, de modo que caracterizar uma política como paternalista é condená-la ou criticá-la. Joel Feinberg (1986, p. 4-5) afirma, por seu turno, que, por duas razões, é um termo infeliz. A primeira razão, de maneira semelhante a Dworkin, é o caráter depreciativo do termo, que tende a ser utilizado de maneira tendenciosa, como um verdadeiro rótulo. Em outra perspectiva, o paternalismo enquanto princípio que legitima intervenções com o objetivo de proteger as pessoas de si mesmas pode ser facilmente confundido com práticas que não são nada benevolentes. O exemplo de Feinberg é ilustrativo: o gerente de uma fábrica que institui uma regra para obrigar os seus empregados a pedir permissão aos superiores para ir ao banheiro pode ser acusado de agir de forma paternalista; o fundamento de sua intervenção, no entanto, não é melhorar o bem-estar dos empregados, mas a pretensão de melhorar a eficiência da fábrica

---

de carro sem cinto de segurança ou pilotasse motocicletas sem capacete, acidentes teriam consequências mais graves e poderiam sobrecarregar o sistema de saúde, o que, por sua vez, geraria danos sociais não internalizados pelos envolvidos nos acidentes.

e obter maiores receitas. Em outros termos, esse tipo de prática não deveria ser vista como paternalismo.

Feitas essas ressalvas, parece ser preferível, em termos de metodologia, que se busque um conceito de paternalismo que seja definido em termos *não normativos*, sem que juízos de valores prévios façam parte das condições que o definem (DWORKIN, 2020). Nesse caso, em uma definição mais clássica de paternalismo, Dworkin (2020) sugere que as seguintes condições sejam levadas em consideração para que se analise se X age paternalisticamente com relação a Y ao fazer (ou deixar de fazer) um determinado ato Z:

a) Z (ou sua omissão) representa uma interferência na liberdade ou na autonomia de Y;
b) X age sem o consentimento de Y;
c) X só pratica o ato Z porque acredita que ele irá melhorar o bem-estar de Y (ou evitar que o seu bem-estar seja reduzido), ou de alguma maneira promoverá os seus interesses, valores ou bens.

Nessa perspectiva, seriam paternalistas políticas como, por exemplo, exigir que as pessoas contribuam para um sistema de seguridade social, exigir que motociclistas utilizem capacetes, proibir as pessoas de nadarem em uma praia sem que salva-vidas estejam presentes, proibir a venda de medicamentos que não sejam efetivos e proibir a venda de drogas nocivas à saúde. O próprio Dworkin (2020) reconhece, no entanto, que definir as condições necessárias para que uma política seja paternalista é o objeto de um julgamento filosófico. O termo "paternalismo", tal como é utilizado em contextos cotidianos, pode ser muito amorfo para servir como base para pensar em problemas normativos em particular. Assim, para Dworkin, é necessário fazer tal escolha a partir de uma análise baseada na hipótese que seja mais útil para pensar a respeito de um conjunto particular de problemas normativos.

Como o objetivo desta dissertação é avaliar a legitimidade de *nudges*, ainda que com fundamentos puramente paternalistas, avaliar os conceitos de paternalismo adotados pelos proponentes de tais medidas parece ser um caminho adequado a ser seguido. Embora reconheça a diversidade de acepções que o termo paternalismo pode ter, Sunstein (2013a, p. 1854, 2014a, p. 54) afirma que o aspecto que unifica as abordagens paternalistas é o governo não acreditar que as escolhas das pessoas promoverão seu próprio bem-estar e, por essa razão, tomar medidas para influenciar ou alterar essas escolhas para o próprio bem dessas pessoas. De maneira semelhante, na recente edição final do livro *Nudge*, Thaler e Sunstein (2021, p. 312), depois de afirmar que evitariam ficar atolados em argumentos semânticos, mencionam que utilizaram o

termo paternalista em sua abordagem apenas para incluir esforços para proteger as pessoas contra seus próprios erros e guiá-las às escolhas que fariam se fossem completamente informadas e livres de vieses comportamentais.

A abordagem é nitidamente mais ampla do que aquela adotada por Dworkin e por outros autores, como Feinberg[134]. Note-se, ainda, que Sunstein e Thaler não se preocupam apenas com medidas tomadas para evitar danos aos sujeitos. Como o bem-estar é o objetivo final – de modo semelhante ao proposto por Dworkin –, são igualmente paternalistas as medidas adotadas para criar benefícios aos agentes.[135] Se, por exemplo, um sujeito deixa de investir seu dinheiro em opções mais rentáveis, ainda que não haja efetivo desembolso de tais quantias e, portanto, não haja prejuízo, seria possível pensar em uma intervenção estatal, para *direcionar* as suas escolhas de investimento para opções que lhe gerem maior benefício.

A definição de paternalismo adotada por Sunstein será adotada nesta dissertação. Embora ela seja mais ampla, esta visão parece englobar de maneira mais clara as medidas propostas pelos paternalismos comportamentais. Por essa razão, essa abordagem é mais útil para as pretensões deste trabalho. Ainda que seja possível dizer que essas medidas, a rigor, não são paternalistas, já que simplesmente buscam direcionar escolhas para se aproximar das reais preferências dos indivíduos – argumento que será mais bem exposto abaixo -, essa discussão extrapola os objetivos desta dissertação. Em outros termos, não se pretende chegar a um consenso a respeito de quais seriam as exatas e necessárias condições do paternalismo - até porque esse consenso não parece ser possível.

Estabelecida essa premissa conceitual, algumas distinções dentro do espectro paternalista são necessárias antes de expor as diferentes correntes que formam o que aqui se denomina paternalismos comportamentais.

Um primeiro ponto envolve a distinção entre o paternalismo rígido (*hard paternalism*) e o paternalismo suave (*soft paternalism*). De acordo com Sunstein (2014a, p. 56-58), haveria duas maneiras distintas de expressar a diferença entre esses dois tipos de paternalismo. Em uma primeira abordagem, seria possível avaliar intervenções paternalistas em termos de um *continuum* entre medidas mais rígidas e medidas mais suaves, a depender da magnitude dos

[134] O conceito de Feinberg é ainda mais restrito do que o de Dworkin. Para Feinberg (1986), o paternalismo jurídico (*legal paternalism*) é a visão segundo a qual prevenir dano (físico, psicológico ou econômico) ao sujeito sempre será um bom e relevante motivo em favor de uma proibição criminal, ainda que haja consentimento desse indivíduo. Como Feinberg se preocupa com proibições criminais, naturalmente, as medidas que o autor tem em vista sempre são coercitivas.

[135] Feinberg (1986, p. 8) separa as normas coercitivas justificadas pela necessidade de proteger as pessoas de danos provocados por elas mesmas – que consistiriam no próprio paternalismo jurídico - e as normas justificadas a partir do desejo de agir de modo a criar benefícios positivos para os agentes, o que Feinberg denomina paternalismo que promove benefícios (*benefit-promoting paternalism*).

custos - de qualquer tipo - impostos aos decisores. Nessa perspectiva, estar diante de um paternalismo rígido significa dizer que os arquitetos de escolha estão impondo grandes custos aos agentes, enquanto um paternalismo suave implicaria custos pequenos aos decisores. A diferença poderia se expressar, por exemplo, ao se comparar a proibição do uso de maconha e uma simples multa civil; ou, por outro lado, a proibição de dirigir um veículo sem cinto de segurança e campanhas educacionais que apresentam avisos gráficos vívidos aos agentes. Por outro lado, Sunstein argumenta que, se o que se busca são distinções categóricas, em vez de um *continuum*, o foco pode se deslocar para a medida em que a intervenção impõe custos materiais aos agentes. Nessa abordagem, o paternalismo rígido estaria relacionado a ações do governo que tentam melhorar o bem-estar das pessoas a partir da imposição de custos materiais às suas escolhas. O paternalismo suave, em contrapartida, envolveria ações do governo que tentam melhorar o bem-estar das pessoas ao influenciar suas escolhas sem impor custos materiais às suas escolhas.

Embora a distinção proposta por Sunstein entre o paternalismo rígido e o paternalismo suave seja útil para a avaliação dos custos que diferentes políticas podem impor aos decisores - o que, na abordagem adotada nesta dissertação, é relevante para a apreciação da legitimidade de diferentes medidas -, essa não parece ser a única abordagem a respeito de tal distinção na literatura. Como expõe Feinberg (1986, p. 12), a distinção entre o paternalismo rígido e o paternalismo suave tem relação com o peso atrelado à voluntariedade da ação ou do consentimento do agente decisor. Nessa perspectiva, o paternalismo rígido aceitaria como razão para a intervenção estatal na vida privada a necessidade de proteger adultos competentes das consequências negativas de suas escolhas voluntárias; o paternalismo suave, por seu turno, defenderia tais intervenções apenas quando a conduta for substancialmente não voluntária ou quando uma intervenção temporária é necessária para se estabelecer se há tal voluntariedade (FEINBERG, 1986, p. 12)[136]. O exemplo clássico de Mill (2017, p. 196-197) esclarece essa distinção. Segundo Mill, se um agente público ou qualquer outra pessoa vê um terceiro tentando cruzar uma ponte que tenha sido considerada insegura e não há tempo para alertá-lo do perigo, eles podem detê-lo e impedir a sua passagem, sem que haja qualquer violação real à

[136] O mesmo sentido é adotado por Dworkin (2020), que menciona que o paternalismo suave justifica a ação paternalista apenas quando ela for necessária para determinar que um indivíduo está agindo voluntariamente e com conhecimento de causa. Conly (2013, p. 45), por sua vez, menciona que o paternalismo rígido é contraposto ao paternalismo suave de duas maneiras: a) ele pode significar forçar as pessoas a agir ou evitar um rumo de ação, em vez de apenas manipular escolhas tal como faz o paternalismo libertário; e b) ele pode significar a imposição de ações em pessoas que não escolheriam ainda que estivessem completamente informadas, enquanto o paternalismo suave estimula que as pessoas tenham pleno conhecimento de causa e decidam em conformidade com a sua vontade.

sua liberdade. Afinal, a liberdade está relacionada a fazer o que se deseja, e o sujeito que atravessaria a ponte não desejava cair no rio. Nesse caso, como não havia plena voluntariedade do rumo de ação que estava sendo tomado, interferir em tal escolha seria uma medida de paternalismo suave legítima, ainda que sob a perspectiva liberal de Mill[137,138].

Outra distinção relevante envolve os conceitos de paternalismo forte e fraco. Segundo Dworkin (2020), um paternalista fraco acredita que há legitimidade em uma intervenção nos meios que os agentes escolhem para atingir determinado fim, se os meios escolhidos acabarem não efetivando esse fim. Por exemplo, se um sujeito prefere segurança à conveniência, exigir o uso de cinto de segurança seria uma medida de paternalismo fraco.[139] Um paternalista forte, por seu turno, como acrescenta Dworkin (2020), acredita que as pessoas podem ter fins equivocados, confusos ou irracionais. Por consequência, haveria legitimidade para uma interferência a fim de evitar que essas pessoas atinjam esses fins. Pode-se imaginar, por exemplo, um sujeito que prefere efetivamente sentir o vento em seus cabelos do que utilizar capacetes enquanto pilota uma motocicleta. Um paternalista forte defenderia que se exigisse a utilização de capacetes, por entender que os fins buscados pelo sujeito em questão são irracionais ou equivocados.

Os mesmos exemplos apresentados no parágrafo anterior permitem que se explore mais uma distinção dentro do conceito de paternalismo. Uma intervenção como a exigência de utilização de capacetes pode ser justificada tanto por fins paternalistas – evitar que as pessoas sofram por conta de atos praticados ou consentidos por elas mesmas –, quanto por outras finalidades, como a redução de externalidades e o aumento do bem-estar social.[140] Nesse caso, quando há razões para intervir que vão além do paternalismo, está-se diante de um paternalismo

---

[137] Feinberg (1986, p. 13-15) não considera o paternalismo suave como parte do paternalismo – que estaria limitado ao que se denomina paternalismo rígido. Em sua abordagem, o paternalismo suave está de acordo com o espírito do liberalismo e, em vez de tentar distinguir o paternalismo rígido do suave, Feinberg considera que seria melhor alargar a definição de liberalismo para considerar o paternalismo suave um dos princípios válidos para limitar a liberdade.

[138] No clássico "*A Theory of Justice*", John Rawls (1999, p. 218-220) argumenta que, na posição original, isto é, no cenário hipotético em que as pessoas se encontram em plena igualdade, sob o véu da ignorância, sem conhecimento das contingências sociais e naturais que podem criar vantagens e desvantagens para elas, os princípios do paternalismo são aqueles que as pessoas reconheceriam como válidos para protegê-las contra fraquezas e enfermidades de sua razão e de sua vontade. Em sua perspectiva, uma intervenção paternalista deve ser justificada por falhas evidentes ou pela ausência de razão e vontade, além de ser guiada pelos princípios da justiça e pelo que se sabe a respeito dos objetivos e preferências mais permanentes dos sujeitos. Ao fim e ao cabo, as pessoas tentariam simplesmente garantir a integridade da sua própria pessoa e de seus objetivos e crenças definitivos, e o paternalismo seria uma proteção contra a nossa própria irracionalidade. Em outras palavras, a perspectiva de Rawls se aproxima do que seria um paternalismo suave na abordagem exposta acima.

[139] O exemplo comumente utilizado por Sunstein (2014a, p. 61) para ilustrar o paternalismo fraco - que ele denomina paternalismo de meios - é um Sistema de Posicionamento Global (GPS). Esse sistema apenas fornece informações para as pessoas sobre como ir de um lugar a outro. As pessoas podem até ignorá-lo, mas correm o risco de prejudicar os seus fins - chegar ao lugar certo - se fizerem isso.

[140] Cf. nota 133 *supra*.

misto (*mixed paternalism*) (FEINBERG, 1986, p. 8). O paternalismo não misto (*unmixed paternalism*), por seu turno, envolve medidas justificadas unicamente por motivos paternalistas: não há outra razão para intervir além de prevenir o dano que o próprio indivíduo causa a si mesmo ou gerar benefícios a ele (FEINBERG, 1986, p. 8).

Do mesmo modo, há distinções entre o paternalismo puro e o paternalismo impuro (DWORKIN, 2020). Em alguns casos, uma medida paternalista será imposta a um terceiro que interage com o beneficiário da medida, com o seu próprio consentimento. É o caso, por exemplo, do produtor de cigarros que vende os seus produtos ao consumidor e deve seguir regras como a proibição de propagandas e a inclusão de avisos chamativos nas caixas do produto. Essa é uma medida de paternalismo impuro, já que a classe de pessoas afetadas pela medida é maior do que a classe de pessoas que estão sendo protegidas. No caso do paternalismo puro, essas classes se confundem: a medida é imposta ao próprio sujeito que por ela será protegido. Um exemplo é a proibição de se nadar em uma praia sem que haja salva-vidas no local. Nesse caso, a medida recai sobre o próprio sujeito a quem se está protegendo.

Por fim, também é preciso diferenciar o paternalismo do moralismo legal, em cuja perspectiva pode ser moralmente legítimo para o Estado proibir certas ações, na medida em que elas sejam consideradas imorais, ainda que não haja dano ou ofensa para qualquer pessoa (FEINBERG, 1988). Nesse caso, o bem-estar do sujeito cujo comportamento é regulado não é um dos fundamentos da intervenção; não há, portanto, paternalismo em medidas que são fundamentadas dessa maneira. A proibição do consumo de maconha, por exemplo, ainda que tenha um componente paternalista e possa ter, ainda, fundamentos em eventuais externalidades, parece ter um largo componente moral. Na perspectiva adotada nesta dissertação, políticas com fundamentos baseados apenas no moralismo legal não têm legitimidade. Esse tópico será mais bem discutido no Capítulo 6.

Expostos esses conceitos, a próxima subseção se dedicará à exposição dos diferentes tipos de paternalismo comportamental.

2.3.2. Paternalismos comportamentais suaves: o paternalismo libertário, o paternalismo assimétrico e o paternalismo leve pragmático

Como já mencionado nesta dissertação, o mapeamento de diversas falhas de mercado comportamentais levou ao desenvolvimento de medidas paternalistas, em situações nas quais a teoria econômica tradicional não prescreveria intervenção alguma. Boa parte dessas medidas são formas de paternalismos suaves, tanto por envolverem medidas de baixo custo e não imporem custos materiais aos agentes, quanto por tentarem direcioná-los àquele que seria o seu verdadeiro interesse, se fossem dotados de completa informação e completa capacidade computacional.

O movimento mais célebre nesse sentido é o paternalismo libertário (*libertarian paternalism*). De acordo com Thaler e Sunstein (2003b, p. 4), por exemplo, inúmeros analistas do direito, incluindo economistas e juristas orientados para a economia, defendem uma posição antipaternalista dogmática, baseada em uma combinação de uma falsa premissa e de duas posições equivocadas.

A premissa equivocada se confunde com a soberania dos consumidores, que passa a ser questionada a partir dos já mencionados fatores supostamente irrelevantes. Afinal, como sustentam Thaler e Sunstein (2003b, p. 5), escolher bem é uma questão empírica e, em geral, as pessoas tomam melhores decisões em contextos nos quais têm experiência e boas informações do que nos contextos em que elas são inexperientes ou ignorantes. Nesse contexto, em razão de falhas em seu comportamento, as pessoas podem prejudicar excessivamente o seu próprio bem-estar. A intervenção estatal pode, então, ser recomendada para amenizar esses prejuízos e, nesse caso, pode ser valorizada pelos próprios indivíduos.

Por outro lado, a primeira concepção equivocada do movimento antipaternalista se expressaria na ideia de que existem alternativas viáveis ao paternalismo (THALER; SUNSTEIN, 2003a, p. 175). De fato, como já mencionado, as pessoas tomam muitas decisões intuitivas e podem ser influenciadas por pequenas mudanças no contexto em que as decisões são tomadas. Em inúmeras circunstâncias, organizações e agentes governamentais são obrigados a escolher a forma como esse contexto será organizado. Por exemplo, a ordem em que a comida é disposta em um refeitório pode afetar as escolhas alimentícias feitas pelos indivíduos (THALER; SUNSTEIN, 2003a, p. 175). Nessa hipótese, qualquer estrutura adotada seria interventiva, uma vez que direcionaria as escolhas das pessoas para alguma das opções – ainda que de forma aleatória. Nesse contexto, entre a aleatoriedade, opções que piorem o bem-estar dos indivíduos e a opção paternalista, que se preocupa com o bem-estar dos decisores, a

opção a ser a escolhida, para Thaler e Sunstein, parece bastante clara[141]. Por essa razão, Thaler e Sunstein concluem que são raras as situações em que existem opções viáveis a algum tipo de paternalismo.

A segunda concepção equivocada do movimento antipaternalista, na visão de Thaler e Sunstein (2003a, p. 175), é a ideia de que o paternalismo está sempre atrelado à coerção estatal. No exemplo mencionado acima, a opção por apenas remanejar a forma como os alimentos estão dispostos em um refeitório não envolve qualquer coerção. Evidentemente, se o diretor escolhe, por razões paternalistas, apresentar alimentos saudáveis em locais nos quais eles serão mais consumidos, não haverá qualquer coerção envolvida. Assim, se, entre os empregados, há indivíduos que não desejam consumir alimentos saudáveis, eles não serão coagidos caso se comportem dessa maneira. Portanto, é possível adotar medidas paternalistas não coercitivas.

A partir de tais conclusões, Thaler e Sunstein começaram a defender que algumas formas de paternalismo deveriam ser aceitáveis mesmo para libertários convictos. Surge, então, o paternalismo libertário. Resumidamente, Sunstein e Thaler (2003b, p. 1198) o conceituam como uma abordagem que, ao mesmo tempo em que preserva a liberdade de escolhas, autoriza que entidades direcionem pessoas para opções que promoverão seu próprio bem-estar. O objetivo seria evitar efeitos aleatórios, inadvertidos ou prejudiciais da arquitetura de escolhas e produzir uma situação que promoverá o bem-estar das pessoas.

O aspecto libertário das estratégias adotadas por Thaler e Sunstein (2009, p. 5, 2021, p. 6) reside na concepção de que as pessoas devem ser livres para fazer o que desejam e para escolher não participar de políticas que não lhes sejam aprazíveis. No mencionado exemplo da cafeteria, não se restringem as opções de escolha. Se, em um dado momento, certo indivíduo deseja consumir alimentos gordurosos, poderá fazê-lo sem maiores interferências em sua estrutura de incentivos. O paternalismo libertário, nesse contexto, é um paternalismo que preserva a liberdade.

O paternalismo libertário não deixa de ser paternalista, no entanto, na medida em que tenta influenciar o comportamento das pessoas a fim de melhorar sua própria vida (THALER; SUNSTEIN, 2009, p. 5, 2021). No caso da cafeteria, ainda que a liberdade de escolha seja preservada, a opção por alimentos saudáveis em posições nas quais eles seriam mais consumidos busca melhorar o bem-estar dos decisores, modificando suas decisões com uma intervenção. A política seria, então, paternalista na medida em que tenta influenciar escolhas a

[141] Outras opções poderiam ser utilizadas para a disposição dos alimentos. Aqui, entram as estratégias tradicionais de comando e controle: banimentos, subsídios, opções obrigatórias etc.

fim de melhorar a situação dos decisores segundo critérios julgados por eles próprios (*as judged by themselves*).

Além disso, o paternalismo libertário é um paternalismo fraco, suave e não intrusivo (THALER; SUNSTEIN, 2021, p. 7). Em outras palavras, na proposta de Thaler e Sunstein, em tese, são propostas intervenções que não tentam alterar os fins dos agentes, mas apenas facilitar que esses agentes adotem os meios mais adequados para atingir esses fins; ao mesmo tempo, esse paternalismo é suave, já que não impõe custos elevados aos decisores e almeja fazer com que esses indivíduos façam as escolhas que fariam se prestassem completa atenção e tivessem completa informação, capacidade cognitiva ilimitada e completo autocontrole (THALER; SUNSTEIN, 2021, p. 7).

A rigor, no entanto, Thaler e Sunstein não parecem defender que esses seriam os únicos fundamentos para a regulação do comportamento privado em termos paternalistas. Ao discutir abordagens para enfrentar internalidades, por exemplo, Sunstein (2014a, p. 38) chega a mencionar que é possível imaginar casos em que incentivos econômicos ou mandados - medidas que impõem custos altos para os decisores e poderiam se qualificar como formas de paternalismo rígido - podem ser as melhores soluções. O autor exemplifica com a proibição de que motoristas enviem mensagens de texto enquanto dirigem, ainda que exista um fundamento, aqui, ligado à proteção de terceiros e ao princípio do dano. Mais adiante, Sunstein (2014a, p. 142) chega a dizer que, se os benefícios de uma imposição forem claramente superiores aos seus custos e, se não houver um problema real do ponto de vista da autonomia, essas medidas de comando e controle parecem ser justificadas, ainda que em termos paternalistas. O seu fundamento, quanto a esse ponto, é diretamente relacionado às consequências da medida. Como Sunstein admite expressamente, medidas de comando e controle não devem ser deixadas de lado simplesmente por serem paternalistas. A questão, na verdade, depende das suas consequências.

No mesmo sentido, mais recentemente, Thaler e Sunstein (2021, p. 333) argumentaram que, em casos em que as pessoas estão tomando decisões estúpidas, míopes e autodestrutivas, medidas de comando e controle não devem ser descartadas. Eles ressaltaram que não se opõem a programas de seguridade social, proibições do consumo de gorduras trans, obrigações ligadas à eficiência energética, exigência de se utilizar capacetes e cintos de segurança ou de utilizar máscaras durante uma pandemia. Mesmo que esses exemplos contenham fundamentos de intervenção não paternalistas, Thaler e Sunstein (2021, p. 333-334) afirmam que, se as escolhas das pessoas impuserem danos aos seus *selves* futuros, políticas mais interventivas do que aquelas propostas pelo paternalismo libertário podem ser defendidas.

Em sua perspectiva, como medidas de comando e controle levantam diversos problemas e preocupações, enquanto as pessoas estiverem tomando decisões informadas sobre como viver suas próprias vidas, deveríamos manter uma atitude de humildade e respeito, incluindo uma presunção em favor da liberdade de escolha. As intervenções, contudo, não são descartadas: basta a demonstração empírica de que tais decisões não são suficientemente informadas, para que intervenções possam ser propostas. Afinal, os próprios autores admitem que a sua "estrela-guia é o bem-estar humano"[142] (THALER; SUNSTEIN, 2021, p. 332).

Como os parágrafos anteriores demonstram, Thaler e Sunstein adotam, nitidamente, uma posição consequencialista e defendem intervenções estatais variadas a depender das consequências dessas medidas. Nesse caso, tanto podem ser defendidas medidas de paternalismo suave, enquanto intervenções que implicam baixos custos para os decisores, quanto propostas mais rígidas, baseadas em medidas de comando e controle, de teor coercitivo. Nesse contexto, um questionamento que surge ao se analisar a obra de Thaler e Sunstein é: se sua proposta é puramente consequencialista, por que, então, defender um paternalismo libertário? Em que medida ele seria necessário? Afinal, Thaler e Sunstein (2021, p. xiii) chegam até mesmo a ressaltar que apenas os dois gostam do termo paternalismo libertário, que gera questionamentos infindáveis, sobre os mais diversos vieses, como será exposto no Capítulo 6.

A avaliação da obra de Sunstein em um caráter mais amplo auxilia a identificar respostas a essas questões. Um dos pontos relevantes de sua obra é a defesa de acordos não completamente teorizados[143]. Nessa perspectiva, quando as pessoas divergem em alguma proposição de relativo alto nível – como a defesa de uma visão filosófica kantiana ou utilitarista - , elas podem ser capazes de entrar em acordos a respeito de níveis mais baixos de abstração (SUNSTEIN, 2018a, p. 37). Julgamentos não completamente teorizados em casos particulares são, para Sunstein, o material ordinário do direito. As pessoas podem concordar, por exemplo, que um limite de velocidade de 65 milhas por hora faz sentido e punir alguém que descumpre essa regra, sem pensar a respeito das teorias criminais da pena (SUNSTEIN, 2018a, p. 38).

Nesse contexto, diante de um impasse a respeito de questões de mais alta ordem - como a defesa de um paternalismo rígido ou de um libertarianismo, ou, mesmo, das teorias filosóficas que podem embasar essas posições, como a deontologia kantiana ou o utilitarismo -, Thaler e Sunstein apresentaram uma proposta de acordo não completamente teorizado: o paternalismo libertário. Nesta perspectiva, ainda que paternalistas e libertários não conseguissem chegar a

[142] Tradução livre. Texto original: "*Our lodestar is human well-being* […]".
[143] Cf., nesse sentido, (SUNSTEIN, 1995, 2018a).

um acordo em termos mais abstratos, quanto aos fundamentos de regulação do comportamento privado, seria possível defender a adoção de medidas que a) almejam aprimorar o bem-estar das pessoas; e b) fazem isso sem coerção, a partir, simplesmente, da exploração de fatores supostamente irrelevantes que influenciam decisões.

Não por outra razão, Thaler e Sunstein (2009, p. 255) esperavam que o paternalismo libertário oferecesse uma verdadeira terceira via, que pudesse superar alguns dos debates menos tratáveis das democracias contemporâneas. Na recente edição final do *Nudge*, inclusive, Thaler e Sunstein (2021, p. 20) acrescentam que, embora o mundo esteja se tornando mais e mais polarizado, eles continuam a acreditar que o paternalismo libertário pode ser uma fundação promissora para o bipartidarismo e para a resolução de problemas simples. Apesar de todas as diferenças entre grupos polarizados, Thaler e Sunstein esperam que pessoas com diferentes convicções políticas estejam dispostas a defender o paternalismo libertário. Em outras palavras, Thaler e Sunstein acreditam no sucesso do seu acordo não completamente teorizado.

As críticas ao paternalismo libertário que serão expostas no Capítulo 6 evidenciam, no mínimo, que esse acordo enfrenta desafios, embora o termo aparentemente paradoxal tenha sido importante para popularizar a teoria de Thaler e Sunstein. Mais do que isso, como será demonstrado, *nudges* não são suficientes por si só para resolver boa parte dos problemas a que se destinam políticas públicas. Como os próprios Thaler e Sunstein reconhecem, há situações que exigem intervenções estatais mais rígidas; em outros casos, a melhor estratégia pode ser deixar com que o mercado funcione livremente e encontre as suas próprias soluções. Fiar-se a um acordo não completamente teorizado entre paternalistas e libertários – que, ressalta-se, acaba por desagradar paternalistas e libertários – como proposta de um consenso para a elaboração de políticas públicas parece uma fundação frágil para a resolução de problemas. Afinal, essa posição acaba por evitar uma questão fundamental: "mercado e Estado não são alternativas um ao outro, e sim, muito pelo contrário, mutuamente dependentes" (TIROLE, 2020). Essas questões serão retomadas e mais discutidas no Capítulos 6.

Feita essa digressão, em uma perspectiva também baseada na racionalidade limitada, outra linha de pesquisas defende a adoção de um paternalismo assimétrico. Na elaboração de sua proposta, Camerer e outros autores (2003, p. 1222) colocam-se no meio do caminho entre pesquisadores conservadores, como Richard Posner, que adota a racionalidade de indivíduos como premissa de sua teoria[144], e pesquisadores influenciados pela economia comportamental, como Jolls, Sunstein e Thaler, que, à época, já criticavam a postura antipaternalista, embora

[144] A premissa de racionalidade é adotada por Posner em seus modelos de análise econômica do direito. Cf., nesse sentido: (POSNER, 1986).

ainda sem uma defesa afirmativa do paternalismo[145]. Nesse contexto, Camerer e outros autores deram um passo além e elaboraram uma defesa afirmativa do paternalismo, embora apenas de medidas paternalísticas que fossem assimétricas.

Embora as intervenções não precisem estar associadas a algum tipo de libertarianismo, a principal contribuição do paternalismo assimétrico é a ênfase na possibilidade de se gerar benefícios para agentes com racionalidade limitada, com pouco ou nenhum prejuízo para agentes completamente racionais (CAMERER *et al*., 2003, p. 1212). No limite, se esses prejuízos não forem gerados, o paternalismo assimétrico seria puro, situação que só traria benefícios aos consumidores (CAMERER *et al*., 2003, p. 1219).

Na visão do paternalismo assimétrico, uma medida específica seria defensável se enquadrar-se ao seguinte modelo: $(p * B) – [(1 – p) * C] – 1 + \Delta\Pi > 0$ (CAMERER et al., 2003, p. 1219). Em outras palavras, o paternalismo assimétrico será configurado desde que uma determinada política crie grandes benefícios (B) para indivíduos cuja racionalidade é limitada (p), enquanto impõe pouco ou nenhum prejuízo (C) para os indivíduos que seriam totalmente racionais. Ao mesmo tempo, é necessário que os custos de implementação (I) sejam baixos e que a política aumente a eficiência econômica, somados os resultados líquidos dos indivíduos que serão empurrados e das empresas ($\Delta\Pi$) que serão afetadas pela política[146]. Dessa forma, uma política fará parte do paternalismo assimétrico se apresentar resultado líquido positivo ao ser enquadrada ao modelo mencionado.

Nessa proposta, a tentativa de se atingir uma terceira via no que concerne aos graus de intervenção estatal, isto é, a busca por um acordo não completamente teorizado, é bastante semelhante àquela empreendida por Thaler e Sunstein, no paternalismo libertário[147]. De fato,

---

[145] De fato, Jolls, Sunstein e Thaler (1998, p. 1541) chegaram a mencionar que a racionalidade limitada conduziria a uma espécie de anti-antipaternalismo, isto é, um ceticismo a respeito do antipaternalismo, embora não defendessem expressamente medidas paternalistas. Posteriormente, Thaler e Sunstein passaram a fazer essa defesa, cunhando a expressão paternalismo libertário, como já exposto nesta subseção.

[146] Note-se que esses custos existirão se a política for uma medida de paternalismo impuro, como exposto na Subseção 2.3.1 acima. Em outras palavras, pode ser que o número de agentes afetado pela medida seja superior ao tamanho da classe de indivíduos que se pretende proteger com a medida. Nesse caso, a medida paternalista pode impor custos para os agentes que interagem com o grupo de indivíduos a ser protegido. Como Sunstein (2015b, p. 424) explica, ao discutir os *nudges* – medidas que implementam o paternalismo libertário e serão objeto de discussão nos Capítulos 3 e 4 –, embora essas políticas preservem a liberdade de escolhas para a população que será afetada, há, paralelamente, medidas de comando e controle para as organizações que organizam a arquitetura de escolhas. Por exemplo, um restaurante que, a partir de uma determinada norma municipal, é obrigado a retirar os saleiros de suas mesas, para reduzir o consumo de sal de seus consumidores é objeto de uma norma de comando e controle e, ao mesmo tempo, direciona o comportamento dos seus consumidores. Na terminologia adotada nesta dissertação, como será demonstrado no Capítulo 4, esses são *nudges* impuros, que se opõem aos *nudges* puros, elaborados diretamente pelo Estado, quando o arquiteto de escolhas é uma entidade pública.

[147] A semelhança não é mera coincidência. Como registrado por Loewenstein e Chater (2017), ambos os movimentos se inspiraram em um evento que ocorreu em 1997, no qual Rabin propôs a ideia de se identificar uma abordagem para políticas públicas que tivesse apelo ao longo de todo o espectro político.

além de se situar no meio do caminho entre autores conservadores e autores ligados à economia comportamental, Camerer e seus coautores (2003, p. 1222) ressaltam que as políticas de sua proposta impõem custos mínimos, caso os conservadores estejam corretos e a maioria das pessoas seja de fato racional, e benefícios máximos, se a racionalidade e o autocontrole forem tão limitados quanto os economistas comportamentais presumem. Não por outra razão, o paternalismo assimétrico é também conhecido como paternalismo para conservadores.

Além disso, como se poderia esperar, boa parte das medidas do paternalismo assimétrico se confundem com propostas do paternalismo libertário. Afinal, como as medidas do paternalismo libertário são pouco custosas e impõem poucos prejuízos para aqueles que não desejam segui-las - uma vez que não há coerção -, em inúmeras situações, elas apresentaram um resultado líquido positivo, se testadas à luz do modelo do paternalismo assimétrico. Organizar um refeitório para que alimentos mais saudáveis fiquem mais próximos dos olhos dos consumidores, por exemplo, pode ser uma proposta defensável tanto em termos paternalistas libertários quanto em termos paternalistas assimétricos. De todo modo, como o paternalismo assimétrico se pauta pelo resultado líquido que uma determinada medida gera, em termos de custos e benefícios, em princípio, ele não excluiria medidas de comando e controle de seu escopo, embora essas propostas possam gerar custos exacerbados para agentes racionais, o que diminui a probabilidade de um resultado líquido positivo.

Outra perspectiva de paternalismo suave comportamental é o paternalismo leve pragmático (*pragmatic light paternalism*) de Loewenstein[148] e Emily Haisley (2008). Inicialmente, Lowenstein e Haisley apresentam as críticas aos diversos critérios de bem-estar que são propostos para substituir a ideia de preferências reveladas pelas escolhas. Esse é um dos desafios impostos aos paternalismos suaves, mas este é um tema que será enfrentado em maiores detalhes no Capítulo 6. O que convém ressaltar, ainda nesta seção, é que, diante de

[148] Embora tenha sido um dos autores da proposta de um paternalismo assimétrico, Loewenstein se afastou expressamente de tal movimento. De fato, Loewenstein e Chater (2017, p. 33) afirmaram que, embora o paternalismo assimétrico tenha feito uma crítica às formas históricas de paternalismo que focam em categorias de indivíduos considerados incompetentes para tomar decisões autointeressadas e defendido um paternalismo com foco em situações em vez de pessoas, essa abordagem acabou por atribuir a necessidade de paternalismo a deficiências individuais, o que Loewenstein e Chater passaram a considerar uma posição equivocada. Em sua visão, considerar que erros na tomada de decisão são deficiências individuais atribui a culpa aos indivíduos por problemas que, muitas vezes, são estruturais em sua natureza. Por exemplo, embora altas taxas de obesidade sejam um fenômeno prevalente nos Estados Unidos atualmente, não parece ser razoável pensar que tenha havido um crescimento recente nos descontos intertemporais ou na prevalência do viés do presente. As causas para a emergência de tais problemas podem, então, ser estruturais, como mudanças nos preços absolutos e relativos de alimentos frescos e empacotados, diminuição dos tempos para preparo de alimentos, maior número de lanches e aumento do tamanho das porções (LOEWENSTEIN; CHATER, 2017, p. 33).

dificuldades para se chegar a um critério claro de avaliação do bem-estar dos indivíduos submetidos a intervenções paternalistas, Loewenstein e Haisley resolvem ser pragmáticos.

Em sua proposta, embora as justificativas para a adoção de paternalismos suaves possam ser mais brandas do que aquelas exigidas para a imposição de paternalismos rígidos, mesmo políticas de paternalismos leves só devem ser adotadas quando julgamentos de bem-estar forem relativamente diretos e simples (LOEWENSTEIN; HAISLEY, 2008, p. 221). A ideia é buscar critérios em que intervenções gerem claramente benefícios de bem-estar, independentemente do critério normativo adotado para avaliá-lo.

Na proposta de Loewenstein e Haisley (2008, p. 222–223), três condições permitem que medidas de paternalismo leve sejam adotadas.

A primeira delas é baseada no critério de dominância: se o resultado de uma estratégia claramente dominar a de outra, o paternalismo leve pragmático recomendaria uma intervenção. No caso da definição do percentual de contribuição de um empregado para um plano de previdência no contexto americano, por exemplo, maximizar a sua contribuição própria até o limite máximo da contribuição correspondente da empresa é uma estratégia dominante para os empregados; qualquer estratégia distinta faria com que eles deixassem de ganhar recursos ao não tirar proveito da contribuição correspondente a ser feita pela empresa (LOEWENSTEIN; HAISLEY, 2008, p. 222).[149]

Além disso, o paternalismo leve pragmático também recomendaria intervenções em casos de resultados claramente negativos. De acordo com Loewenstein e Haisley (2008, p. 222), em determinadas circunstâncias, as pessoas tomam decisões que claramente as levam para um caminho prejudicial e, indiscutivelmente, indesejado, não pretendido e contrário aos próprios interesses do decisor. Nesse caso, não seria necessário identificar um critério de bem-estar

[149] Como Thaler e Sunstein (2021, p. 181) explicam, nos Estados Unidos, em 1980, um novo tipo de plano de previdência baseado no empregador foi desenvolvido – o plano conhecido como 401(k) em razão da seção da lei que o tornou possível. Esse plano fez com que as empresas mudassem para planos com contribuições definidas, o que significa que os planos estipulavam quanto empregados e empregadores deveriam contribuir para uma conta protegida de impostos em nome dos empregados. Embora as contribuições fossem definidas, os benefícios recebidos pelos empregados em sua aposentadoria dependiam do nível de poupança e de investimento que eles adotavam, assim como do desempenho de tais investimentos. Nesses planos, o empregador deveria fazer uma contribuição correspondente a algum percentual da contribuição do empregado, até um determinado limite, como 6% do seu salário (THALER; SUNSTEIN, 2021, p. 184). Daí decorre o fato de que contribuir até o limite máximo da contribuição correspondente da empresa é uma estratégia dominante para os empregados; contribuições menores fariam com que a empresa não atingisse o seu limite máximo de contribuição correspondente. Inclusive, diante dos baixos índices de contribuição, motivados por fatores supostamente irrelevantes, Thaler e Shlomo Benartzi (2004) desenvolveram um programa de sucesso, que explorou fatores como o viés do *status quo* e aversão à perda para contribuir para um aumento considerável no número de contribuições e nas taxas de contribuição adotadas pelos empregados. O programa foi denominado *Save More Tomorrow* e chegou a ser convertido em lei nos Estados Unidos. Como será discutido na próxima subseção, no entanto, há quem questione o sucesso de tal programa. Cf., nesse sentido, (BUBB; PILDES, 2014).

preciso: as pessoas estariam em melhor condições se uma intervenção fosse proposta. O exemplo de Loewenstein e Haisley é a utilização de empréstimos consignados (*payday loans*), que aumentam as chances de um sujeito ir à bancarrota. Sua insolvência seria um resultado nitidamente negativo que fundamentaria uma intervenção paternalista.

Por fim, o terceiro fundamento elencado por Loewenstein e Haisley (2008, p. 222) é a autoarbitragem (*self-officiating*), isto é, a possibilidade de que as pessoas, tendo escolhido previamente seus próprios objetivos, recomendem restrições, incentivos ou o fornecimento de informações para aprimorar seu autocontrole. Na medida em que essas escolhas não sejam feitas no calor de alguma emoção intensa, a fim de evitar a lacuna de empatia quente-fria, isto é, a tendência a se subestimar a influência de estados emocionais, como a raiva, a dor ou a fome no comportamento ou nas preferências futuras[150], esse seria um fundamento adequado para intervenções paternalistas segundo Loewenstein e Haisley. Se, por exemplo, um sujeito acima do peso acredita piamente que estaria em condições melhores se perdesse peso e, consistentemente, reporta que uma intervenção paternalista leve poderia ajudá-lo, uma medida como essa seria justificável.

Feitos esses apontamentos, convém destacar ainda que a abordagem pragmática de Loewenstein e Haisley (2008, p. 223) é expressamente baseada na satisfação das preferências dos indivíduos. O economista, nessa visão, aqui visto como um planejador de políticas públicas, deve se atentar às afirmações verbais de preferências como informações úteis, ainda que elas não estejam alinhadas às escolhas efetivamente feitas pelo indivíduo. Sua posição, na analogia de Loewenstein e Haisley, seria como a de um psicoterapeuta que leva em consideração os desejos de seu paciente de ficar mais feliz, mais sociável e menos ansioso, ainda que ele se comporte de maneiras que levem a resultados opostos. Se o sujeito professa o interesse de parar de fumar, por exemplo, mas não consegue efetivar esse interesse, uma medida paternalista pode ser útil.

2.3.3. O paternalismo coercitivo comportamental: fatores supostamente irrelevantes comprometem a noção de autonomia

Como já mencionado acima, o mapeamento de fatores supostamente irrelevantes que influenciam as decisões não levou ao desenvolvimento apenas de paternalismos suaves. De

[150] De acordo com Loewenstein (2005), quando as pessoas estão em estados emocionais "quentes", isto é, fortemente influenciadas por estímulos emocionais, elas subestimam a influência desses estados e, assim, acabam superestimando a estabilidade de suas preferências naquele momento.

fato, esses mesmos fatores vêm sendo utilizados para embasar propostas de paternalismo coercitivo, por motivos comportamentais. Na Subseção 2.3.2, demonstrou-se, por exemplo, que Thaler e Sunstein não rejeitam essas propostas por si só; na verdade, chegam a aventar a possibilidade de prescrevê-las, inclusive, informando-se por critérios consequencialistas. A mesma hipótese pode, em tese, ser defendida sob fundamentos paternalistas assimétricos.

No entanto, os proponentes dos paternalismos suaves, em geral, não levam adiante análises de custos e benefícios envolvendo comparações entre suas propostas que preservam as escolhas e medidas de comando e controle. Realmente, não se costuma demonstrar com clareza que as ferramentas dos paternalismos suaves, como regras *default*, obrigações de transparência e mudanças de enquadramento (*framing*), são as ferramentas ótimas para enfrentar uma determinada falha de mercado comportamental (BUBB; PILDES, 2014, p. 1598).

Nesse contexto, Ryan Bubb e Richard H. Pildes (2014) argumentam que a análise econômico-comportamental do direito (*behavioral law and economics*), que está por trás dos paternalismos suaves mencionados, exclui ferramentas regulatórias tradicionais de seu arcabouço, como imposições, impostos e outras medidas que limitam as escolhas. Essa omissão, para Bubb e Pildes, é resultado de duas características da análise econômico-comportamental do direito. A primeira delas é a aspiração política de boa parte desses movimentos: como já mencionado, os paternalistas suaves muitas vezes querem oferecer uma terceira via política, capaz de gerar um consenso político amplo ao preservar escolhas. Além de tal aspiração política, outro motivo pode ser uma simples predisposição filosófica em favor da liberdade, o que leva os pesquisadores de tais vertentes a direcionar seu enfoque para ferramentas que preservam escolhas.

Os fatores supostamente irrelevantes mapeados pela economia comportamental, todavia, podem apontar que a solução ótima seja prescrever algum tipo de medida de comando e controle. No caso da poupança para a aposentadoria, por exemplo, Bubb e Pildes sustentam que uma análise econômico-comportamental minuciosa poderia indicar a necessidade de se instituir um sistema obrigatório de contribuições à seguridade social.[151] A rejeição de tais políticas pelos paternalistas suaves, porém, não seria motivada necessariamente por critérios de

[151] Bubb e Pildes (2014) criticam propostas baseadas na mudança de padrões de consentimento e em aumentos nas taxas padrões de contribuição, como o *Save More Tomorrow* (Cf. nota 149 *supra*), por entender que o sucesso de medidas como a proposta de Thaler e Benartzi é superestimado. Embora o número de contribuintes possa ter aumentado, muitos dos empregados acabavam por se vincular à taxa padrão em vez de aumentar suas contribuições. Mesmo o plano de aumentos periódicos proposto por Thaler e Benartzi, vinculados a aumentos salariais, também não seria ótimo, já que as contribuições voltariam ao mínimo a cada vez que o sujeito trocasse de emprego. Finalmente, como os sujeitos eram direcionados para contas livres de impostos, Bubb e Pildes mencionam que as distorções geradas por tais subsídios tributários podem ter causado prejuízos superiores a eventuais aumentos no número de contribuintes.

bem-estar, mas, em certos casos, é quase pressuposta. Embora Bubb e Pildes não rejeitem as propostas dos paternalismos suaves comportamentais, eles propõem uma visão mais ampla que, a partir de falhas de mercado comportamentais, também abre espaço para formas de paternalismo coercitivo comportamental.

Em outra perspectiva, há propostas de paternalismo coercitivo quase que por princípio, isto é, praticamente desvinculadas propriamente de um sopesamento claro e casuístico de custos e benefícios. O exemplo mais notável foi desenvolvido por Conly (2013). Seu argumento é bastante direto: em muito mais casos do que admitimos, a sociedade deveria agir e obrigar as pessoas a fazer o que é bom para elas. Segundo Conly, a base para valorizar a liberdade seria a crença de que as pessoas são agentes racionais. A existência de déficits cognitivos, no entanto, contesta essa premissa e, para Conly, retira o valor que normalmente é atribuído à autonomia. Em sua perspectiva, a existência de fatores supostamente irrelevantes justifica a adoção de um paternalismo coercitivo, para proteger as pessoas de fazer coisas que lhes são prejudiciais.

Conly compara a situação de um sujeito que toma decisões enviesadas com situações em que tradicionalmente se aceita uma intervenção coercitiva de terceiros, como casos de ignorância e de incompetência. No caso da ignorância, os custos de uma decisão ruim são muito altos e o conhecimento de um *expert* é necessário e está disponível. As pessoas reconhecem, então, que é melhor delegar a escolha. É o caso, por exemplo, de um paciente que só pode ter acesso a determinados medicamentos a partir de prescrições médicas. Já a incompetência justifica intervenções de terceiros na medida em que as pessoas não conseguem lidar adequadamente com os fatos e precisam de ajuda, ainda que não queiram. É o caso de uma criança que precisa de seus pais para tomar decisões por conta própria.

Na visão de Conly, a diferença entre a racionalidade do ignorante e do incompetente e a de uma pessoa adulta em suas faculdades normais é apenas de grau e não de espécie; ainda que em menor medida, o ser humano médio, em sua perspectiva, é frágil, erra frequentemente e gera prejuízos a si mesmo. Os fatores supostamente irrelevantes expostos no Capítulo 1 são uma evidência de tal argumento. Por essa razão, Conly argumenta que, assim como um paciente precisa de um médico e um incompetente precisa de seus pais para tomar decisões, as pessoas médias precisam da ajuda de intervenções governamentais.

Diante da suposta necessidade de ajuda, Conly compara as possíveis soluções que poderiam ser adotadas pelo governo. Ao avaliar medidas de educação ou aprendizado a partir da experiência, Conley pondera, em primeiro lugar, que vieses cognitivos não são de fácil controle ou apreensão. Em sua visão, não se poderia exigir que adultos aprendessem toda uma nova área da psicologia para ter consciência de suas limitações, enquanto crianças

simplesmente não teriam a capacidade de receber tais ensinamentos. Aprender com os próprios erros, por outro lado, pode não ser mais possível, como quando uma doença já foi diagnosticada, ou simplesmente pode não ser factível, diante de vieses como o excesso de otimismo. Uma outra possibilidade – já discutida na seção anterior – seria o paternalismo libertário, que também é rechaçado por Conly. Desta vez, os argumentos estão ligados à sua manipulabilidade e à sua baixa efetividade – questões que serão discutidas em maiores detalhes nos Capítulos 5 e 6.

Nesse contexto, a solução restante para que o governo protegesse os indivíduos de si próprios seria a prevenção coercitiva de determinados rumos de escolha. Assim, ainda que ações afetassem apenas o próprio indivíduo, como no caso de consumo abusivo de cigarros, alimentação excessivamente gordurosa e consumismo excessivo, Conly defende que uma intervenção de paternalismo coercitivo seria justificada. Exemplos de medidas seriam a proibição de cigarros, do consumo de gordura trans ou o controle da quantidade de dívidas que um indivíduo poderia assumir.

Embora sustente que a verdadeira necessidade de intervenção só pode ser descoberta pela análise empírica, não é essa a impressão que se tem do seu texto. Conly defende que, por princípio, não há áreas em que essa análise não poderia ser feita. No entanto, Conly defende a primazia de medidas coercitivas em comparação a outras ferramentas de intervenção sem se fundamentar em uma análise concreta de custos e benefícios. Em outras palavras, presume-se que as medidas coercitivas são, de fato, o melhor caminho para proteger as pessoas.

Conly também chega a reconhecer que legislações paternalistas só deveriam intervir quando as pessoas provavelmente tomam decisões que interfiram séria e irrevogavelmente em sua habilidade de atingir seus próprios objetivos; ou seja, embora proponha um paternalismo rígido, Conly também defende um paternalismo fraco, destinado apenas a corrigir os meios que as pessoas equivocadamente adotam para atingir os fins que elas mesmas elegem. No entanto, o espírito da proposta de Conly indica um espaço consideravelmente amplo para intervenções coercitivas, em contextos em que hoje elas não são aceitas sem apresentar evidências empíricas que justifiquem essa posição. Ao fim e ao cabo, sua proposta é, definitivamente, o paternalismo coercitivo comportamental por excelência.

2.4. Atos substancialmente autônomos: noção empiricamente factível de autonomia[152]

Toda a discussão exposta acima será retomada no Capítulo 6, quando serão discutidas as críticas ao paternalismo libertário e à utilização de *nudges* e será elaborada a tese proposta nesta dissertação. De todo modo, antes de finalizar este Capítulo, sobretudo após apresentar o paternalismo coercitivo comportamental de Conly, é fundamental estabelecer o que se entende por autonomia neste trabalho.

Ao se deparar com as falhas de mercado comportamentais, um questionamento que surge naturalmente - como agora já parece claro - é a medida em que essa constatação enfraquece a noção de autonomia. Afinal, por trás de políticas paternalistas, ainda que sem coerção, está a noção de que os indivíduos não são plenamente capazes de decidir livres de influências que vão além do seu controle nem conseguiriam se orientar com base em motivos racionais. As decisões tomadas por essas pessoas são autônomas? Ou, colocando a pergunta de outro modo, a racionalidade dos agentes é a base da proteção da autonomia?

Como visto na Subseção 2.3.3 acima, o argumento contra a autonomia é justamente aquele desenvolvido por Conly (2013). Sunstein (2014a, p. 125-127), por outro lado, em uma posição nitidamente consequencialista, vê a autonomia como um ingrediente do bem-estar em algumas situações. A lógica é simples: as pessoas gostam de tomar decisões, de modo que, em muitas circunstâncias, elas obtêm ganhos de bem-estar ao permanecerem livres para decidir. Assim, a análise da legitimidade de uma determinada intervenção em termos de autonomia - seja na forma de um *nudge*, seja por meio de uma medida de comando e controle - passaria por apurar se o ganho de bem-estar supera as perdas correspondentes. Outros sustentam, ainda, que os indivíduos atribuem um valor intrínseco à possibilidade de decidir livremente os rumos de ação que afetam o seu próprio bem-estar[153].

Como já mencionado na Introdução, este trabalho não pretende se debruçar sobre o profundo e interminável debate filosófico em torno das diferentes concepções de autonomia. De qualquer maneira, uma etapa fundamental desta dissertação é expor a sua concepção de autonomia. Afinal, boa parte dos argumentos filosóficos contrários e favoráveis a diferentes

[152] Agradeço ao Prof. Dr. Brunello de Souza Stancioli pelos comentários que mudaram os rumos desta seção.

[153] Cf., *e.g.*, (REBONATO, 2014; WALDRON, 2014). De maneira polêmica, Sunstein (2013b) argumenta que a deontologia poderia ser considerada uma heurística moral. Segundo o autor, o julgamento deontológico é um atalho mental para a análise moral correta, que seria consequencialista. Sunstein reconhece, contudo, que as evidências nas pesquisas que corroboram esse entendimento não são suficientes para se estabelecer que julgamentos deontológicos são errados ou falsos.

graus de regulação do comportamento privado que serão discutidos neste trabalho dependem dessa concepção.

Antes de apresentar essa concepção, duas ressalvas são necessárias. Em primeiro lugar, um conceito de autonomia deve ser empiricamente possível (DWORKIN, 1988, p. 7–8). Algumas concepções de autonomia, no entanto, são excessivamente exigentes em um universo no qual as pessoas têm racionalidade limitada (ENGELEN; NYS, 2020). Se, por exemplo, a autonomia significar que o sujeito deve determinar suas decisões sem ser influenciado por elementos alheios ao seu controle ou que as decisões devam ser tomadas pelos motivos corretos, a maior parte das decisões cotidianas seriam heterônomas. Em segundo lugar, entende-se que a autonomia não deve ser analisada em uma perspectiva puramente instrumental. Ganhos de bem-estar são, certamente, importantes, mas as pessoas atribuem um valor intrínseco ao direito de escolher os rumos que pretendem seguir[154]. Esse valor deve ser considerado em qualquer análise que tenha a autonomia por objeto.

Feitas essas ressalvas, é necessário mencionar, novamente com base em Dworkin (1988, p. 10-12), que a autonomia pode funcionar como um ideal moral, social ou político. Como um ideal moral, a ênfase da autonomia recai sobre a maneira como o indivíduo forma suas opiniões quanto ao mérito de restrições legais. Essa noção diz respeito à necessidade ou à desejabilidade de que indivíduos escolham ou aceitem seu próprio código moral. Como um ideal social, por sua vez, a autonomia representa um conjunto de preocupações ligado à maneira pela qual instituições não políticas afetam os valores, as atitudes e as crenças dos membros da sociedade. Por fim, a autonomia enquanto ideal político é a base para se questionar a elaboração e o funcionamento de instituições políticas que tentam impor um conjunto de fins, valores e atitudes sobre as pessoas. O interesse desta dissertação recai sobre a autonomia política.

Um primeiro ponto relevante para a definição do conceito de autonomia é a distinção entre a capacidade de agir de maneira autônoma e o agir de maneira autônoma. Reconhecer que um sujeito é autônomo significa simplesmente dizer que ele tem a capacidade de agir de forma autônoma. No entanto, ter essa capacidade não é garantia de que uma escolha autônoma será feita efetivamente (FADEN; BEAUCHAMP, 1986, p. 237). Nesse caso, como o interesse deste trabalho é avaliar a legitimidade de políticas com impacto em atos específicos das pessoas, parece mais adequado e empiricamente realizável discutir a autonomia de atos individuais. Em

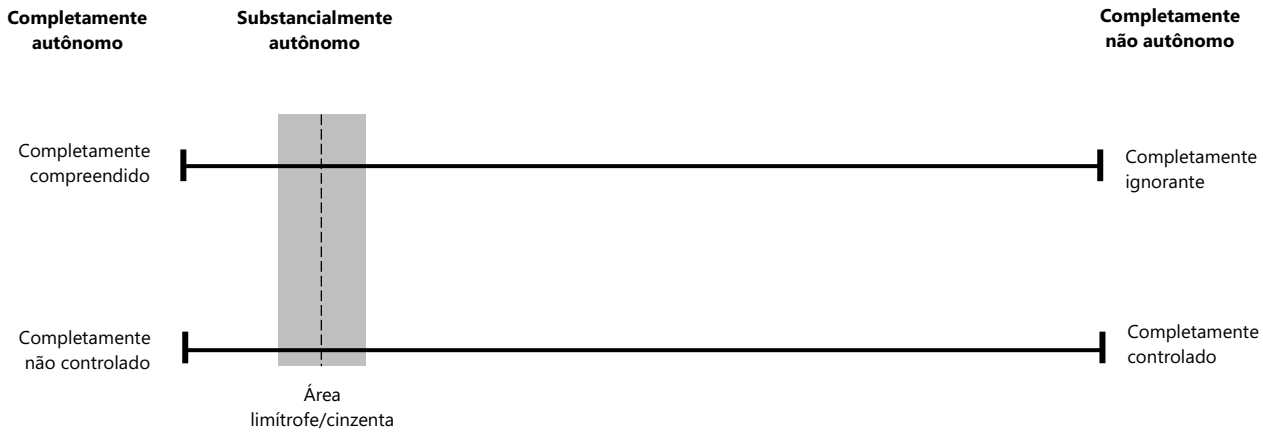
[154] Björn Bartling, Ernst Fehr e Holger Herz (2014), por exemplo, avaliaram se direitos de decisão são instrumentais ou apresentam um valor intrínseco. Em seu experimento, os autores identificaram que boa parte das pessoas atribui um valor intrínseco positivo aos direitos de escolher.

outras palavras, desloca-se o questionamento "do agente autônomo para o agir autônomo. Desvincula-se da busca do sujeito e centra-se na análise do ato (STANCIOLI, 2004, p. 34)".

Nessa perspectiva, Ruth R. Faden e Tom L. Beauchamp (1986) defendem que três critérios devem ser utilizados para avaliar se um ato é autônomo: (i) a compreensão; (ii) a ausência de influências controladoras; e (iii) a intencionalidade. Com base nesses critérios, um ato autônomo é uma ação intencional, completamente compreendida e completamente livre de influências controladoras de terceiros (FADEN; BEAUCHAMP, 1986, p. 237).

À medida que algum desses elementos deixa de prevalecer, um ato se torna gradualmente menos autônomo. No limite, a depender do nível em que um dos critérios deixa de ser observado, o comportamento deixa de ser autônomo. Faden e Beauchamp não buscam a definição de um ato completamente autônomo - cenário que dificilmente seria empiricamente realizável. Em vez disso, o objetivo é identificar atos que sejam substancialmente autônomos, limite a partir do qual se pode inferir que um ato efetivamente reflete a autonomia do indivíduo (FADEN; BEAUCHAMP, 1986, p. 241). Há, portanto, espaço para diferentes graus de autonomia ao se avaliar uma ação intencional, como demonstrado no esquema abaixo.

Figura 4 - Graus de autonomia de ações intencionais



Fonte: Faden e Beauchamp (1986, p. 239)

Apresentada essa estrutura geral, é necessário definir, então, o que significam os critérios utilizados por Faden e Beauchamp para definir um ato autônomo; ou seja, o que significam a compreensão, o não controle e a intencionalidade?

A princípio, Faden e Beauchamp (1986, p. 248) consideram que uma ação não pode ser autônoma se o sujeito não consegue ter adequada compreensão de sua ação. Embora não haja um consenso na literatura da filosofia e da psicologia quanto à definição de compreensão, Faden

e Beauchamp (1986, p. 252) consideram que um sujeito tem completa compreensão de uma ação se tiver um entendimento completamente adequado de todas as afirmações relevantes e que descrevam: a) a natureza da ação; e b) as consequências e resultados previsíveis que possam se originar de sua ação ou omissão. À medida em que esses requisitos não forem completamente satisfeitos, o sujeito não terá uma compreensão completa e, logo, o ato analisado também não será completamente autônomo. Há, portanto, espaço para diferentes graus de compreensão de uma ação e haverá autonomia se um ato for substancialmente compreendido pelo sujeito.

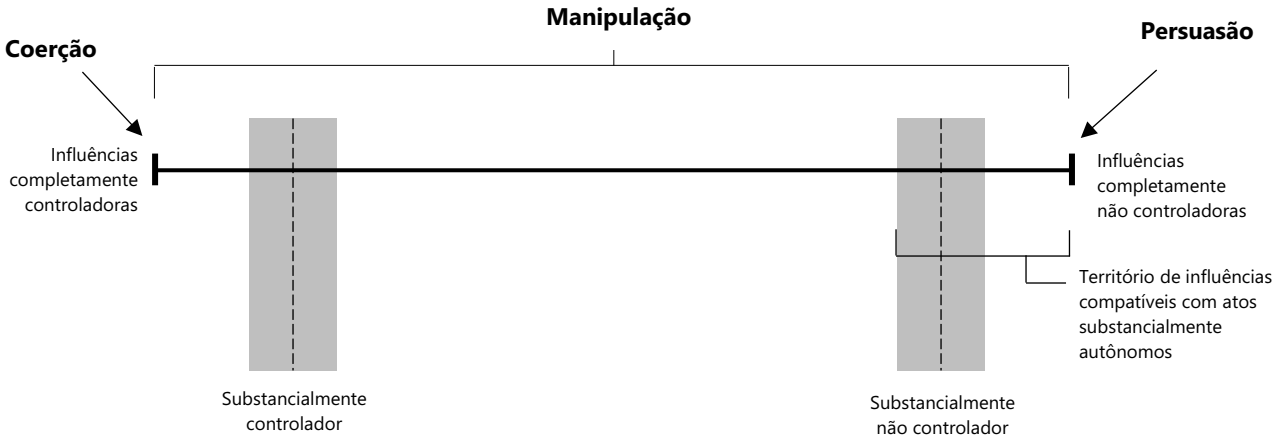
Além disso, para serem autônomas, as pessoas devem estar livres do controle, isto é, devem agir independentemente de influências de terceiros capazes de surrupiar a sua auto-orientação (FADEN; BEAUCHAMP, 1986, p. 256). Para Faden e Beauchamp (1986, p. 258), um ato é completamente não controlado se não tiver sido alvo de uma tentativa de influência, ou, caso tenha sido, se a tentativa for malsucedida ou incapaz de privar o agente de fazer ou acreditar no que deseja. O controle do sujeito, no entanto, pode ser reduzido a depender do grau de influência que lhe é imposto.

Nem toda influência, no entanto, é igualmente controladora (FADEN; BEAUCHAMP, 1986, p. 258-260). Em um extremo, há influências que controlam completamente o sujeito: são medidas coercitivas que sempre comprometem a autonomia, já que efetivamente controlam o rumo de ação adotado pelo agente. É o caso, por exemplo, do assaltante que coloca uma arma à cabeça do sujeito e o obriga a fazer uma transferência bancária. No extremo oposto, a persuasão pode ser considerada um tipo de influência que, por definição, é não controladora. O médico que recomenda fortemente que o paciente utilize um determinado medicamento o persuade, sem, contudo, controlar seus rumos de ação.

Por fim, há tipos de intervenção que admitem graus diferentes de controle: as manipulações. Se o sujeito é enganado, seduzido ou doutrinado por um terceiro, haverá influências que podem ser tão controladoras quanto a coerção, mas também podem ser não controladoras, como a persuasão. Trata-se de uma categoria ampla que inclui qualquer influência intencional e bem-sucedida de alterar as escolhas disponíveis para o sujeito de maneira não coercitiva ou alterar as suas percepções sobre essas escolhas de maneira não persuasiva (FADEN; BEAUCHAMP, 1986, p. 261). Se, por exemplo, um governo resolve apresentar à população o número absoluto de pacientes curados por uma doença com alto percentual de cura, em vez de indicar a progressão de novos casos e mortes, há uma tentativa velada de influenciar as percepções da população, sem se valer de métodos coercitivos ou persuasivos.

O *continuum* de influências controladoras e não controladoras pode ser dividido no seguinte esquema, sendo que influências no mínimo substancialmente controladoras comprometem a autonomia, enquanto influências substancialmente não controladoras permitem que atos sejam autônomos:

Figura 5 - O continuum de influências de controladoras a não controladoras



Fonte: Faden e Beauchamp (1986, p. 243)

O último critério para avaliar a autonomia de um ato específico é a intencionalidade. Para Faden e Beauchamp, esse critério representa a intenção específica de executar uma determinada ação; não se trata de ações que simplesmente acontecem por acidente, inadvertidamente ou, mesmo, por equívoco. Na verdade, uma ação intencional é uma "ação pretendida em conformidade com um plano, quer o ato seja desejado ou não" (FADEN; BEAUCHAMP, 1986, p. 243).[155] Em resumo, atos intencionais exigem planejamento: para um ato ser intencional, ele deve corresponder à concepção do agente do que representa aquele ato, ainda que as suas consequências não se materializem (FADEN; BEAUCHAMP, 1986, p. 242). Nesse caso, seguindo os seus próprios planos, os agentes podem obter consequências indesejadas, como a cicatriz que sabidamente será o resultado de uma cirurgia feita para curar um câncer. Como o resultado indesejado é parte do plano, para Faden e Beauchamp (1986, p. 245–247), também esse ato - os resultados tolerados, como a cicatriz - é intencional.

Além disso, Faden e Beauchamp (1986, p. 247-248) entendem que atos intencionais não se encontram num *continuum*, já que não existiriam graus de ação intencional: ou se está diante de um ato intencional, ou não. Faden e Beauchamp até especulam razões que poderiam levá-los a avaliar diferentes graus de intencionalidade, mas acabam por rejeitá-las. A princípio,

[155] Tradução livre. Texto original: "[...] *action willed in accordance with a plan, whether the act is wanted or not*".

alguns atos poderiam ser mais intencionais com base na atenção dada pela vontade do agente à ação específica; ou seja, as ações seriam menos intencionais em algum grau se fossem executadas automaticamente, com menos consciência cognitiva. Outra possibilidade seria diferenciar níveis de intencionalidade com base na correspondência entre uma ação e o plano de ação que representa o que o agente realmente gostaria de fazer e como ele gostaria de efetivar essa intenção.

Faden e Beauchamp rejeitam ambas as propostas. Em sua visão, mesmo atos comparativamente não deliberados, como ativar um interruptor de luz, ainda podem ser tão intencionais quanto o consentimento dado para uma cirurgia. De todo modo, Faden e Beauchamp sustentam ser improvável que consentimentos ou recusas deixem de ser autônomos por não seguirem os dois critérios apresentados acima, que justificariam a proposta de se considerar diferentes graus de intencionalidade.

Essa última visão, no entanto, não parece estar completamente alinhada com o marco teórico adotado nesta dissertação. Além de ter sido justificada com argumentos que não parecem muito convincentes - afinal, por que ativar um interruptor de luz é tão intencional quanto o consentimento dado a uma cirurgia? -, Faden e Beauchamp parecem não dar a devida importância à influência de fatores supostamente irrelevantes nos planos e nas ações implementadas pelos agentes. Por três motivos diversos, parece viável também tratar a intencionalidade como um *continuum* entre atos mais ou menos intencionais.

Inicialmente, além de ações - como os consentimentos ou recusas que Faden e Beauchamp abordam em sua obra dedicada à teoria do consentimento informado –, *omissões* também podem ser bastante relevantes. Quando, por exemplo, um sujeito deixa de se posicionar ativamente para contribuir em maior proporção com um determinado plano de previdência, a sua omissão é intencional? Os planos do sujeito, em um sentido mais amplo, poderiam até envolver poupar mais recursos para o futuro; pode ser, no entanto, que, por falta de atenção ou pelo viés para o presente, o sujeito deixa de agir. Em outras palavras, o ato de não contribuir - a sua omissão – pode não ser intencional e, por consequência, pode não ser suficientemente autônomo.

Parece, portanto, ser possível afirmar que omissões - ainda que inadvertidas - também possam ser consideradas mais ou menos autônomas. No entanto, por mais que uma omissão possa ser tida como intencional, nitidamente, ela não tem o mesmo grau de intencionalidade que uma ação ativa. É menos provável que o sujeito trace um plano com relação às suas omissões e há uma maior chance de que ele seja influenciado por vieses que o vinculam a determinada opção de escolha. Por essa razão, por mais que não pareça razoável impedir que

omissões sejam consideradas autônomas, também parece ser válido ressaltar que elas não têm o mesmo grau de intencionalidade que uma ação ativa. Esse é o primeiro argumento que leva a se pensar a intencionalidade enquanto um *continuum* nesta dissertação.

Além disso, como Faden e Beauchamp[156] até chegam a admitir, alguns atos são praticados automaticamente - são baseados no Sistema 1 -, enquanto outros são mais lentos e envolvem processos de deliberação e raciocínio - envolvem processos ligados ao Sistema 2. Considerando essa distinção, não parece razoável considerar que um ato automático, como ser levado a adquirir impulsivamente um doce na fila do supermercado, tem o mesmo grau de intencionalidade que uma decisão extremamente deliberada, como a escolha de um apartamento para morar em uma nova cidade. Na primeira decisão, o sujeito pode sequer ter pensado a respeito da escolha que está fazendo ou, até mesmo, pode ter sido manipulado a fazê-la[157]. A segunda, por sua vez, envolve análise de custos e benefícios, alta deliberação e raciocínio a longo prazo. Parece razoável, portanto, argumentar que essas decisões têm níveis de intencionalidade diversos, a depender do grau em que elas são deliberativas. Esse é o segundo motivo que justifica a análise da intencionalidade dentro de um *continuum*.

Por fim, como, mais uma vez, Faden e Beauchamp[158] reconhecem, alguns atos são mais correspondentes aos verdadeiros planos do sujeito do que outros. Em outros termos, pode ser que a preferência revelada pela escolha do agente não corresponda às preferências de segunda ordem do sujeito, isto é, à sua capacidade de refletir e adotar com relação aos seus desejos, intenções e aspirações de primeira ordem (DWORKIN, 1988, p. 15). Afinal, além de ter preferências imediatas, humanos têm preferências sobre essas preferências[159]. Em outras

---

[156] Ao discutir um dos argumentos para considerar a intencionalidade como um *continuum*, Faden e Beauchamp (1986, p. 247-248) mencionam que, enquanto alguns atos são executados automaticamente, outros exigem um verdadeiro esforço da vontade e podem envolver planejamento deliberativo, monitoramento cuidadoso e, talvez, até mesmo notas escritas para que o ato tenha sucesso. O exemplo dado por Faden e Beauchamp mostra como essas distinções se aproximam das noções de Sistema 1 e Sistema 2, já expostas no Capítulo 1. Ao se aprender a dirigir, os atos necessários para se aprender a controlar e reagir exigem concentração, deliberação e consciência interna ativa. Com o tempo, ao se tornar um motorista experiente, os mesmos atos são executados sem pensar.

[157] O grau de manipulação a que o sujeito está submetido, em linha com o que já foi discutido acima, impacta, contudo, no elemento da ausência de controle sobre a ação. O que se argumenta, aqui, é simplesmente que decisões automáticas são menos intencionais que decisões deliberadas, independentemente do grau de influência que recai sobre essas decisões automáticas.

[158] Faden e Beauchamp (1986, p. 248) mencionam que o grau de correspondência entre uma ação realizada e o plano de ação depende não apenas do número de discrepâncias para o plano, mas também da importância de cada discrepância. No limite, uma ação pode se desviar tanto dos planos de ação que o ato pode quase não ser reconhecido como aquilo que o sujeito realmente gostaria de fazer. Esse poderia ser um critério para diferenciar a intencionalidade de duas ações distintas.

[159] Harry Frankfurt considera os desejos de segunda ordem uma característica peculiar dos humanos. Como afirma o autor: "Parece ser peculiarmente característico dos humanos, no entanto, que eles sejam capazes de formar o que eu devo chamar de 'desejos de segunda-ordem'. Além de querer e escolher e ser movido a isso ou aquilo, os homens podem também querer ter (ou não ter) certos desejos e motivos. Eles são capazes de querer ser diferentes, em suas preferências e propósitos, do que eles são. […]" (FRANKFURT, 1971, p. 7, tradução livre).

palavras, ao mesmo tempo em que o sujeito pode preferir tomar uma garrafa de cerveja na segunda-feira após o trabalho (preferência de primeira-ordem), ele pode desejar não ter essa preferência (preferência de segunda ordem).

Frequentemente, há um descompasso entre as preferências de primeira ordem e as preferências de segunda ordem. Problemas de autocontrole são uma das razões claras para esse descompasso. O sujeito pode, por exemplo, desejar parar de fumar e, ao mesmo tempo, sentir vontade de fumar um cigarro e acabar por implementar essa ação. Nesse contexto, parece razoável considerar que ações que se conformam às preferências de segunda ordem do indivíduo são mais intencionais do que ações que se distanciam dessas preferências.

A ideia, aqui, aproxima-se da noção de autenticidade[160,161]. Faden e Beauchamp (1986, p. 263-268) até chegam a avaliar se a autenticidade deve ser uma condição necessária à autonomia, mas rejeitam esse critério, por entender que as tentativas de incluí-lo reduziriam demais o escopo da autonomia ou criariam dificuldades desnecessárias.[162] Nesse caso, Faden e Beauchamp consideram mais promissora a ideia de incluir a autenticidade como elemento do critério do não controle. Em outros termos, exige-se que a ação autônoma seja livre do controle de terceiros e, ao mesmo tempo, independente do controle de compulsões neuróticas, adições e outros desordens psiquiátricas.

Essa não parece ser a abordagem mais promissora. Independentemente de desordens psiquiátricas, humanos, em geral, são influenciados o tempo todo por fatores supostamente

---

[160] Charles Taylor (1991, p. 28-29), um dos mais célebres filósofos que trabalha com a noção de autenticidade, menciona que ter um tipo de contato com si próprio e com a própria natureza interna é um ideal moral crucial. Se o sujeito não age de maneira verdadeira a si próprio, ele perderia o próprio significado do que é ser humano para ele. E ser verdadeiro a si próprio significa respeitar a sua própria originalidade, que apenas o sujeito pode articular e descobrir. Nessa articulação, o sujeito se define. Embora essa não seja propriamente a visão adotada nesta dissertação, a teoria de Taylor é relevante e deve ser mencionada.

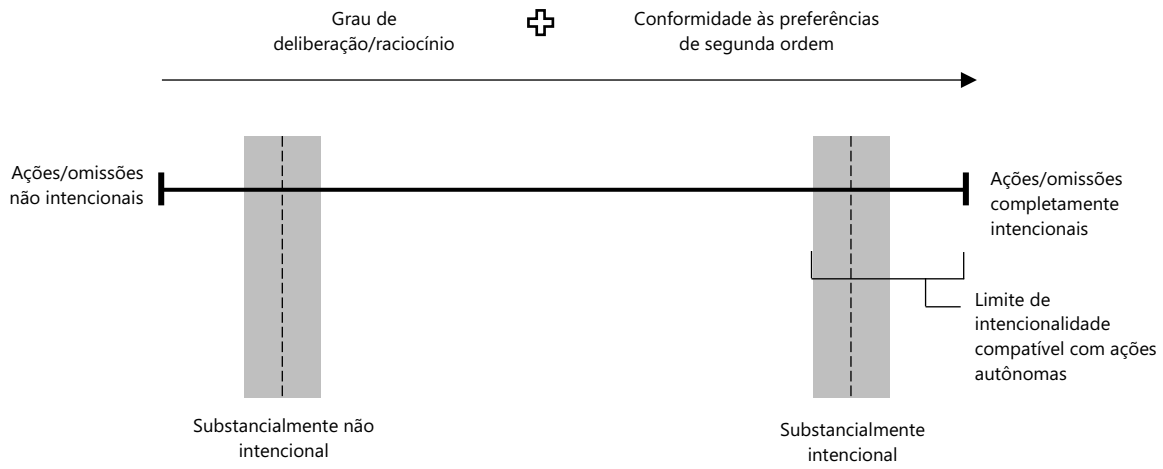
[161] A influente teoria da autonomia desenvolvida por Dworkin é amplamente baseada na autenticidade. Para Dworkin (1988, p. 20), a autonomia é concebida como uma capacidade de segunda ordem pela qual as pessoas refletem de maneira crítica sobre as suas preferências de primeira ordem, desejos e aspirações, e a capacidade de aceitar ou tentar mudar estes desejos diante das preferências de mais alta ordem e valores. É pelo exercício de tal capacidade que as pessoas definem sua natureza e dão significado à própria vida.

[162] Inicialmente, Faden e Beauchamp (1986, p. 263-268) consideram a possibilidade de incluir uma condição de autenticidade à autonomia, que exigiria que, para serem autônomas, as ações fossem consistentes com os valores e comportamentos reflexivamente aceitos pelos sujeitos. Em sua visão, porém, um componente de autenticidade nesses termos restringiria demais o escopo de proteção de inúmeras ações vistas como autônomas, o que diminuiria a importância do princípio de respeito à autonomia. Em um segundo sentido, a condição da autenticidade poderia ser reformulada, para se considerar, como condição da ação autônoma, a estabilidade ou a consistência nos valores subjacentes à escolha ou, em outra perspectiva, o não repúdio a esses valores. Também nestes casos, a inclusão do componente de autenticidade é rejeitada por Faden e Beauchamp. Com relação à exigência de estabilidade e consistência, os autores ponderam que as pessoas frequentemente agem de maneira contrária aos seus valores estáveis e alguns desses atos, são autônomos. Embora mais promissor, o não repúdio também geraria problemas, especialmente pela dificuldade de se conhecer o que realmente é ou pertence ao self do sujeito e de se explicar corretamente o comportamento humano em termos de causas ou motivos subjacentes. Por exemplo, o sujeito que é viciado em heroína deve ser considerado autônomo a não ser que repudie o seu comportamento? Diante de tais dificuldades, Faden e Beauchamp rejeitam a autenticidade enquanto critério para que um ato seja autônomo.

irrelevantes. Essa influência faz com que a ação de pessoas sadias também se distancie de suas preferências de segunda ordem. Tal distanciamento, no entanto, não significa, necessariamente, que há maior controle sobre essas ações. O sujeito pode, simplesmente, agir de maneira inadvertida e, por impulso, comportar-se de maneira contrária aos valores que ele almejaria seguir. Não há maior controle, mas há claro afastamento daquelas que seriam as suas reais intenções.[163] Se a intencionalidade é agir com base em planos prévios, parece decorrer de tal conceito a diminuição do grau de intencionalidade caso a ação praticada pelo sujeito se distancie desses planos, isto é, de suas preferências de segunda ordem. Também por essa razão, opta-se, nesta dissertação, por analisar a intencionalidade como um *continuum*.

Os diferentes graus de intencionalidade encontram-se refletidos no esquema abaixo:

Figura 6 - O *continuum* da intencionalidade



Fonte: Elaboração do Autor, com inspiração nos esquemas apresentados por Faden e Beauchamp (1986)

[163] Kant (1997, p. 363), um dos mais célebres filósofos que adotam a autenticidade como critério definidor da autonomia, menciona que, embora as pessoas tenham poderes para direcionar a mente, normalmente há empecilhos que criam resistência à força e à livre escolha das pessoas enquanto sujeitos pensantes. Por exemplo, a preguiça e a satisfação sensual (*sensual indulgence*). Esses elementos precisam ser dominados (*mastered*) e não apenas direcionados. A autoridade para compelir a mente, apesar desses impedimentos, é o que Kant chama de *autocracia*. Em interpretação dessa teoria, Paul Guyer (2003) menciona que, nas condições empíricas da vida, Kant não sugere que a autonomia seria realizada pela completa eliminação das inclinações. Na verdade, para Kant, a autonomia seria posta em prática pelo desenvolvimento da autocracia, também conhecida como autodomínio (*self-mastery*). Do mesmo modo, como mencionado por Bart Engelen e Thomas Nys (2020), determinar os fins que são importantes para si - conceito de autonomia utilizado pelos autores – é diferente de realmente motivar-se de acordo com esses fins para realizá-los (autocracia). Embora se adote outro conceito de autonomia nesta dissertação, com enfoque em atos autônomos em vez de sujeitos autônomos, esse raciocínio pode ser utilizado para explicar como as pessoas podem agir de maneira contrária às suas preferências de segunda ordem.

2.5. Principais conclusões do Capítulo 2

O objetivo deste Capítulo foi expor a maneira pela qual a identificação de fatores supostamente irrelevantes levou ao desenvolvimento de novos fundamentos para a regulação do comportamento privado. Inicialmente, foram apresentadas as falhas de mercado tradicionais, que fundamentam intervenções do Estado de acordo com os pressupostos da economia neoclássica. Em seguida, foram discutidas as falhas de mercado comportamentais, os diferentes tipos de paternalismo comportamental, assim como a noção de autonomia adotada neste trabalho. As principais conclusões deste Capítulo foram as seguintes:

1. O mapeamento de inúmeros fatores supostamente irrelevantes que influenciam as decisões exige o ampliamento da noção de falha de mercado, a fim de identificar razões para a regulação do comportamento privado de caráter comportamental. Como exemplos, foram apresentados as internalidades, o otimismo excessivo, o impacto de atributos ocultos na tomada de decisões e problemas com estatística. Políticas efetivas envolvem uma combinação de tais *insights* comportamentais com a análise econômica tradicional;
2. Uma política deve ser considerada paternalista se incluir esforços para proteger as pessoas contra seus próprios erros e guiá-las às escolhas que fariam se fossem completamente informadas e livres de vieses comportamentais. Paternalismos suaves envolvem a imposição de baixos custos aos decisores ou intervêm em escolhas que não são propriamente voluntárias. Ao mesmo tempo, paternalismos fracos tentam influenciar apenas os meios escolhidos pelos sujeitos para atingir seus próprios fins;
3. As falhas de mercado comportamentais dão origem a mais oportunidades para se prescrever políticas paternalistas. Essa constatação levou ao desenvolvimento de uma série de paternalismos comportamentais. Em uma primeira perspectiva, foram desenvolvidos paternalismos suaves, como o paternalismo libertário, o paternalismo assimétrico e o paternalismo pragmático leve. De outro lado, também há propostas de paternalismos coercitivos, inspiradas por insights comportamentais;
4. Um sujeito que tem a capacidade de agir de maneira autônoma pode tomar decisões não autônomas. Por isso, a análise da autonomia de atos específicos parece ser mais útil. Nesse caso, com inspiração em Faden e Beauchamp, os seguintes critérios foram adotados para definir se um ato é autônomo: (i) a compreensão; (ii) a ausência de influências controladoras; e (iii) a intencionalidade. Todos esses critérios podem estar

> presentes em maior ou menor grau na avaliação de um ato específico. Embora a autonomia completa não seja empiricamente factível, níveis suficientes de compreensão, ausência de controle e intencionalidade podem levar a atos substancialmente autônomos.

Estabelecidas essas premissas, é necessário ressaltar que, além de novos fundamentos para a intervenção na vida privada, a identificação de fatores supostamente irrelevantes também originou novas ferramentas de intervenção no arcabouço regulatório. Os *nudges* são as mais célebres delas, mas também há outras como *boosts*, *thinks*, *nudges plus* e *sludges*. Há, ainda, considerável ruído na definição de tais conceitos na literatura. Os próximos dois Capítulos apresentarão esses conceitos, assim como o atual estado da literatura a seu respeito.

3. A ELABORAÇÃO E O DESENVOLVIMENTO DA DEFINIÇÃO DE *NUDGE*: PROPOSIÇÃO ORIGINAL E REFINAMENTOS CONCEITUAIS

Um sujeito entra em uma livraria de rua e decide escolher um livro. As estantes e os mostradores podem ser organizados de inúmeras formas. Seguro dizer que os livros que mais ficam à mostra serão escolhidos com maior probabilidade, como demonstra a exposição dos inúmeros fatores supostamente irrelevantes que impactam decisões feitas ao longo dos Capítulos 1 e 2. Imagine, então, que um governo obrigue as livrarias a expor os livros escritos por João Guimarães Rosa em suas prateleiras de destaque. O objetivo dessa política – uma imposição às livrarias – é direcionar o comportamento dos consumidores às opções que lhes seriam mais aprazíveis, mas que, por fatores supostamente irrelevantes, não estavam sendo escolhidas. Segundo o governo que propôs essa política, tal medida teria por objetivo aumentar o bem-estar dos indivíduos submetidos a ela – isto é, os consumidores –, sem coerção ou mudança substancial dos incentivos envolvidos na escolha. Em outras palavras, seria uma espécie dos paternalismos suaves já expostos no Capítulo 2.

Independentemente da legitimidade de tal intervenção, no fim do dia, o livreiro deve tomar alguma decisão sobre a forma como os livros serão organizados - seja essa decisão determinada pelo governo, seja ela uma decisão da própria livraria. Essa decisão, no entanto, influenciará as futuras escolhas dos consumidores. Em todas as situações em que as pessoas tomam decisões, a arquitetura de escolhas, isto é, a forma como é organizado o contexto social em que as pessoas decidem, influencia as opções que são efetivamente escolhidas. Em tais contextos, instituições especificam quando, se e como as pessoas devem escolher (SUNSTEIN, 2015a, p. 5)[164].

A partir da reorganização da arquitetura de escolhas, diversos pesquisadores passaram a utilizar a observação de traços do comportamento humano para influenciar a elaboração de políticas públicas. Em outras palavras, se a identificação de falhas de mercado comportamentais serviu como base para a elaboração de novos fundamentos à regulação do comportamento privado, esses mesmos fatores comportamentais originaram novas ferramentas para a intervenção estatal. Afinal, ao se considerar uma determinada justificativa para uma intervenção do governo, a questão subsequente é definir como o governo deve intervir (BUBB; PILDES, 2014, p. 1602). Nesse sentido, o trabalho de Thaler e Sunstein, em especial, é notável também por demonstrar sistematicamente a possibilidade de se direcionar escolhas, sem métodos

[164] Thaler e Sunstein (2009, p. 3) definem esses agentes como arquitetos de escolhas, isto é, indivíduos que têm a responsabilidade de organizar o contexto em que as pessoas tomam decisões.

coercitivos ou mudanças na estrutura de incentivos. A essa estratégia, Thaler e Sunstein deram o nome de *nudge* - traduzido, nesta dissertação, como estímulo comportamental, como já mencionado na Introdução.

No modelo econômico neoclássico, intervenções estatais normalmente tomam a forma de medidas de comando e controle. De fato, como apontam Bubb e Pildes (2014, p. 1602), o arcabouço neoclássico de intervenção regulatória apresenta recomendações rigorosas de políticas públicas. Para controlar externalidades negativas, por exemplo, como já mencionado no Capítulo 2, um conjunto de medidas ligadas a impostos, regras de responsabilidade e a definição de direitos de propriedade costuma ser empregado. Já com relação à assimetria de informações, mandados de transparência com relação às informações (*disclosure mandates*) ou a própria produção e disseminação de informações pelo próprio governo podem ser ferramentas úteis. Monopólios naturais, por sua vez, são controlados com a regulação de tarifas, o controle de acesso ou, até mesmo, a propriedade dos ativos pelo próprio governo.

Nesse contexto, entender o surgimento dos *nudges* é um elemento-chave para a compreensão do surgimento predominante de paternalismos suaves em meio aos paternalismos comportamentais expostos no Capítulo 2. Afinal, se as ferramentas neoclássicas se baseavam, em geral, na coerção, a manipulação da arquitetura de escolhas acrescenta um novo elemento ao cardápio de ferramentas regulatórias disponíveis ao regulador. Em vez de coagir, na forma de obrigações diretas ou indiretas, cujo descumprimento levaria a sanções juridicamente executáveis, *nudges* abrem maior espaço para paternalismos suaves. O regulador pode tentar intervir sem criar custos significativos àqueles que não querem se vincular à intervenção.

De todo modo, antes de enfrentar essas críticas, uma análise mais detida a respeito da definição de *nudge* parece ser necessária. Quando se tem um primeiro contato com a obra de Thaler e Sunstein, supor que *nudges* podem ser antiéticos, manipulativos, ilegais ou, até mesmo, inconstitucionais, pode parecer algo radical; talvez, até mesmo, oriunda de uma posição excessivamente libertária. Afinal, boa parte das medidas propostas por Thaler e Sunstein são realmente muito brandas: simplificação de requisitos burocráticos para procedimentos como pedidos de financiamento estudantil; lembretes via e-mail ou mensagem de texto; evocação de intenções de implementar uma determinada opção; divulgação de informações nutricionais de um alimento ou de informações de consumo de energia elétrica etc. (SUNSTEIN, 2019c).

Em ordenamentos em que medidas como a exigência de capacetes para pilotar motocicletas ou cintos de segurança para dirigir automóveis são acatadas sem muita discussão, questionar a legitimidade de medidas menos interventivas como os *nudges* mencionados pode, de fato, parecer uma discussão supérflua. Apesar disso, embora essas medidas sejam relevantes

e tenham efetividade, sobretudo se considerados os baixos custos para implementá-las, como será demonstrado neste Capítulo, elas estão distantes do que a maior parte dos críticos têm em mente, quando criticam a utilização de *nudges*.

Em geral, o conceito adotado pelos críticos é mais restrito e costuma se limitar a medidas que exploram vieses comportamentais - ou, na terminologia predominantemente adotada nesta dissertação, fatores supostamente irrelevantes - e que acabam, assim, por frequentemente passar despercebidas. A exploração dessas falhas na tomada de decisão é um elemento central da maior parte das críticas. Como será demonstrado, no entanto, Thaler e Sunstein não veem esse elemento como condição suficiente para que uma intervenção seja um *nudge*.

Para apresentar essa discussão, este Capítulo pretende mapear três momentos distintos da literatura da análise econômico-comportamental do direito (*behavioral law and economics*). O primeiro deles se confunde com o início da elaboração dos paternalismos comportamentais e o surgimento de técnicas de intervenção não coercitivas, baseadas em *insights* comportamentais. Em seguida, será apresentada a elaboração original do conceito de *nudge* e a amplificação do seu escopo, especialmente em anos mais recentes. O terceiro momento, por sua vez, terá por objetivo apresentar refinamentos conceituais elaborados na literatura, que originaram ao menos duas definições estritas de *nudges*, além de classificações em diferentes categorias segundo os mecanismos psicológicos explorados pela intervenção ou a técnica de influência empregada. Feita essa exposição, este Capítulo pretende justificar a opção metodológica desta dissertação por uma definição estrita de *nudge*, ainda que ela não seja a definição adotada por Thaler e Sunstein, os proponentes do conceito.

3.1. *Nudges* antes dos *nudges*: diferentes técnicas de intervenção que exploram elementos do comportamento humano

Thaler e Sunstein foram os responsáveis por cunhar o termo *nudge* para representar todo um conjunto de ferramentas regulatórias que, de maneira geral, afastavam-se de medidas coercitivas e impositivas. Apesar disso, antes de tal desenvolvimento teórico, políticas públicas baseadas em *insights* comportamentais já existiam. Antes mesmo da publicação de *Nudge*, em que o conceito foi apresentado, ao propor a incorporação de insights comportamentais à análise econômica do direito e, depois, ao elaborar o paternalismo libertário, Thaler e Sunstein já

discutiam políticas que podem ser consideradas *nudges*[165]. Na mesma época, diversos outros autores também já propunham medidas de intervenção baseadas em insights comportamentais.

Como já mencionado no Capítulo 2, os paternalismos comportamentais surgem a partir do mapeamento de inúmeras falhas de mercado comportamentais. Nada mais natural que, para enfrentá-las, os mecanismos comportamentais que lhes são subjacentes sejam levados em consideração para elaborar a técnica de intervenção específica. Nesse contexto, junto de propostas de novos fundamentos para se regular o comportamento privado, já eram apresentadas técnicas de regulação que se distanciavam do tradicional arcabouço regulatório neoclássico. Boa parte dessas medidas inclui-se no que, hoje, é reconhecido como um *nudge*. Como será demonstrado abaixo, há, de fato, inúmeros exemplos que evidenciam políticas como alteração de regras *default*, mudança de *framings*, criação de pontos de decisão, alteração de pontos de partida, exigência de períodos de reflexão (*cooling-off periods*) etc[166].

Nesta seção, pretende-se demonstrar como exatamente essas políticas foram descritas na literatura no contexto anterior à elaboração do conceito de *nudge*, a fim de apreender quais seriam os seus principais traços comuns. A seleção utilizada nesta subseção é meramente exemplificativa e pretende apenas expor, em linhas gerais, como essas políticas eram descritas em alguns trabalhos bastante relevantes desse período. Não há, portanto, uma intenção de fazer uma revisão bibliográfica exaustiva da literatura da época.

Ao começar a abordar a interação entre o direito e as ciências comportamentais, ainda de forma incipiente, Sunstein (1997) pontuou que a incorporação de *insights* comportamentais ao direito poderia impactar a análise prescritiva, positiva e normativa do direito.[167] Do ponto de vista prescritivo, discutido neste Capítulo, já naquela oportunidade, Sunstein apontava estratégias baseadas na manipulação do contexto em que as escolhas são feitas para direcionar o comportamento. Essas intervenções baseavam-se explicitamente na exploração de fatores

---

[165] Em um *tweet* recente, Sunstein (2022) mencionou que o texto *Libertarian Paternalism is not an oxymoron* (THALER; SUNSTEIN, 2003b), abordado em mais detalhes abaixo, é algo como *nudges* antes de *nudges* (*nudges before nudges*), corroborando a afirmação feita nesta dissertação.

[166] Boa parte dessas políticas, inclusive, já era proposta antes da sistematização dos *fatores supostamente irrelevantes*. A experimentação, afinal, pode evidenciar práticas que funcionam, ainda que não se tenha propriamente teorizado a respeito dos motivos que explicam o seu sucesso. Akerlof e Shiller (2015), por exemplo, demonstram que, na história da publicidade, publicitários sempre exploraram elementos do comportamento humano e, inclusive, desenvolveram técnicas de estatística para medir a eficácia de suas intervenções. Ainda que esses publicitários não tivessem um conhecimento profundo sobre a forma como humanos reagiriam a intervenções, eles conseguiram, a partir da tentativa e erro, definir sistematicamente o que fazia as pessoas comprarem os produtos a serem vendidos.

[167] A análise positiva se preocupa com previsões a respeito de qual será o comportamento das pessoas. Do ponto de vista prescritivo, pretende-se demonstrar como a sociedade pode atingir os seus objetivos compartilhados. A análise normativa, por fim, destina-se a avaliar o que um sistema jurídico deve fazer (SUNSTEIN, 1997, p. 1177–1178).

supostamente irrelevantes, mapeados pela economia comportamental. Após descrever a aversão à perda e fazer referência ao viés do *status quo*, por exemplo, Sunstein (1997, p. 1180) mencionou que, na regulação ambiental, seria possível manipular pontos de referência (*framing*) para influenciar o comportamento; em vez de dizer que uma proposta tenta melhorar a qualidade do ar e da água, o arquiteto de escolhas poderia destacar que a política estaria tentando recuperar a qualidade do ar e da água para atingir os níveis de qualidade de uma data prévia. Essa abordagem teria maior sucesso em fazer com que as pessoas reduzissem efetivamente a poluição.

Em aprofundamento dessa abordagem, Jolls, Sunstein e Thaler (1998, p. 1522–1541) descreveram em maiores detalhes como, a partir da incorporação de *insights* comportamentais, o direito poderia ser mais bem estruturado para atingir objetivos específicos. Por exemplo, após descrever como o viés retrospectivo (*hindsight bias*)[168] impacta o julgamento de jurados em ações de reparação civil nos Estados Unidos[169], Jolls, Sunstein e Thaler prescrevem a manipulação do conjunto de informações oferecido aos jurados – como no caso em que a origem do dano que está sendo discutida no processo é ocultada dos jurados – para aprimorar as avaliações de probabilidade dos eventos que poderiam causar o dano[170]. Já com relação à redução de assimetrias informacionais, Jolls, Sunstein e Thaler ponderam que, a partir de uma análise comportamental, é essencial pensar em *como* apresentar informações às pessoas, além de simplesmente fornecer mais informações. Nesse contexto, em variados contextos, quando o

[168] De acordo com o viés retrospectivo, demonstrado originalmente por Baruch Fischhoff e Ruth Beyth (1975), depois de um evento acontecer, as pessoas superestimam as probabilidades prévias de que ele viesse a se materializar e aparentam não estar surpresas. Por exemplo, hoje, após a Rússia ter invadido a Ucrânia, a probabilidade prévia de que uma guerra viesse a eclodir entre os dois países pode parecer maior do que se estimava, de fato, antes da invasão ser efetivada. Como descrito no experimento original, envolvendo estimativas da probabilidade dos possíveis resultados de viagens do então Presidente Nixon a Moscou e Beijing, ao relembrar-reconstruir probabilidades anteriormente estimadas de que um desses resultados viesse de fato a ocorrer, os participantes atribuíam probabilidades mais altas para eventos que acreditavam ter acontecido. Embora, em suas previsões originais, os sujeitos tivessem superestimado probabilidades baixas e subestimado probabilidades altas, suas probabilidades relembradas-reconstruídas raramente indicavam surpresas com os resultados que aconteceram na prática. Como explica Kahneman (2012, p. 253): "Solicitadas a reconsiderar suas antigas crenças, as pessoas lembram-se, em vez disso de suas atuais – um caso de substituição – e muitas vezes não conseguem acreditar que um dia acharam outra coisa. Sua incapacidade para reconstruir crenças passadas inevitavelmente o levará a subestimar em que medida você foi surpreendido por eventos passados".

[169] Afinal, os jurados frequentemente têm que avaliar a probabilidade de um evento que acabou ocorrendo. Tendo esse conhecimento prévio, os jurados acabam vinculando uma alta probabilidade ao evento apenas porque ele acabou acontecendo de fato, sem fazer uma verdadeira avaliação de probabilidade (JOLLS; SUNSTEIN; THALER, 1998, p. 1523-1524).

[170] O exemplo de Jolls, Sunstein e Thaler (1998, p. 1527-1528) envolve uma companhia de processamento de alimentos que está sendo acusada de ter utilizado um produto químico em seu processo produtivo de maneira negligente. Esse produto teria causado câncer em pessoas que moravam nos arredores da fábrica. A intervenção proposta pelos autores faria com que os jurados não soubessem o que exatamente a empresa teria feito. Seria possível dizer, por exemplo, que a empresa estava decidindo se utilizaria ou não o produto químico , e, diante desse cenário, os jurados teriam que avaliar *ex ante* qual seria a probabilidade de o produto químico provocar reações cancerígenas nas pessoas.

governo quer desencorajar determinado comportamento a partir de campanhas publicitárias, as informações devem ser manipuladas (*framing*), para explorar a aversão à perda[171] e a saliência[172], mas, ao mesmo tempo, evitar as armadilhas do otimismo excessivo[173].

Ao propor o paternalismo libertário, Thaler e Sunstein (2003b) também fazem considerações sobre as medidas que seriam utilizadas para implementá-lo. Inicialmente, Thaler e Sunstein (2003b, p. 2) ressaltam que a sua proposta de paternalismo, de espírito libertário, parte da generalização de dois experimentos envolvendo a alteração de regras *default*. Ambos os casos estavam ligados à inscrição automática de empregados em programas de previdência e poupança para a aposentadoria[174]. Em seguida, Thaler e Sunstein baseiam-se no já mencionado exemplo da alteração da ordem em que os alimentos estão dispostos em uma cafeteria para explicar a sua proposta de paternalismo. Assim como nas propostas de inscrição automática de empregados em planos de contribuição para a aposentadoria, no caso da cafeteria, a intervenção eventualmente proposta explora nitidamente elementos comportamentais subjacentes às escolhas dos agentes. Se no primeiro caso esses elementos estão ligados ao viés do *status quo* e à inércia, no segundo, eles se relacionam à saliência.

As demais intervenções mencionadas por Thaler e Sunstein (2003b) ao longo do mesmo artigo também exploram explicitamente fatores supostamente irrelevantes. De maneira geral, um dos argumentos mais convincentes de Thaler e Sunstein à sua proposta é a ausência de alternativas a um paternalismo libertário. Afinal, como os autores enfatizam, "[...] em muitos domínios, as pessoas não têm preferências claras, estáveis e bem-ordenadas. O que elas escolhem é um produto de efeitos de enquadramento, pontos de partida e regras default [...]"

---

[171] Nesse caso, deve-se enquadrar a informação em termos de perda em vez de ganhos, a fim de ter mais efetividade na mudança do comportamento (JOLLS; SUNSTEIN; THALER, 1998, p. 1536). Em uma campanha para promover o autoexame de mama, por exemplo, pode-se destacar as consequências negativas de não fazer autoexames – menor chance de descobrir um tumor tratável – em vez de seus benefícios.

[172] A ideia é apresentar informações de forma vívida e pessoal, em vez da mera evidência estatística. O exemplo de Jolls, Sunstein e Thaler (1998, p. 1537) é uma campanha antidrogas que apresenta um ovo sendo fritado, com a mensagem esse é seu cérebro sob a influência de drogas.

[173] Considerando o viés do otimismo, o governo poderia destacar o comportamento de terceiros, em vez de focar no próprio indivíduo. Em uma campanha para estimular a direção defensiva, por exemplo, como os sujeitos superestimam as suas habilidades como motoristas, o governo poderia destacar algo como Dirija defensivamente! Cuidado com as outras pessoas! (JOLLS; SUNSTEIN; THALER, 1998, p. 1537).

[174] O primeiro experimento, elaborado por Brigitte C. Madrian e Dennis F. Shea (2001), constatou que a inscrição automática aumenta consideravelmente a participação de empregados em planos de poupança para a aposentadoria. Ao mesmo tempo, a taxa de contribuição e a alocação de investimentos padrões tiveram considerável influência no comportamento dos participantes. Em geral, eles se mantiveram nos padrões adotados, embora as taxas de contribuição e a alocação de investimentos de empregados contratados antes da inscrição automática fossem diversas. Esse efeito poderia levar, inclusive, a menores taxas de contribuição. A outra intervenção, também ligada à inscrição automática, é o programa Save More Tomorrow, elaborado por Thaler e Benartzi (2004) e já mencionado nesta dissertação. Cf., nesse sentido, a nota 149 *supra*.

(THALER; SUNSTEIN, 2003b, p. 3)[175]. Nesse contexto, Thaler e Sunstein apresentam inúmeros exemplos de como regras *default*[176], pontos de partida (âncoras)[177] e enquadramentos específicos (*framing*)[178] influenciam o comportamento humano de maneira generalizada e são inevitáveis. Dessa forma, todas essas estratégias poderiam ser utilizadas para atingir o objetivo do paternalismo libertário: promover o bem-estar dos indivíduos, nos cenários em que suas preferências não são bem formadas.

Esses cenários em que o paternalismo - ainda que em uma forma fraca - seria inevitável formam, para Thaler e Sunstein (2003b, p. 29), um conjunto de intervenções reunidas, naquela ocasião, sob a alcunha de paternalismo mínimo. Essas são as intervenções mais promissoras dessa espécie de paternalismo suave, já que, além de seus efeitos positivos, segundo Thaler e Sunstein, enquanto não houvesse custos ou houvesse custos mínimos para escolher outras opções, elas seriam libertárias em máximo grau.

Ao lado do paternalismo mínimo, Thaler e Sunstein enumeram outros tipos de intervenção paternalista. Em vez de tentar direcionar as escolhas, o regulador pode, por exemplo, obrigar as pessoas a decidirem explicitamente, no que Thaler e Sunstein denominam

---

[175] Tradução livre. Texto original: "[...] *in many domains, people lack clear, stable, or well-ordered preferences. What they choose is a product of framing effects, starting points, and default rules* […]".

[176] Thaler e Sunstein mencionam outras propostas, além dos já mencionados experimentos de Madrian e Shea (2001) e Thaler e Benartzi (2004). No contexto das relações de emprego, o próprio Sunstein (2001, 2002) já havia explorado a possibilidade de regras *default* criarem um efeito dotação (*endowment effect*), de modo que a alocação original de direitos acaba determinando a forma como empregados e empregadores organizam suas relações. Nesse contexto, Sunstein argumentou que uma mudança nessa alocação, com a regra padrão atribuindo direitos renunciáveis aos empregados em vez de aos empregadores, poderia aumentar o fluxo de informações aos empregados, além de poder gerar bem-estar em contextos como a poupança para aposentadoria, na forma demonstrada nos experimentos já mencionados. Por outro lado, Eric Johnson e outros autores (1993) demonstraram a influência de regras *default* em um experimento quase natural envolvendo as leis securitárias dos Estados da Pensilvânia e de Nova Jérsei. Em Nova Jérsei, as seguradoras tinham responsabilidade limitada por danos sofridos por motoristas – apenas deveriam indenizar os danos graves; se desejassem, os motoristas poderiam adquirir ativamente o direito de responsabilizar integralmente as seguradoras por quaisquer tipos de danos. Na Pensilvânia, por outro lado, a alocação inicial de direitos envolvia a responsabilidade integral das seguradoras; os motoristas, por sua vez, poderiam renunciar a essa integralidade a troco de preços mais baratos. Enquanto apenas 20% dos motoristas de Nova Jérsei adquiriam o direito de responsabilizar integralmente as seguradoras, 75% dos motoristas da Pensilvânia mantinham esse direito. Camerer (2000, p. 295) menciona esse experimento e conclui que tornar uma opção o *status quo* parece estabelecer um ponto de referência, do qual as pessoas são relutantes em abrir mão.

[177] Nesse ponto, Thaler e Sunstein (2003b, p. 20) destacam que, em estudos de valoração contingente, método usado para atribuir valor a bens regulatórios, como maior segurança ou proteção ambiental, as pessoas não costumam ter preferências bem formadas, de modo que os pontos de partida (âncoras) geram consequências importantes para a tomada de decisões.

[178] Frederick (2003), por exemplo, demonstrou que, embora a literatura da época indicasse que pessoas do presente valorizavam a vida das gerações futuras em muito menor grau do que a de seus contemporâneos, alterações nas perguntas feitas a essas pessoas (*framing*) poderiam evocar respostas em sentidos distintos. Por exemplo, se um componente de equidade fosse inserido na pergunta – comparação entre um programa que salva 300 vidas na geração atual e nenhuma nas gerações subsequentes e um programa que salva 100 vidas na geração atual, 100 vidas na geração dos filhos dos participantes e 100 vidas na geração dos seus netos -, 80% dos participantes escolhiam o segundo programa (FREDERICK, 2003, p. 45-46).

escolhas coercitivas. Ao mesmo tempo, em um cenário mais intervencionista, o governo poderia adotar restrições procedimentais, para garantir que as escolhas dos indivíduos são realmente voluntárias e completamente racionais. Nesse caso, haveria custos adicionais (*fricções*) para deixar de seguir a regra *default*. Essas medidas poderiam ser justificadas em caso de problemas sérios de racionalidade limitada ou autocontrole.[179] Um exemplo é a exigência de períodos de reflexão (*cooling-off period*), antes que as pessoas tomem alguns tipos de decisão, como requerer um divórcio ou, mesmo, adquirir um produto em vendas porta a porta (THALER; SUNSTEIN, 2003b, p. 28). Em geral, essas medidas tentam evitar a lacuna de empatia quente-frio e evitar que emoções fortes levem a uma decisão inadequada. Aproximam-se, de certa forma, do conceito de *sludge*, que viria a ser desenvolvido posteriormente, como será discutido no Capítulo 4.

Por fim, o regulador também poderia impor restrições substantivas às escolhas, determinando os termos dentro dos quais escolhas poderiam ser feitas. Thaler e Sunstein exemplificam essas restrições com normas trabalhistas dos Estados Unidos, como o direito dos empregados de só serem demitidos por justa causa e as formas pelas quais a renúncia a esse direito pode ser consentida.

De maneira geral, portanto, as medidas propostas no âmbito do paternalismo libertário de Thaler e Sunstein são explicitamente informada por insights comportamentais. Mais do que isso, como em outros paternalismos suaves, boa parte de suas intervenções explora aspectos do comportamento que prejudicam a obtenção de decisões ótimas - como, por exemplo, a aversão à perda, os descontos intertemporais hiperbólicos e o viés do *status quo* - para aprimorar o processo de tomada de decisões (LOEWENSTEIN; HAISLEY, 2008, p. 22).[180] Em outras palavras, em tais políticas, além de irregularidades comportamentais serem o fundamento da intervenção, a técnica de intervenção é elaborada para explorar essas irregularidades, que não existiriam no universo do *Homo economicus*. Por isso, um padrão comum entre as intervenções dos paternalismos suaves comportamentais é redirecionar os padrões de comportamento das pessoas para fazer com que eles as ajudem, em vez de prejudicá-las, como poderia ocorrer na ausência de intervenção (LOEWENSTEIN; HAISLEY, 2008, p. 24).

---

[179] No mesmo sentido, Camerer e outros autores (2003, p. 1238-1250) apresentam, como exemplos do paternalismo assimétrico, exigências de períodos de reflexão, para evitar que as pessoas tomem decisões em estágios "quentes" emocionalmente, ou limitações nas escolhas de consumidores, como a imposição de prazos para combater a procrastinação.

[180] A exceção são as restrições procedimentais citadas acima, que, em vez de explorar fatores supostamente irrelevantes, buscam evitar que eles influenciem a tomada de decisões.

Como os exemplos mencionados acima demonstram, esse padrão não é identificado apenas na obra de Thaler e Sunstein. Na verdade, trata-se de algo comum em toda a literatura que prescreve a adoção de políticas de paternalismo leve, informado por insights comportamentais. A tabela abaixo reúne alguns dos exemplos citados acima, assim como outras propostas de intervenção identificadas na literatura que ilustram essa afirmação:

Tabela 3 - *Nudges* antes dos *nudges*: propostas de intervenção mapeadas na literatura

Autores	**Técnica de intervenção**	**Elemento comportamental explorado[181]**
Sunstein (1997)	Mudança no enquadramento de políticas ambientais com ênfase em recuperar um estágio anterior da qualidade da água e do ar, em vez de simplesmente melhorá-la (*framing*).	Aversão à perda e viés do *status quo*
Jolls, Sunstein e Thaler (1998)	Manipular a informação disponível para jurados avaliarem probabilidades (*framing*)	Viés retrospectivo
Jolls, Sunstein e Thaler (1998)	Manipular informações de campanhas governamentais para desencorajar certos comportamentos (*framing*).	Aversão à perda, saliência e otimismo excessivo
Thaler e Sunstein (2003b)	Inscrição automática de empregados em programas de poupança para aposentadoria (regra *default*).	Inércia, sugestionamento (influência social), efeito dotação e má formação de preferências
Thaler e Sunstein (2003b)	Alteração da ordem de alimentos saudáveis em um refeitório, para torná-los mais visíveis.	Saliência
Thaler e Sunstein (2003b)	Mudanças das âncoras (pontos de partida) em valorações contingentes de bens públicos.	Má-formação de preferências e ancoragem a ajuste
Thaler e Sunstein (2003b)	Manipulação de informações a respeito da avaliação de programas para salvar vidas no presente e no futuro (*framing*).	Descontos intertemporais e má-formação de preferências
Madrian e Shea (2001)	Inscrição automática de empregados em programas de poupança para aposentadoria (regra *default*).	Procrastinação, efeito dotação, viés do *status quo* e ancoragem e ajuste
Thaler e Benartzi (2004)	Inscrição automática de empregados em programas de poupança para aposentadoria (regra *default*).	Racionalidade limitada (dificuldade para definir o limite ótimo de poupança), descontos hiperbólicos, procrastinação, inércia e aversão à perda
Sunstein (2001)	Alteração da alocação inicial de direitos renunciáveis em favor dos empregados (regra *default*).	Efeito dotação e otimismo excessivo
Sunstein (2002)	Alteração da alocação inicial de direitos renunciáveis em favor dos empregados (regra *default*).	Efeito dotação e otimismo excessivo
Johnson e outros autores (1993)	Diferença na alocação inicial de direitos de ação – responsabilidade integral ou limitada de seguradoras (regra *default*).	Efeito dotação
Frederick (2003)	Manipulação em textos referentes a programas para proteção de vidas no futuro e no presente, para elicitar diferentes resultados (*framing*).	Descontos intertemporais e saliência
Camerer e outros autores (2003)	Alteração de regras *default* em contextos como o de seguros e a poupança para aposentadoria.	Viés do status quo, originado da aversão à perda, da procrastinação e do viés da omissão[182]

[181] Evidentemente, muitos elementos comportamentais podem ser explorados ao mesmo tempo. Nesta tabela, serão expostos alguns dos fatores elencados pelos autores para explicar o funcionamento de suas técnicas de intervenção.
[182] O viés da omissão é a tendência de favorecer omissões, como deixar que alguém morra, em vez de comissões equivalentes, como assassinar alguém de forma ativa (RITOV; BARON, 1990). Embora em vários contextos esse favorecimento faça sentido, há casos em que ele existe ainda que não existam diferenças entre as omissões e as

Camerer e outros autores (2003)	Exigência de que empresas reordenem seus contratos ou forneçam informações adicionais aparentemente irrelevantes, em contextos como a proteção do consumidor ou a proteção dos investidores.	Ignorância de fatores relevantes na tomada de decisões, como atributos ocultos ou taxas-base, ou influência de fatores irrelevantes como diferenças entre a forma como as opções são escolhidas
Camerer e outros autores (2003)	Períodos de reflexão antes de tomar decisões custosas ou irreversíveis, em contextos como a proteção do consumidor, o direito de família e mediações.	Proteção contra a empatia quente-frio[183]
Camerer e outros autores (2003)	Limitações nas escolhas dos consumidores, como a imposição de prazos para determinadas decisões.	Procrastinação
Loewenstein e Haisley (2008)	Contratos de depósito, em que sujeitos que querem perder peso depositam um valor e recebem um pagamento de duas vezes o valor depositado para cada dia que seu peso atingir uma determinada meta.	Aversão à perda
Benartzi e outros autores (2007)	Programa de diversificação automática de ativos, para reduzir a participação societária de empregados na empresa em que trabalham, se essa participação exceder um determinado percentual de seu portfólio de investimentos. Em todo caso, empregados poderiam optar por agir de forma diversa (regra *default*).	Inércia e procrastinação

Fonte: Elaboração do Autor.

Portanto, o traço em comum da maior parte das técnicas de intervenção que passaram a ser prescritas como remédios a falhas de mercado comportamentais é a tentativa de *explorar* traços do comportamento humano, que tradicionalmente poderiam criar prejuízos, para tentar aprimorar o processo de tomada de decisão. Como será demonstrado na próxima subseção, esse traço distintivo foi essencial para a elaboração do conceito de *nudge*, por Thaler e Sunstein.

3.2. A elaboração do conceito de *nudge*: um rótulo comum às diferentes técnicas de intervenção comportamentais

A partir de todo o desenvolvimento teórico resumidamente apresentado no item anterior, Thaler e Sunstein (2009, 2021)[184] propuseram um conceito que teria por objetivo reunir as políticas que, em geral, implementavam boa parte dos *paternalismos comportamentais*[185], sob a mesma alcunha. Nascia, ali, uma nova técnica de intervenção: os *nudges*.

comissões. Ilana Ritov e Jonathan Baron (1990, p. 265) exemplificam com escolhas a respeito da eutanásia ativa (droga letal) ou passiva (ordens para não ressuscitar alguém).

[183] Nesse caso, esse elemento comportamental não é propriamente explorado. Os mecanismos psicológicos por trás da intervenção são distintos: há uma tentativa de proteger o decisor de um viés comportamental.

[184] Enquanto esta dissertação estava sendo escrita, Thaler e Sunstein (2021) publicaram uma nova edição revisada de *Nudge*, com inúmeras alterações no texto. Por isso, tanto a obra original quanto a sua edição final serão referenciadas nesta dissertação, com destaques para eventuais modificações. De todo modo, para os fins desta seção, convém ressaltar que a definição de *nudge* se manteve fiel à elaboração original, embora o conceito tenha sido ampliado em obras acadêmicas de Sunstein, como ainda será demonstrado.

[185] Como será demonstrado no Capítulo 4, Thaler e Sunstein (2003b) admitem que essas políticas não precisavam ser necessariamente paternalistas. Políticas *welfaristas*, com o objetivo de aumentar o bem-estar social de maneira geral, também podem ser implementadas a partir das técnicas de intervenção propostas por Thaler e Sunstein.

Como já mencionado, em sua proposta original, um *nudge* é qualquer aspecto da arquitetura de escolha que altera o comportamento das pessoas de maneira previsível sem proibir quaisquer opções ou modificar substancialmente os incentivos econômicos envolvidos (THALER; SUNSTEIN, 2009, p. 6, 2021, p. 8). São intervenções simples e fáceis de se evitar. *Nudges* não são coercitivos e permitem que os indivíduos mudem de opinião e decidam de forma diferente da opção proposta pela política, sem qualquer sanção ou desincentivo econômico. Logo, *nudges* se diferenciam tanto de medidas de comando e controle (proibições ou imposições), quanto de subsídios e multas.

Ainda na definição original, depois de tecer considerações a respeito de limitações no comportamento humano, em comparação ao padrão de comportamento esperado de *Econs* – termo utilizado pelos autores para representar o *Homo economicus* –, Thaler e Sunstein acrescentam outro elemento que parece fazer parte da definição dos *nudges*. De acordo com os autores: "nudges incluem intervenções que alteram significativamente o comportamento de Humanos, mas seriam ignoradas por Econs" (THALER; SUNSTEIN, 2009, p. 8, 2021, p. 11)[186]. Nesse sentido, Thaler e Sunstein (2009, p. 8, 2021, p. 11–12) explicam que, enquanto *Econs* respondem principalmente a incentivos, humanos também são influenciados por *nudges*. Pontuam, ainda, que incentivos podem ser representados de diversas formas, inclusive sob a forma de custos cognitivos ou emocionais; apesar disso, tais custos podem ser criados a partir de um *nudge*, desde que, obrigatoriamente, os custos impostos sejam baixos – *i.e.*, não alteram substancialmente os incentivos dos decisores.

Essa é, basicamente, a forma como Thaler e Sunstein apresentam a definição de *nudge*. Embora os livros sejam marcados por inúmeros exemplos, que ainda serão mencionados abaixo, a definição original de *nudge* parece conter uma série de condições suficientes e necessárias para que uma determinada intervenção seja considerada um *nudge*[187].

A principal característica de um nudge é não ter coerção no que se refere às pessoas cujo comportamento se quer influenciar[188] - o que serve, inclusive, para Thaler e Sunstein demonstrarem que uma das concepções equivocadas do movimento antipaternalista é atrelar a coerção ao paternalismo, como já discutido no Capítulo 2. Em outros termos, *nudges* não forçam as pessoas a escolher de uma ou outra forma - seja por meio de proibições, seja a partir

---

[186] Tradução livre. Texto original: "[...] *nudges include interventions that significantly alter the behavior of Humans, even though they would be ignored by Econs*".

[187] Para uma das discussões mais aprofundadas a esse respeito, como ainda será mencionado neste Capítulo, Cf. (HANSEN, 2016)

[188] Como será demonstrado no Capítulo 4, *nudges* podem se originar de imposições para arquitetos de escolha do setor privado e, assim, ter um elemento coercitivo. Na terminologia adotada nesta dissertação, estes são os *nudges* impuros.

de determinações -, sob pena de sanção. Daí decorre uma primeira condição negativa para que uma intervenção seja um *nudge* em sua concepção original: a não coercitividade.

Ao mesmo tempo, como mencionado, um *nudge* não deve impor custos significativos aos agentes que decidem agir de maneira contrária à intervenção. Essa constatação expõe uma segunda condição negativa para que uma intervenção seja um *nudge*: a ausência de alteração substancial da estrutura de incentivos.

Além disso, como o conceito exposto acima expõe, *nudges* envolvem aspectos da arquitetura de escolhas, isto é, do contexto em que as pessoas tomam decisões (THALER; SUNSTEIN, 2009, p. 3, 2021, p. 3). Como Sunstein (2019a, p. 115) costuma repetir frequentemente, a arquitetura de escolhas é inevitável, ainda que as pessoas não a vejam.[189] Afinal, as pessoas tomam decisões o tempo inteiro e são influenciadas por diferentes aspectos dessa decisão. Como um aspecto da arquitetura de escolhas, um *nudge* busca influenciar alguma dentre as várias escolhas que um agente faz em determinado contexto. Mais do que um aspecto, enquanto técnica de regulação do comportamento privado, *nudges* são alterações da arquitetura de escolhas. Nesse caso, uma condição positiva para que uma intervenção seja um nudge parece ser a proximidade ao momento em que uma escolha é feita. Assim, quanto maior a distância entre a intervenção e a escolha, menos a intervenção se aproximaria de um *nudge*.

Ao mesmo tempo, *nudges* devem influenciar o comportamento de maneira previsível. A arquitetura de escolhas sempre influencia as decisões - o que leva Sunstein (2019a, p. 115) a afirmar, por exemplo, que o clima é, em si, um tipo de arquitetura de escolhas, já que influencia o que as pessoas escolhem. Um *nudge*, no entanto, não influencia o comportamento ao acaso; há, na verdade, uma tentativa de explorar aspectos previsíveis do comportamento humano e direcionar escolhas específicas em uma determinada direção. Em outras palavras, a previsibilidade é outra condição positiva para que uma intervenção seja um nudge, o que exige que a intervenção seja implementada intencionalmente pelo Estado.

Por fim, ao dizer que *nudges* seriam intervenções que influenciariam humanos, mas seriam ignorados por *Econs*, Thaler e Sunstein, para além de reforçar que *nudges* não funcionam com alterações significativas da estrutura de incentivos, parecem indicar que *nudges* atuariam a partir da exploração de fatores supostamente irrelevantes para *Econs*.[190] Em outros termos, o

---

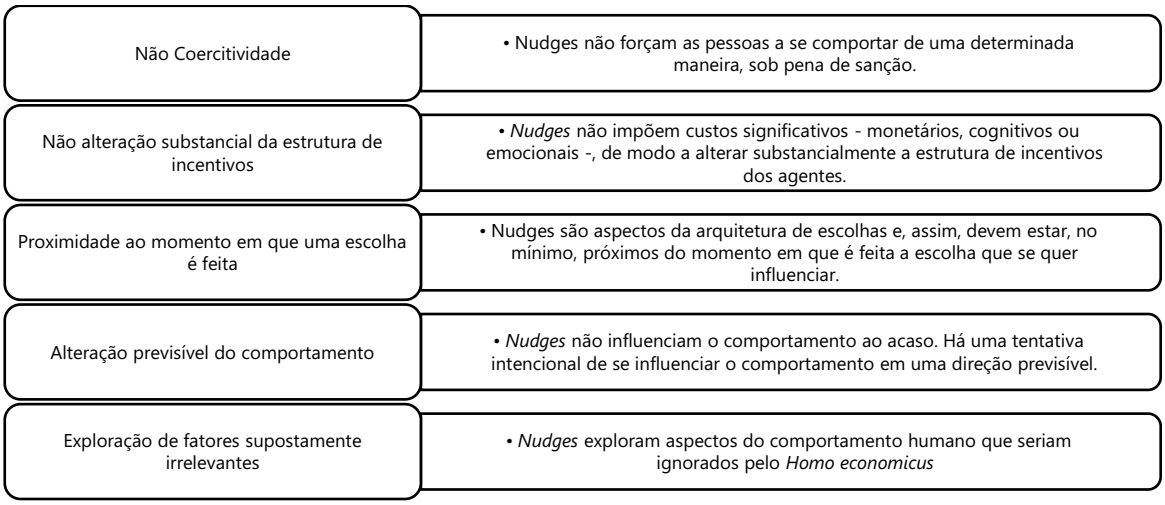
[189] Sunstein costuma ilustrar esta afirmação com uma anedota contada pelo escritor David Foster Wallace (2012, p. 263), no início de seu discurso de paraninfo na Kenyon College, em 21 de maio de 2005: "Dois peixinhos estão nadando e cruzam com um peixe mais velho que vem nadando no sentido contrário, que os cumprimenta dizendo: 'Bom dia, meninos. Como está a água?'. Os dois peixinhos continuam nadando por mais algum tempo, até que um deles olha para o outro e pergunta: 'Água? Que diabo é isso?'".

[190] Nesse sentido, Cf.: (HANSEN, 2016).

diferencial dos *nudges* seria explorar aspectos do comportamento humano que não se conformam ao modelo racional-maximizador do *Homo economicus*, assim como as políticas identificadas na Seção 3.1. Nesse caso, para que uma intervenção seja vista como um *nudge*, uma condição positiva parece ser explorar aspectos comportamentais que seriam ignorados pelo *Homo economicus*.

As condições suficientes e necessárias para que uma intervenção seja um *nudge*, de acordo com a definição originalmente proposta e descrita acima, encontram-se resumidas na imagem abaixo.

Figura 7 - Condições necessárias a um *nudge* na definição original



Fonte: Elaboração do Autor.

Boa parte das propostas apresentadas por Thaler e Sunstein como *nudges* na obra original cumpre essas condições. Um dos exemplos de *nudges* que Thaler e Sunstein (THALER; SUNSTEIN, 2009, 2021) mencionam é o próprio programa *Save More Tomorrow*, já destacado na seção anterior. Como Thaler e Sunstein (2009, p. 114, 2021, p. 189) explicam, o *Save More Tomorrow* é um sistema de arquitetura de escolhas baseado em princípios comportamentais, como aversão à perda, a inércia, a ilusão do dinheiro[191] e a maior facilidade de se aceitar restrições no autocontrole se elas forem implementadas no futuro. Diversas outras propostas mencionadas no livro *Nudge* também exploram aspectos comportamentais como:

[191] De acordo com Thaler e Sunstein (2009, p. 114, 2021, p. 189), a ilusão do dinheiro representa o fato de perdas serem sentidas em valores nominais, isto é, não ajustados pela inflação. Nesse caso, um aumento de 3% é visto como aumento, ainda que a inflação esteja em 4%.

a) tornar transparentes atributos ocultos, como taxas e custos de financiamentos imobiliários (THALER; SUNSTEIN, 2021, p. 225);
b) exigir que instituições financeiras tenham algumas modalidades padrão de financiamento imobiliário, a que seriam atribuídas um selo para que os consumidores soubessem de imediato as características do financiamento (THALER; SUNSTEIN, 2021, p. 225-226);
c) estímulo à tomada de decisões a respeito do consentimento para doação de órgãos, quando, ao renovar documentos, as pessoas recebem a sugestão de expressar seu consentimento quanto à doação de órgãos, sem qualquer obrigação de fazê-lo (THALER; SUNSTEIN, 2021, p. 265-268);
d) alteração da regra *default* para doação de órgãos, a fim de que os indivíduos sejam presumidamente doadores a não ser que se manifestem em sentido contrário (THALER; SUNSTEIN, 2009, p. 179-181, 2021, p. 261-264)[192].

Neste ponto, uma ressalva é necessária: as definições conceituais trazidas acima foram apresentadas em livros de divulgação científica.[193] Não há, propriamente, um rigor conceitual. De fato, Thaler e Sunstein (2021, p. 312) expressamente admitem que um livro inteiro poderia ser escrito a respeito de objeções conceituais, éticas, empíricas e de outras vertentes a respeito dos *nudges*; nesse caso, para conseguir enfrentar parte das críticas, os autores admitem que não se atolariam em argumentos semânticos. Mais do que isso, Thaler (2015, p. 324) admite que, originalmente, ele e Sunstein pretendiam simplesmente escrever um livro sobre o paternalismo libertário; o termo *nudge* surgiu como uma sugestão de um editor que avaliava a proposta de publicar o livro, já que parecia capturar aquilo que Thaler e Sunstein pretendiam fazer. Nesse caso, pode ser que essas obras não sejam a fonte mais confiável para se definir o que Thaler e Sunstein realmente consideram condições necessárias e suficientes para que uma intervenção seja um *nudge*.

[192] Na edição original do *Nudge*, Thaler e Sunstein (2009, p. 179–181) enfatizaram bastante o poder da presunção de consentimento para a doação de órgãos. Os autores destacavam, com expressa influência pelo trabalho de Johnson e Goldstein (2003), como a presunção de consentimento levava a níveis muito mais elevados de consentimento para doações de órgãos. Como destacam na edição final de *Nudge*, entretanto, Thaler e Sunstein (2021) não consideram o consentimento presumido a melhor estratégia para lidar com doações de órgãos; preferem, na verdade, o estímulo à tomada de decisões. Segundo os autores, embora a presunção de consentimento pareça promissora, a doação de órgãos sempre envolverá os familiares do doador; nesse caso, na prática, o consentimento raramente é presumido. Nos países que adotam o *opt-out* para doação de órgãos, a vontade dos familiares acaba sendo preponderante. Aspectos de infraestrutura também são preponderantes à efetividade de políticas nesse âmbito. Nesse contexto, Thaler e Sunstein indicam que estimular a tomada de decisões expressa por parte dos potenciais doadores parece ser uma estratégia mais eficaz ao aumento das doações de órgãos.
[193] Agradeço à Professora Dra. Rute Saraiva por chamar a minha atenção a esse ponto.

Um caminho para identificar os traços que constituem um *nudge* é a análise das obras subsequentes de Thaler e Sunstein, após a publicação original de *Nudge*. Embora Thaler tenha poucas publicações posteriores a esse respeito, Sunstein tem uma produção profícua a respeito de questões conceituais e filosóficas envolvendo os *nudges*.

Segundo Thaler (2015, p. 326), um *nudge* seria um pequeno aspecto do ambiente que atrai a atenção das pessoas e influencia o seu comportamento. Em linha com a definição apresentada originalmente, Thaler afirma que *nudges* são efetivos para *humanos*, mas não são para *Econs*, uma vez que eles já estão tomando decisões corretas. *Nudges*, assim, são fatores supostamente irrelevantes que influenciam escolhas para aprimorar a vida das pessoas. Em outra oportunidade, Thaler (2018b, p. 1283) registrou que *nudges* são aspectos da arquitetura de escolhas que influenciam decisões sem alterar recompensas objetivas ou incentivos. Para o *Homo economicus*, *nudges* não fariam diferença; no universo do *Homo sapiens*, todavia, *nudges* importam, e a boa arquitetura de escolhas pode facilitar a navegabilidade – ponto que Sunstein costuma reiterar frequentemente, como será demonstrado abaixo. De qualquer maneira, na obra de Thaler, parece claro que *nudges* são elementos que não seriam relevantes para *Econs*, embora alterem as decisões de humanos.

Por outro lado, na obra de Sunstein, de maneira geral, a definição de *nudge* parece ser mais ampla do que aquela descrita no livro *Nudge* e na própria obra de Thaler. De fato, Sunstein (2014d, p. 583, 2015c, p. 110, 2015b, p. 417, 2019a, p. 59, 2019c, p. 127) chega a afirmar que *nudges* seriam abordagens que preservam a liberdade e que conduzem as pessoas a determinadas direções, mas também permitem que elas sigam o seu próprio caminho. Dessa forma, na abordagem de Sunstein, nudges seriam definidos apenas por condições negativas: a não coercitividade e a ausência de alteração substancial na estrutura de incentivos.

Ao partir de um conceito mais amplo, Sunstein elabora subdivisões e tenta explicar como *nudges* podem funcionar de maneira tão diversa e, ainda assim, serem reunidos como uma única técnica de intervenção. Inicialmente, Sunstein (SUNSTEIN, 2014d, 2015c, 2019a, 2019c) também apresenta uma lista não exaustiva de 10 diferentes exemplos de intervenção que poderiam ser considerados *nudges*.

O primeiro exemplo mencionado são as regras *default*, talvez o *nudge* mais eficaz – e, também, como será demonstrado no Capítulo 6, o que gera maiores desafios éticos.[194] Em

[194] Inúmeras políticas baseadas em alteração de regras *default* já foram apresentadas na seção anterior. De todo modo, convém ressaltar que, segundo Johnson e Goldstein (2003, p. 1338), a maioria das escolhas em políticas públicas têm um padrão que será imposto no caso de os indivíduos não terem feito qualquer escolha explícita. Analisando o exemplo de doações de órgãos, os autores mencionam que os países europeus possuem duas políticas

seguida, Sunstein menciona o uso de normas sociais, isto é, ressaltar que a maioria de um grupo apresenta determinado comportamento, a fim de fazer com que o indivíduo também se comporte daquela maneira. Também são mencionados por Sunstein: a) a simplificação, em que ajustes são feitos em programas existentes para torná-los mais intuitivos; b) o aumento de facilidade e de conveniência, quando é dada visibilidade a certas opções mais vantajosas para o decisor; c) a divulgação (*disclosure*) de informações úteis de maneira compreensível e acessível; d) os alertas, dramáticos ou não, como os que são encontrados em embalagens de cigarro; e) as estratégias de compromisso prévio, nas quais as pessoas se comprometem com uma ação específica em um futuro determinado; f) os lembretes, por e-mail ou mensagem de texto; g) a evocação de intenções de implementação, quando se questiona ao decisor se ele pretende implementar determinado rumo de ação; e, por fim, h) as informações sobre a natureza e as consequências de escolhas passadas, podendo indicar que a forma como o sujeito vem agindo está sendo excessivamente prejudicial.

A partir de tais distinções, Sunstein (2015b, p. 426) ressalta que *nudges* não são necessariamente elaborados para conter ou explorar o Sistema 1 e os vieses comportamentais. Em sua perspectiva, alguns *nudges* tentariam contra-atacar o Sistema 1 ao fortalecer o Sistema 2, com tentativas de ampliar o papel da deliberação no julgamento das pessoas, por meio, por exemplo, da divulgação de informações relevantes, de técnicas de desenviesamento e do uso de estratégias de pré-compromisso (SUNSTEIN, 2015b, p. 425).

A lição é simples: alguns *nudges* teriam como fim aprimorar a navegabilidade, isto é, fazer com que as pessoas consigam chegar à destinação desejada de maneira mais fácil. No mesmo sentido, Sunstein, em conjunto com Lucia A. Reisch, (2019, p. 124), afirma expressamente que definir *nudges* em termos da exploração de vieses comportamentais é uma receita para confusão, já que muitos *nudges* tentam simplesmente ampliar a navegabilidade, como faria um GPS. Nesse caso, alguns *nudges* poderiam ser úteis ainda que na ausência de qualquer viés (SUNSTEIN, 2015b, p. 426).

Esse ponto é retomado quando Sunstein (2015b, p. 426–427) distingue diferentes tipos de *nudge*, com enfoque na avaliação de questões éticas – tema que será abordado no Capítulo 6. Sunstein apresenta três distinções: a) *nudges* paternalistas e *nudges* que enfrentam falhas de mercado; b) *nudges* educativos, que tentam informar as pessoas para que elas possam tomar

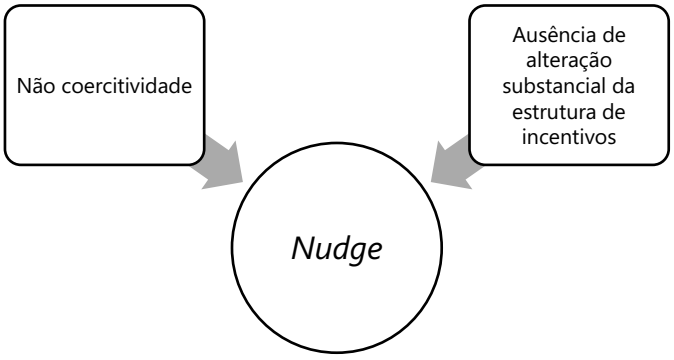
---

de *default* diferentes. Em países em que o consentimento é presumido, as pessoas são presumidamente doadoras de órgãos e, caso assim não desejem, devem se posicionar ativamente para modificar sua decisão. Por outro lado, em países de consentimento explícito, as pessoas só serão doadoras caso se registrem para tanto. Johnson e Goldstein (2003, pág. 1339) demonstraram que há substancial aumento no percentual de consentimento, quando o *default* é o consentimento presumido.

melhores decisões por conta própria, e *nudges* que não têm elementos educativos; e c) *nudges* que exploram vieses comportamentais e *nudges* que promovem decisões deliberativas. A primeira distinção está mais relacionada aos fundamentos que levam à regulação do comportamento privado e será abordada no Capítulo 4. As outras duas distinções deixam claro que, para Sunstein, *nudges* não precisam explorar fatores supostamente irrelevantes e, tampouco, serem indiferentes para *Econs*. Afinal, o fornecimento de mais informações ou educação é uma medida que poderia aprimorar até mesmo as decisões de agentes racionais, que maximizam a utilidade de suas escolhas.

Como será demonstrado na próxima seção, há um vasto debate em torno da definição de *nudge*. A amplificação do conceito na obra de Sunstein contribui para alimentar esse debate. Sunstein não ignora que essa discussão exista. De fato, ao discutir a distinção entre *nudges* e *sludges*, Sunstein (2020a) menciona que a definição usual de *nudge* é controversa. Ele menciona que, como um *nudge* não tem um significado autoevidente, interpretações baseadas em bases linguísticas ou lógicas serão inevitáveis. Apesar disso, Sunstein (2020a, p. 7) sugere que a vantagem da definição tradicional é ter um termo que capture todo o rol de intervenções que preserva a liberdade de escolhas e, ao mesmo tempo, direciona as escolhas das pessoas em algumas direções. Estas parecem ser as condições definidoras de um *nudge* para Sunstein e a sua visão favorece, claramente, um conceito bastante amplo.

Figura 8 - A definição ampla e agregadora de *nudge* na obra de Sunstein



Fonte: Elaboração do Autor.

Como é possível perceber, obter clareza a respeito das condições necessárias e suficientes para que uma intervenção seja considerada um *nudge* não é uma tarefa fácil. Como Sunstein é quem mais se dedicou à questão, as suas colocações a respeito da definição parecem ser mais relevantes que as de Thaler. Mais do que isso, a definição mais ampla de *nudge* parece

capturar uma de suas características importantes: *nudges* nada mais são do que um rótulo comum para diferentes tipos de intervenção baseadas em *insights* comportamentais.

Esse rótulo foi essencial para popularizar essa técnica de intervenção na vida privada. De qualquer forma, toda a fluidez em torno do conceito gera inúmeros debates. Na próxima seção, serão apresentadas as principais controvérsias existentes na literatura acerca da definição de *nudge*.

3.3. *Nudges* em sentido estrito e diferentes tipos de *nudges*: debates conceituais em torno de *nudges*

Naturalmente, considerando a fluidez do conceito de *nudge* e as condições bastante abertas que o definem, como exposto na subseção anterior, há inúmeras controvérsias em torno de qual deveria ser a verdadeira definição de *nudge*[195]. Reforça essas controvérsias o fato de a definição de *nudge* adotada por Sunstein reunir sob o mesmo espectro técnicas de intervenção muito diversas. Como já mencionado, Sunstein (2020a, p. 7) sugere que a vantagem de ter uma definição mais ampla é ter um termo que capture todo o rol de intervenções que preserva a liberdade de escolhas e, ao mesmo tempo, direciona as escolhas das pessoas em algumas direções. No entanto, se *nudges* são qualquer intervenção que preserve a liberdade de escolhas e direcione as pessoas para alguns caminhos específicos, há risco de se incluir dentro dessa definição práticas como, em um extremo, colocar drogas de maneira sub-reptícia na bebida de uma pessoa e, em outro extremo, deixar Bíblias dos Gideões Internacionais[196] em um quarto de hotel (HAUSMAN, 2018).

Ainda que o exemplo de Daniel M. Hausman seja exagerado e incorreto tecnicamente – drogar uma pessoa, afinal, reduz a sua liberdade de escolhas –, com uma definição ampla, Sunstein acaba por reunir sob a mesma definição técnicas de regulação que geram diferentes tipos de dilemas éticos. De fato, boa parte das medidas propostas por Thaler e Sunstein são muito brandas: simplificação de requisitos burocráticos para procedimentos como pedidos de financiamento estudantil; lembretes via e-mail ou mensagem de texto; evocação de intenções de implementar uma determinada opção; divulgação de informações nutricionais de um

[195] Cf., por exemplo: (BALDWIN, 2014; BARTON; GRÜNE-YANOFF, 2015; BERTHET; OUVRARD, 2019; BOVENS, 2009; CONGIU; MOSCATI, 2020, 2021; FRENCH, 2011; HANSEN, 2016; HANSEN; JESPERSEN, 2013 HAUSMAN, 2018; HAUSMAN; WELCH, 2010; LIN; OSMAN; ASHCROFT, 2017; MONGIN; COZIC, 2018; MÜNSCHER; VETTER; SCHEUERLE, 2016; OLIVER, 2015; SAGHAI, 2013; TOR, 2016).

[196] Os Gideões Internacionais são uma organização internacional dedicada à distribuição de Bíblias em hotéis, escolas, hospitais etc. Cf.: https://www.gideons.org/. Acesso em: 25 abr. 2022.

alimento ou de informações de consumo de energia elétrica etc. (SUNSTEIN, 2014d, 2015d, 2019a, 2019c). Outras técnicas, no entanto, exploram processos cognitivos mais rápidos que contornam a consciência e, com isso, geram maiores preocupações. É o caso, por exemplo, de regras *default* e de alterações do enquadramento (*framing*).

Como Sunstein chega a admitir - o que já foi mencionado na última seção –, um *nudge* não tem um significado autoevidente. Não por outra razão, há, na literatura, inúmeras tentativas de tornar a definição de *nudge* mais clara, sob bases linguísticas, lógicas e epistemológicas. Em geral, essas tentativas levam a definições reduzidas de *nudge*, mais próximas à definição adotada originalmente no próprio livro *Nudge* por Thaler e Sunstein. Outras propostas, adotando uma definição mais ampla, assim como Sunstein tem feito, acabam por desenvolver diferentes maneiras de classificar *nudges* em categorias distintas.

Inicialmente, é possível identificar, de maneira geral, pelo menos dois conjuntos de tentativas de refinamento da definição de *nudge*. Em ambos os casos, as condições necessárias para que uma intervenção seja considerada um *nudge* são mais rigorosas do que aquelas apontadas por Sunstein em sua obra mais recente. Em outras palavras, ambas as tentativas de refinamento conceitual de *nudge* apresentam definições estritas de *nudge*, em contraposição à definição ampla utilizada por Sunstein.

A primeira e mais relevante tentativa de refinamento da definição de *nudge* retoma um elemento incluído por Thaler e Sunstein na definição original do conceito de *nudge* e já mencionado na subseção anterior: a exploração de fatores supostamente irrelevantes. Como mencionado anteriormente, Thaler e Sunstein afirmam que "nudges incluem intervenções que alteram significativamente o comportamento de Humanos, mas seriam ignoradas por *Econs*" (THALER; SUNSTEIN, 2009, p. 8, 2021, p. 11). Em linha com essa afirmação, seriam afastadas da definição de *nudge* intervenções que também afetassem agentes racionais. Afinal, na proposta original de Thaler e Sunstein, para serem um *nudge*, as intervenções deveriam ser racionalmente neutras, ou seja, irrelevantes normativamente para um agente econômico racional (TOR, 2016, p. 5).

Segundo Hansen (2016), autor que desenvolveu esse ponto minuciosamente, esse seria o princípio da economia comportamental e seria um critério fundacional à definição de *nudge*. Em linha com esse princípio, a definição de *nudge* deveria estar ligada às limitações cognitivas e aos vieses comportamentais catalogados nas pesquisas da economia comportamental. Para alinhar a definição de *nudge* ao princípio da economia comportamental, Hansen (2016), de maneira semelhante a Hausman e Brynn Welch (2010), acrescentou duas condições positivas à definição de *nudge*: (i) ser utilizado em razão de limitações na tomada de decisões que criam

dificuldades para as pessoas agirem racionalmente em conformidade com o seu próprio interesse; e (ii) explorar essas limitações como parte integral da tentativa de direcionar o comportamento das pessoas de maneira previsível.

De maneira geral, a definição de *nudge* da qual parte a maioria dos críticos e autores que discutem *nudges* em termos teóricos inclui elementos como esses. Ainda que sem teorizar com tanta profundidade, esses autores costumam partir de uma concepção estrita de *nudge*, que liga essa técnica de intervenção à exploração ou à manipulação de aspectos automáticos e inconscientes do comportamento.

Riccardo Rebonato (2014), por exemplo, cujos argumentos têm sua qualidade reconhecida inclusive por Sunstein[197], afirma expressamente ser difícil encontrar na literatura uma definição única e concisa que enfrente todas as condições necessárias e suficientes do paternalismo libertário - ou dos *nudges*, que são tratados como sinônimos por Rebonato. Em sua visão, *nudges* englobariam um conjunto de intervenções que busca superar vieses cognitivos e inadequações decisórias, para, a partir da exploração de tais elementos comportamentais, influenciar as decisões de um indivíduo de maneira reversível em direção às escolhas que ele faria se tivesse tempo e informações ilimitadas, além das habilidades analíticas de um agente racional (REBONATO, 2014, p. 359).

No mesmo sentido, Gigerenzer (2015, p. 363) afirmou que, após a publicação do livro *Nudge*, quase tudo o que afeta o comportamento tem sido reconhecido como um *nudge*, o que esvazia o conceito de sentido. Por isso e para tentar se vincular à definição original de *nudge*, Gigerenzer adota expressamente o conceito exposto por Rebonato. Do mesmo modo, Evans Selinger e Kyle Whyte (2011, p. 928) sustentam que um *nudge* não tenta informar o sistema automático, mas sim trabalhar com a influência inevitável de vieses comportamentais. Engelen e Nys (2020), por seu turno, consideram *nudges* como modificações deliberadas na arquitetura de escolhas com a intenção de influenciar previsivelmente o comportamento das pessoas, explorando mecanismos psicológicos a-racionais[198]. Ao justificar a sua escolha conceitual, Engelen e Nys afirmam que um conceito estrito de *nudge* se alinha com a maior parte da literatura que discute *nudges*, assim como com o espírito da obra de Sunstein que parte, em primeiro lugar, da observação de que pessoas comuns nem sempre são tão racionais quanto elas gostariam, em razão de heurísticas e vieses que influenciam o seu comportamento.

---

[197] Na primeira nota de rodapé do artigo *The Ethics of Nudging*, Sunstein (2015b, p. 414) admite que algumas das objeções mais fortes ao uso de *nudges* e da arquitetura de escolhas como ferramentas regulatórias podem ser encontradas na obra de Rebonato.

[198] Engelen e Nys (2020) utilizam a expressão a-racional por considerarem que exigências de racionalidade não se aplicam às heurísticas cognitivas, mas apenas aos agentes que tomam decisões e suas crenças, preferências e ações.

Nesse contexto, ainda na discussão desse primeiro conjunto de tentativas de refinar a definição de *nudge*, um questionamento inevitável se volta aos efeitos da adoção de uma definição estrita, que vincula a exploração de fatores supostamente irrelevantes, isto é, a neutralidade racional, à definição de *nudge*. Afinal, que tipos de técnicas de intervenção deixariam de ser vistas como *nudges*?

A principal distinção entre o conceito amplo de *nudge* adotado por Sunstein e o conceito estrito ora exposto se refere à medida em que o fornecimento ou a divulgação de informações poderia ser considerado um *nudge* (BERTHET; OUVRARD, 2019, p. 2). Em diferentes obras, Sunstein (2014a, 2015b, 2019c) considera o fornecimento de informações factuais e a persuasão racional como *nudges*. Em resposta a argumentos de que *nudges* exploram vieses comportamentais, por exemplo, Sunstein (2018b, p. 64) afirma, como já mencionado, que muitos *nudges* buscam apenas aumentar a navegabilidade, isto é, facilitar o caminho para que as pessoas cheguem ao seu destino preferido. A divulgação de informações, por exemplo, seria útil ainda que nenhum viés estivesse presente.

Apesar disso, autores como Hansen (2016) apontam que o fornecimento de informações factuais ou a persuasão racional não tentam explorar vieses cognitivos e, por isso, excluem essas técnicas da definição de *nudge*[199]. Afinal, mesmo *Econs* poderiam ser beneficiados por tais intervenções. A separação entre o fornecimento de informações e os *nudges* é encontrada, inclusive, na obra de autores que prescrevem a utilização de *nudges*[200].

O fornecimento de informações, no entanto, pode ser exigido em razão de vieses comportamentais ou de limitações nas capacidades de processamento cerebral. De fato, pode ser que informações tenham que ser fornecidas pela incapacidade de o sujeito absorvê-las por conta própria. Nesse caso, porém, de acordo com Hansen (2016), o fornecimento de informações seria apenas uma estratégia de desenviesamento (*debiasing*)[201], já que não há uma tentativa de exploração de vieses cognitivos para influenciar as pessoas de forma previsível.

---

[199] Nem toda persuasão é racional: é possível utilizar falácias lógicas para influenciar a decisão de um indivíduo (HANSEN, 2016). Nesse caso, a razão de intervir será uma falha cognitiva – a limitada capacidade de processamento de informações – e, ao mesmo tempo, a intervenção explorará vieses cognitivos.

[200] Dilip Soman (2015, p. 26-34), por exemplo, ao descrever as diferentes possibilidades de se influenciar escolhas para que um agente opte por outra opção, apresenta quatro conjuntos de técnicas: a) regulação, incluindo proibições, exigências e regras de conformidade; b) incentivos econômicos, incluindo impostos, penalidades, bônus e subsídios; c) informação e persuasão, que incluem publicidade, divulgação de informações e materiais promocionais; e d) *nudges* e arquitetura de escolhas, incluindo regras *default*, simplificação e alteração dos padrões de consentimento (*opt-in versus opt-out*).

[201] Técnicas de desenviesamento são testes elaborados para tentar descrever as condições para a observação e os processos psicológicos que precisam ser evocados ou controlados para evitar um viés comportamental (FISCHHOFF, 1982, p. 422). Quando bem-sucedidas, essas técnicas podem aprimorar os julgamentos. *Nudges* serão comparados com técnicas de desenviesamento no Capítulo 4.

Outra situação, para Hansen (2016), envolve técnicas como o envio de lembretes. Nesse caso, para o autor, não há apenas o fornecimento de informações factuais; a intervenção tem por fundamento uma limitação cognitiva: a memória limitada. Ao mesmo tempo, porém, ela explora vieses cognitivos, ao agir com base na saliência. Para Hansen, então, essa estratégia poderia ser um *nudge*, já que influencia o comportamento de humanos, sem impactar o comportamento de *Econs*.

Também nesse sentido, Luc Bovens (2009) tenta diferenciar os *nudges* de peças publicitárias organizadas pelo governo, que envolvem o fornecimento de informações. Em sua visão, embora possa haver alguma sobreposição, *nudges* exploram mecanismos comportamentais dos quais as pessoas não têm controle completo; a publicidade social, por sua vez, pode envolver o simples fornecimento de informações e levar a uma alteração do comportamento que é completamente autônoma. Em alguns casos, porém, a publicidade pode ser elaborada para explorar esses mecanismos do comportamento, quando, por exemplo, a informação é enquadrada (*framing*) de forma a ter maior impacto na tomada de decisões, sem, contudo, atrair uma decisão deliberativa.

De maneira alternativa, Bovens (2009, p. 209) sugere que *nudges* poderiam ser diferenciados da publicidade a partir da estipulação de que um *nudge* deva afetar a situação atual de escolha. Dessa forma, um cartaz com pulmões afetados pelo câncer não é um *nudge*, mas um pacote de cigarros com a mesma imagem seria um *nudge*, já que a imagem seria evocada no momento em que o sujeito fosse fumar. Para Bovens, essa condição estaria alinhada à maior parte dos exemplos de Thaler e Sunstein.

Alguns apontamentos adicionais feitos por Hansen (2016) auxiliam a compreender os limites da definição estrita de *nudge* que é adotada por ele e por outros autores. O conceito original de *nudge*, tal como proposto por Thaler e Sunstein (2009, p. 8, 2021, p. 8), não engloba o veto a qualquer opção de escolha. Não é claro, porém, se o acréscimo de opções deveria ser considerado um nudge. Para Hansen (2016), essa definição dependeria da relevância racional da opção acrescida.

Em um primeiro aspecto, há acréscimos de opções que poderiam levar o sujeito a escolher uma opção que lhe traz mais utilidade. Imagine, por exemplo, que o sujeito precisa escolher uma entre duas opções de leitura: “Terra Sonâmbula”, de Mia Couto, e “Não é um rio”, de Selva Almada. Entre estas duas obras, o sujeito aufere maior utilidade com a leitura de Terra Sonâmbula. Seu livro favorito, porém, é “O Mestre e Margarida”, de Mikhail Bulgákov; se essa opção for acrescida, ela será a sua escolha final, a partir de uma decisão racional. Nesse caso, a intervenção - *i.e.,* o acréscimo da opção - não se justifica em razão de limitações na

tomada de decisões nem explora essas limitações. Na abordagem de Hansen, então, o acréscimo de opções racionalmente relevantes não deveria ser considerado um *nudge*.

Em outra perspectiva, Hansen (2016) aponta que o acréscimo de opções pode ser utilizado para explorar limitações cognitivas e direcionar a escolha de um indivíduo. As preferências de humanos reais são muito influenciáveis pelo contexto e podem ser alteradas pelo acréscimo de opções irrelevantes. Por isso, um dos vieses comportamentais mais explorados no *marketing* é o efeito chamariz (*decoy effect*).

Ariely (2008) exemplifica esse efeito com um experimento que utilizava planos de subscrição da revista *The Economist*. Em um primeiro cenário, duas opções estavam disponíveis: a assinatura *online* por US$ 59,00; e a assinatura impressa e *online* por US$ 125,00. Nesse contexto, a maior parte dos participantes optava pela assinatura *online*. Em um segundo cenário, uma terceira opção foi acrescida: a assinatura exclusivamente impressa, também por US$ 125,00. O acréscimo dessa opção, irrelevante em termos de utilidade, levou a um aumento substancial do volume de assinaturas no plano impresso e *online*: em comparação com o plano apenas impresso; o plano impresso e *online* parecia vantajoso, já que continha mais benefícios pelo mesmo valor. Nesse caso, o acréscimo de opções é motivado claramente por limitações cognitivas e as explora para direcionar as escolhas de indivíduos. Em outras palavras, na abordagem de Hansen, o acréscimo de opções racionalmente irrelevantes deve estar englobado na definição de *nudge*[202].

Além do acréscimo de informações, outra questão relevante é a alteração não substancial de incentivos. Thaler e Sunstein (2009, p. 8, 2021, p. 11-12) deixam claro que apenas excluem do conceito de *nudge* mudanças de incentivos significativas. Selinger e Whyte (2011, p. 927-928), nesse contexto, pontuam que algumas intervenções que utilizam a alteração de incentivos - como o site *Sticck.com*[203] ou o programa britânico que altera os custos associados ao abandono do cigarro ou à realização de atividades saudáveis[204] - são consideradas *nudges* de maneira

[202] A remoção de opções irrelevantes pode ser utilizada para criar um efeito inverso ao do efeito chamariz já descrito. No exemplo do *The Economist*, diante do segundo cenário, com as três opções de assinatura, o arquiteto de escolhas poderia excluir a opção de assinatura exclusivamente digital e assim direcionaria as pessoas à assinatura exclusivamente digital. Essa remoção de opções racionalmente irrelevantes também poderia ser considerada um *nudge*.

[203] Como explicam Thaler e Sunstein (2009, p. 233-234, 2021, p. 56-57), o *Stickk.com* é um site em que as pessoas fazem compromissos financeiros e não financeiros para cumprir um determinado objetivo. No caso dos compromissos financeiros, o sujeito transfere uma determinada quantia e faz o compromisso de cumprir um objetivo até uma certa data; se o objetivo for cumprido, ele recebe o dinheiro de volta; caso contrário, o dinheiro é doado à caridade. Compromissos não financeiros envolvem a pressão dos pares – e-mails à família e aos amigos anunciando o sucesso ou as falhas – e o monitoramento conjunto do objetivo do sujeito a partir de um *blog*.

[204] O *Behavioural Insights Team* considera um *nudge* um programa que oferece *kits* antifumo aos cidadãos que contêm adesivos e *vouchers* que oferecem descontos em planos de academia, aulas de natação e outras atividades saudáveis (SELINGER; WHYTE, 2011, p. 926).

equivocada. Embora Selinger e Whyte tenham um ponto - já que algumas de tais estratégias envolvem alterações substanciais de custos -, Hansen (2016) aponta que alterações pouco substanciais de incentivos podem ser consideradas *nudges*.

Estratégias de *marketing* que definem o preço de um produto como R$ 0,99, em vez de R$ 1,00, por exemplo, não teriam qualquer relevância a um agente racional. Para humanos, no entanto, a estratégia pode ter impactos relevantes. Nessa estratégia, o relevante para a tomada de decisões não é a alteração nos incentivos; de fato, a mudança do preço nesse caso é, em maior medida, uma reorganização da arquitetura de escolhas que torna o valor baixo mais saliente para o consumidor. Para tornar esse ponto mais claro na definição de *nudge*, Hansen (2016) propõe que, em vez de afastar dessa definição quaisquer incentivos significantes, *nudges* devem ser considerados como uma função da arquitetura de escolhas que altera o comportamento das pessoas e trabalha independentemente de mudanças de incentivos.

De maneira semelhante a Hansen, Avishalom Tor (2016, p. 7) argumenta que algumas das aplicações dos *nudges* tentam influenciar o comportamento a partir de ferramentas que não são racionalmente neutras – e que, portanto, não seriam classificadas como *nudges*. O exemplo de Tor é o uso de normas sociais. Em sua visão, um agente racional seria plenamente capaz de mudar seu comportamento para evitar atrair a fúria de seus pares ou para obter seus favores. Os custos e benefícios de sua interação social, inclusive, poderiam se estender ao domínio econômico, já que as pessoas podem ser relutantes em fazer transações com pessoas que não seguem as normas sociais. Nesse caso, incluir o uso de normas sociais na definição de *nudge* seria um equívoco de Thaler e Sunstein, em linha com sua própria teoria.

Outra tentativa de refinamento da definição de *nudge* foi elaborada por Yashar Saghai (2013). Inicialmente, Saghai (2013, p. 488) sustenta que uma condição necessária a um *nudge* é a preservação do conjunto de escolhas, isto é, se A tenta intervir nas escolhas de B, o conjunto de escolhas será preservado se permanecer inalterado ou for expandido, se comparado à situação prévia à influência. Para Saghai, no entanto - e aqui reside o principal diferencial de sua abordagem -, a preservação do conjunto de escolhas não é suficiente para garantir a liberdade de escolhas; afinal, se, na prática, as pessoas não puderem resistir à influência, não haverá liberdade de escolhas concreta. A partir dessa premissa, Saghai acrescenta uma nova condição necessária à definição de *nudge*: ser substancialmente não controladora.

Essa condição é expressamente inspirada na abordagem de Faden e Beauchamp (1986, p. 243), que, como discutido no Capítulo 2, exige influências, pelo menos, substancialmente não controladoras, para que um ato possa continuar sendo substancialmente autônomo. Nessa abordagem, uma influência é substancialmente não controladora se o sujeito influenciado

consegue facilmente deixar de agir em conformidade com a influência, se assim desejar. Como explica Saghai (2013, p. 489), essa visão é influenciada pelo critério da resistibilidade. Uma influência seria facilmente resistível – e, portanto, substancialmente não controladora – se o sujeito a quem ela se direciona for capaz de se opor sem esforços ao rumo de ação indicado pela influência, se não quiser escolher agir desse modo.

Assim como Hansen, Saghai (2013, p. 489) considera que *nudges* se baseiam em processos cognitivos rasos, termo que engloba processos não deliberativos ou não completamente deliberativos com três propriedades em comum: a) são rápidos; b) consomem menos recursos; e c) geram respostas que não são resultados de completa deliberação. Considerando essas propriedades, Saghai questiona em que circunstâncias uma influência que explora processos cognitivos poderia ser facilmente resistível. Para ter essa habilidade, segundo Saghai, um sujeito B que sofre a influência de A para agir de um determinado modo φ deverá:

a) ter a capacidade de se tornar consciente da pressão de A para levá-lo a φ (capacidades de *direcionar a atenção*);
b) ter a capacidade de *inibir* a propensão à φ gerada pela influência de A (capacidades de *inibição*);
c) não estar sujeito a uma influência ou colocado em circunstâncias que prejudicariam significativamente o exercício fácil (*effortless*) das capacidades de chamar a atenção e de inibição.

Se, diante de uma intervenção, as pessoas perderem as capacidades indicadas acima ou se o seu exercício se tornar muito custoso, a intervenção não será substancialmente não controladora, e, por consequência, ela não será um *nudge*. Com base nessa argumentação, Saghai (2013, p. 491) chega à seguinte definição de *nudge*:

> A *empurra* (*nudges*) B quando A torna mais provável que B irá φ, sobretudo por desencadear processos cognitivos rasos de B, sendo que a influência de A preserva o conjunto de escolhas de B e é substancialmente não controladora (i.e., preserva a liberdade de escolhas de B)[205].

Em oposição aos *nudges*, de acordo com Saghai (2013, p. 491), intervenções substancialmente controladoras, que, assim como nudges, também exploram processos

[205] Tradução livre. Texto original: "*A nudges B when A makes it more likely that B will w, primarily by triggering B's shallow cognitive processes, while A's influence preserves B's choice-set and is substantially noncontrolling (ie, preserves B's freedom of choice)*".

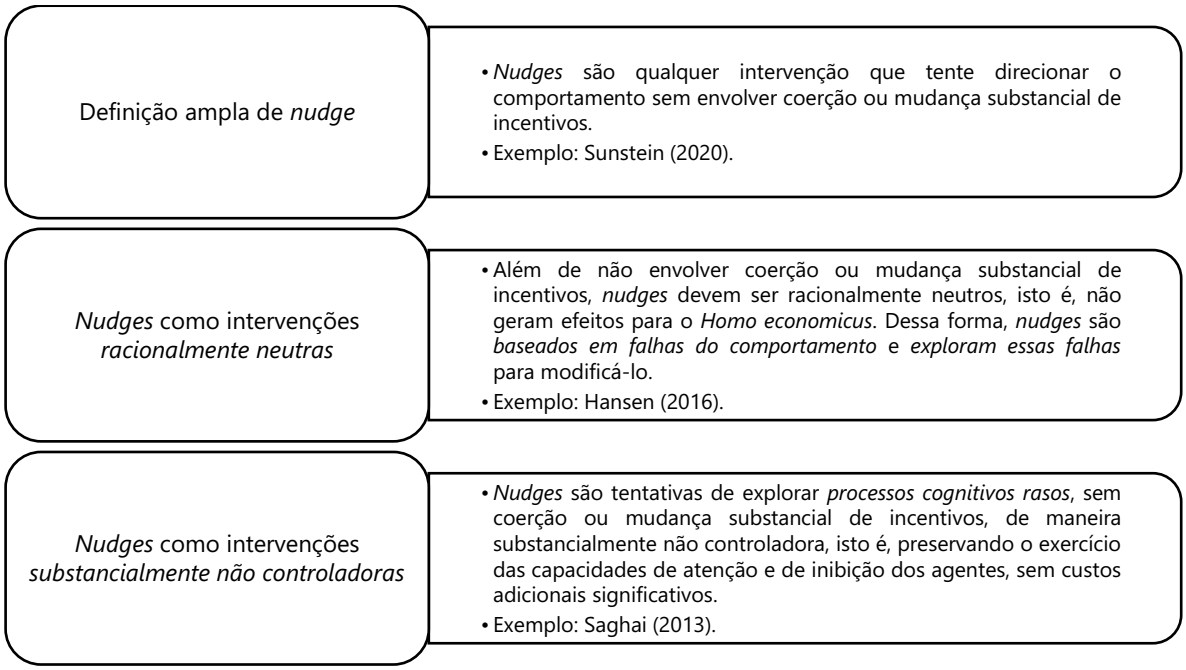
cognitivos rasos, são cutucões comportamentais (*behavioural prods*). Na abordagem de Saghai, assim como *nudges*, cutucões comportamentais não envolvem coerção ou alteração substancial de incentivos; essas intervenções, no entanto, são mais controladoras; não há, diante delas, fácil resistibilidade.

Como visto, a definição de Saghai é mais reduzida do que a definição de Sunstein; em sua abordagem, a exploração de processos cognitivos rasos é um elemento integrante da definição de *nudge*. Sua abordagem é, inclusive, ainda mais reduzida do que a abordagem de Hansen e dos outros autores que veem os *nudges* como ferramentas que exploram vieses comportamentais ou outros processos cognitivos rasos. Afinal, na abordagem de Saghai, além de explorar processos cognitivos rasos, uma intervenção precisa ser substancialmente não controladora para ser considerada um *nudge*. Em razão disso, na prática, muitas intervenções que são vistas como *nudges* deixariam de sê-lo na abordagem de Saghai.

Em suma, além da definição ampla de *nudge* adotada por Sunstein, é possível encontrar duas principais tentativas de refinamento conceitual na literatura. Ambas excluem do escopo da definição de *nudge* intervenções coercitivas ou que alteram substancialmente a estrutura de incentivos dos agentes. Esse é o elemento unificador de todas essas abordagens. Além disso, a primeira tentativa de refinamento, ilustrada pela abordagem de Hansen (2016), considera que *nudges* são estratégias racionalmente neutras, nos termos de Tor (2016), que exploram processos automáticos do comportamento e que tentam resolver problemas causados por essas mesmas falhas. Para a outra tentativa de refinamento, elaborada por Saghai (2013), *nudges* exploram processos cognitivos rasos e influenciam o comportamento de maneira substancialmente não controladora, isto é, preservando o exercício das capacidades de atenção e de inibição, sem que sejam exigidos maiores esforços para efetivar tais capacidades.

A imagem abaixo sumariza os principais pontos de todas essas abordagens:

Figura 9 - Três definições de *nudge*



Fonte: Elaboração do Autor.

Além de refinamentos da definição de *nudge*, autores que adotam o conceito amplo, adotado por Sunstein, acabam tendo que reconhecer a existência de inúmeros tipos de *nudges* distintos. Com isso, assim como ocorre na obra de Sunstein, já mencionada na seção anterior, esses autores acabam desenvolvendo diferentes tipologias de *nudges*. Como sugerem Vincent Berthet e Benjamin Ouvrard (2019, p. 3), essas tipologias se dividem em: a) categorias orientadas pelo mecanismo; e b) categorias orientadas pela técnica.

As categorias orientadas pelo mecanismo buscam identificar os mecanismos psicológicos que são explorados pelos nudges.

Robert Baldwin (2014), por exemplo, sugere que *nudges* poderiam ser classificados em três patamares distintos a depender dos impactos que trazem para a autonomia. *Nudges* de primeiro grau melhoram a tomada de decisões reflexivas através do fornecimento de mais informações. Já *nudges* de 2º grau baseiam-se em limitações comportamentais para empurrar a decisão à direção desejada; ainda assim, o agente decisor mantém a capacidade de perceber o empurrão e avaliar os seus efeitos. O exemplo dado por Baldwin são as regras *default*. *Nudges* de 3º grau, por fim, intervêm na liberdade de escolhas por meio da manipulação do comportamento. Os exemplos dados por Baldwin são *framing*, saliência de opções de escolha ou estratégias emocionais que moldam as preferências dos indivíduos, o que dificultará a identificação para os indivíduos de que há uma intervenção.

Philippe Mongin e Mikaël Cozic (2018), por seu turno, afirmam que um *nudge* pode se referir a três significados distintos. O primeiro – *nudge* 1 – seria representado por intervenções que interferem minimamente nas condições de escolha, simplesmente deixando de proibir qualquer opção ou de impor mudanças significativas nos incentivos econômicos. Embora possa haver algum viés comportamental subjacente à intervenção, ela não o explora intencionalmente, como parte integrante da intervenção. Em uma segunda acepção – *nudge* 2 –, as intervenções utilizam falhas na racionalidade de maneira instrumental, de maneira semelhante à proposta por Hansen (2016) e Hausman e Welch (2010). Finalmente, em uma terceira acepção – *nudge* 3 –, *nudges* seriam intervenções que promovem o bem-estar e contra-atacam as falhas na racionalidade que afetam os processos de tomada de decisão individual.

Adrien Barton e Till Grüne-Yanoff (2015, p. 343–344), por outro lado, após adotar expressamente a definição ampla de *nudge* elaborada por Sunstein, distinguem entre três diferentes tipos de *nudge*. A primeira categoria reúne *nudges* que, para direcionar o comportamento, exploram o conhecimento acerca de processos cognitivos não deliberativos (Sistema 1), para desencadear o uso de heurísticas específicas. É o caso, por exemplo, da alteração de posições em uma lista, da alteração de regras *default* e de mudanças de enquadramentos (*framing*). Outros *nudges*, em vez de desencadear o uso de heurísticas, atuam para bloquear essas heurísticas. A exigência de períodos de reflexão (*cool-off periods*) antes de uma determinada aquisição, como já mencionado, busca evitar os efeitos da empatia quente-frio e a influência de processos emocionais em decisões sérias. Há, por fim, *nudges* que atuam simplesmente a partir do fornecimento de informações, a fim de alterar o comportamento das pessoas de maneira previsível. De acordo com Barton e Grüne-Yanoff, uma placa informando que um rio é perigoso para nadar, exemplo clássico de Mill para ilustrar o paternalismo suave, seria considerada um *nudge* nessa modalidade.

Em outra perspectiva, Hansen e Andreas Maaløe Jespersen (2013)[206] sugerem uma distinção entre *nudges* de tipo 1 e *nudges* de tipo 2.[207] Inicialmente, Hansen e Jespersen (2013, p. 14) não tratam escolhas como sinônimo de comportamento. Em linha com a teoria da ação, escolhas (*choices*) seriam interpretadas como o resultado de uma intervenção do sistema reflexivo; comportamentos (*behaviors*), por seu turno, incluem "ações não voluntárias", que não seriam resultado de deliberação - ou de escolhas -, mas seriam "causados" por outros

[206] Como já mencionado, Hansen (2016) é autor de uma tentativa minuciosa de refinamento conceitual de *nudge*. Essa tentativa é posterior à obra escrita com Jespersen e pode indicar, assim, uma evolução em seu pensamento a respeito do tema.

[207] Hansen e Jespersen fazem, ainda, uma distinção entre *nudges* transparentes e não transparentes. Como essa distinção está mais ligada à avaliação normativa de *nudges*, ela será retomada no Capítulo 6.

eventos; na visão de Hansen e Jespersen, esses processos cognitivos só seriam ações, em um sentido derivativo, como um comportamento que estava sob o potencial controle de um determinado agente. Dessa forma, Hansen e Jespersen usam o termo comportamento, para reunir todos os movimentos corporais e processos cognitivos; os conceitos escolha e ação, contudo, são destinados apenas a movimentos e processos que resultam da reflexão.

A partir de tais definições teóricas, Hansen e Jespersen (2013, p. 14-16) distinguem *nudges* de tipo 1 e *nudges* de tipo 2. Em sua visão, ambos os tipos de *nudge* teriam por objetivo influenciar maneiras automáticas de pensamento – o que implica um conceito mais reduzido do que o adotado por Sunstein.

Nesse contexto, *nudges* de tipo 1 envolveriam influências a comportamentos mantidos pelo pensamento automático ou por suas consequências, sem envolver qualquer forma de pensamento reflexivo. Em outras palavras, *nudges* de tipo 1 afetam comportamentos que não envolvem deliberação, julgamento ou escolhas. A mudança do tamanho padrão de pratos em uma cafeteria, para estimular um menor consumo de alimentos, assim como a redução da distância entre as linhas de uma rodovia, para reduzir a velocidade dos veículos[208], são exemplos de *nudges* de tipo 1.

Por outro lado, *nudges* de tipo 2 teriam como foco influenciar a atenção e as premissas do pensamento reflexivo, a partir de influências aos processos automáticos de pensamento. Nos termos do Sistema Dual de Kahneman, esses *nudges* explorariam o Sistema 1, a fim de atrair a atuação do Sistema 2. Dessa forma, *nudges* de tipo 2 influenciariam comportamentos que podem ser caracterizados como ações, isto é, o resultado de deliberação, julgamentos e escolhas. Um exemplo de *nudges* de tipo 2 seria o clássico caso da mosca desenhada em mictórios do aeroporto de Amsterdam[209]. De acordo com Hansen e Jespersen (2013, p. 15), esse *nudge* tenta capturar processos visuais guiados pelo pensamento automático e, assim, atrair a atenção reflexiva, para que o sujeito permaneça focado no ato de urinar.

As categorias orientadas pela técnica, por seu turno, foram criadas para enfrentar a questão da efetividade dos *nudges*, a fim de identificar quando e por que *nudges* funcionam (BERTHET; OUVRARD, 2019, p. 3). O objetivo, aqui, é identificar diferentes técnicas de nudge, independentemente dos mecanismos psicológicos subjacentes à intervenção.

---

[208] Thaler e Sunstein (2009, p. 37–39) mencionam que a *Lake Shore Drive*, que contorna o Lago Michigan em Chicago, é uma das vias mais conhecidas do mundo. Nessa via, há uma série de curvas em *S* que são perigosas e muitos motoristas não conseguem reduzir a velocidade. Nesse contexto, para encorajar essa redução da velocidade, foram pintadas faixas horizontais na pista e a distância entre essas faixas se torna gradativamente menor. Com a diminuição da distância, cria-se a sensação de que a velocidade está aumentando e, em resposta, os motoristas reduzem a velocidade.

[209] Vide: https://worksthatwork.com/1/urinal-fly. Acesso em: 02 mai. 2022.

Nesse sentido, uma primeira tentativa relevante de divisão da taxonomia da arquitetura de escolhas foi feita por Robert Münscher, Max Vetter e Thomas Scheuerle (2016). Os autores sugeriram três técnicas básicas de intervenção ligadas à arquitetura de escolhas: a) informação da decisão (*decision information*); b) estrutura da decisão (*decision structure*); e c) auxílio da decisão (*decision assistance*).

A primeira categoria – informação da decisão – reúne estratégias que têm como alvo a forma como são apresentadas as informações relevantes à tomada de decisões, sem que as opções sejam propriamente alteradas. Nesse conjunto de intervenções, encontram-se técnicas como: a) traduzir informações, a partir, por exemplo, do reenquadramento (*reframe*) ou de simplificações da forma como as informações estão dispostas; b) tornar informações mais visíveis, o que inclui técnicas de *feedback* e de transparência quanto às informações que podem estar invisíveis, como tarifas de cartão de crédito; e c) fornecer pontos de referência sociais, a partir, por exemplo, de referências ao comportamento dos pares e ao comportamento de influenciadores de opinião (*opinion leader*).

As intervenções baseadas na estrutura da decisão, por outro lado, envolvem alterações no arranjo de opções e no formato de tomada de decisão, para modificar as opções disponíveis no momento de decidir. De acordo com Münscher, Vetter e Scheuerle, são técnicas de estrutura de decisão: a) a alteração de *defaults* de escolha, o que inclui o estabelecimento de uma regra padrão para casos em que não houver ação, assim como o estímulo a escolhas ativas (*prompted choices*); b) a alteração do nível de esforço referente à escolha de uma determinada opção, o que inclui o aumento ou a diminuição na percepção do esforço físico ou financeiro[210]; c) a alteração da extensão ou da composição das opções, que pode ser implementada com técnicas como a mudança de categorias, a alteração das consequências de uma opção, a conexão de uma decisão a um benefício ou custo e a alteração das consequências sociais de uma escolha.

Por fim, intervenções baseadas no auxílio de decisões ajudam os decisores a seguir as suas próprias intenções. Münscher, Vetter e Scheuerle consideram técnicas de auxílio de decisões: a) o fornecimento de lembretes; e b) técnicas que facilitam o compromisso, como

[210] Para Münscher, Vetter e Scheuerle (2016), o termo "esforço financeiro" é utilizado de maneira análoga ao esforço físico em decisões que envolvem transações financeiras. A ideia é que, ainda que sem alterações nos incentivos, é possível alterar a forma como as pessoas percebem uma despesa como mais ou menos custosa. Se os custos forem diluídos ao longo do tempo, por exemplo, as pessoas irão considerar a transação como menos custosa. Inúmeras variações podem ser pensadas nesse sentido. Essa noção está ligada ao fenômeno da *dor do pagamento* (*pain of payment*), isto é, o desprazer direto e imediato que se segue ao ato de realizar um pagamento (ZELLERMAYER, 1996). O que Münscher, Vetter e Scheuerle denominam alteração do esforço financeiro está ligado ao fenômeno do acoplamento (*coupling*), isto é, à medida em que o consumo é temporalmente associado ao efetivo pagamento de uma determinada quantia (RAGHUBIR; SRIVASTAVA, 2008). Quanto maior o desacoplamento entre pagamento e consumo, menor será o desconforto atrelado à respectiva transação (dor do pagamento).

intervenções que favorecem o autocompromisso ou que exigem que o sujeito faça um compromisso público com uma determinada opção de escolha.

Eric Johnson e outros autores (2012) também fornecem uma divisão da arquitetura de escolhas entre técnicas que estruturam a tarefa de escolha e técnicas que descrevem as opções de escolha.

A primeira categoria - estrutura da tarefa de escolha - envolve técnicas como: a) a definição do número de alternativas disponível; b) a utilização de tecnologia da informação, como ferramentas de busca e sistemas de recomendação de produtos, além de ferramentas que auxiliam a decisão, como comparadores de alternativas de escolhas; c) a definição de regras *default*; d) alterações na estrutura intertemporal de uma tarefa, como atrair a atenção para opções do futuro ou incluir janelas temporais limitadas para que uma oportunidade seja exercida; e e) o estabelecimento de diferentes fases para que diferentes atributos de uma escolha sejam escolhidos.

Já a descrição de opções de escolha reúne intervenções como: a) a divisão de opções e atributos em grupos ou categorias; e b) a forma como atributos específicos de uma opção são elaborados, de modo mais ou menos saliente. Johnson e seus colaboradores destacam, ainda, que o *design* dos atributos de uma opção de escolha pode ser alterado segundo diferentes técnicas. Nesse contexto, os autores mencionam: a) a parcimônia, que envolve a redução do esforço cognitivo necessário para a compreensão das informações ligadas a uma determinada alternativa; b) a linearidade, em que relacionamentos não lineares de um atributo da decisão são apresentados de maneira mais linear; c) a utilização de escalas comuns para avaliar atributos semelhantes ou valores quantitativos; d) a tradução de atributos, em que as consequências de uma determinada opção de escolha são expressamente mapeadas para o agente decisor; e (e) a expansão de um atributo, como a expansão da escala de um determinado atributo para que o agente dê mais peso a esse fator ao decidir[211].

Luca Congiu e Ivan Moscati (2020), por seu turno, argumentam que, independentemente da definição de *nudge* que se adote, todos os elementos da arquitetura de escolha podem ser decompostos em duas dimensões: Mensagem (*Message*) e Ambiente (*Environment*).

[211] No exemplo de Johnson e de seus colaboradores (2012, p. 496), o consumo de gasolina de um veículo é apresentado com numeradores mais elevados, o que dá a impressão de que as diferenças entre opções distintas são maiores, aumentando o peso que esse fator teria na decisão de adquirir um determinado veículo. Assim, em vez de apresentar o consumo de gasolina em pequenas distâncias (100 milhas), o arquiteto de escolhas poderia apresentá-lo em grandes distâncias (10.000 milhas), o que tornaria mais clara a distinção entre um veículo econômico e outro que consumisse um alto volume de gasolina.

A dimensão Mensagem reúne intervenções baseadas na comunicação verbal, oral e escrita, que o arquiteto de escolhas pode enviar aos decisores e que descreve a situação de escolha ou fornece informações a seu respeito. Essas intervenções, segundo Congiu e Moscati, podem enfrentar motivações racionais ou tentar explorar limitações cognitivas, vieses e hábitos. Em sua visão, mensagens como "Fumar causa câncer" tentam enfrentar motivações racionais, enquanto exemplos como "Inscreva-se e cancele a assinatura online a qualquer tempo", exploram elementos comportamentais, tal como a inércia dos consumidores.

Em outra perspectiva, a dimensão Ambiente reúne os elementos e aspectos da situação de escolha que o decisor pode encontrar nos seus arredores físicos ou virtuais, perceber por seus próprios sentidos ou interagir. Incluem a maneira como as alternativas são organizadas, a presença de restrições, obstáculos, opções *default* e a exposição a estímulos sensoriais, como música ou imagens. Assim como as Mensagens, o Ambiente também pode, segundo Congiu e Moscati, explorar aspectos racionais e falhas de racionalidade; de fato, enquanto lombadas construídas em uma estrada diminuirão a velocidade de motoristas racionais que não desejam danificar seus veículos, a alteração de uma regra *default* explorará *fatores supostamente irrelevantes*.

3.4. Definição de *nudge* adotada: alteração da arquitetura de escolhas para explorar fatores supostamente irrelevantes e reduzir as fricções associadas a uma opção de escolha

Como já discutido, um *nudge* não tem um significado autoevidente, de modo que interpretações baseadas em bases linguísticas ou lógicas sempre serão inevitáveis (SUNSTEIN, 2020a). O conceito, na verdade, é impreciso em sua própria origem; Thaler e Sunstein pretendiam escrever um livro sobre o paternalismo libertário, quando um editor sugeriu que a palavra *nudge* parecia capturar exatamente o que eles queriam propor (THALER, 2015, p. 324). O seu objetivo, assim, voltava-se à identificação de novos fundamentos para intervenção estatal, que pudessem representar uma terceira via entre o paternalismo e o libertarianismo. Não havia propriamente uma tentativa de definição de uma nova técnica de regulação do comportamento privado, no sentido da especificação de condições necessárias e suficientes para que uma intervenção fosse considerada um *nudge*.

De qualquer forma, o termo *nudge* foi incorporado à literatura e ganhou cientificidade. De fato, essa expressão é amplamente utilizada na literatura para representar um conjunto de intervenções bastante heterogêneo. Considerando a imprecisão mencionada, no entanto, chegar a um consenso acerca de qual seria a melhor definição de *nudge* é uma tarefa impossível. Apesar

disso, considerando que esta dissertação pretende avaliar normativamente se nudges são capazes de gerar bem-estar e se geram riscos à autonomia, é essencial evidenciar qual a definição de nudge que parece mais adequada à luz da literatura..

Essa definição de *nudge* adotada deve permitir que as críticas sejam enfrentadas de maneira direta e que *nudges* sejam devidamente comparados com estratégias alternativas também informadas por *insights* comportamentais. A comparação dos *nudges* com estratégias alternativas como técnicas de desenviesamento, *boosts*, *thinks*, *nudges plus* e *sludges* será realizada no Capítulo 4. De todo modo, comparar as diferentes definições expostas neste Capítulo é o primeiro passo para optar pela definição mais apropriada às finalidades desta dissertação.

Inicialmente, com relação à definição ampla, é necessário enfatizar que, como Sunstein (2020a) argumenta, uma expressão abrangente – como *nudge* – pode, de fato, ter a vantagem de reunir o conjunto de intervenções não coercitivas e que não envolvem alterações substanciais na estrutura de incentivos. Essa reunião é importante para facilitar a popularização de intervenções sutis, que podem ter resultados efetivos e com custo-benefício significativo. Ela evidencia, além disso, que técnicas coercitivas ou alterações na estrutura de incentivos nem sempre serão necessárias. Apesar disso, tamanha amplitude gera certa incerteza a respeito da avaliação normativa de técnicas específicas, uma vez que *nudges* seriam quase que um gênero dentro do qual se inserem várias espécies de intervenção distinta. Nesse caso, não faria muito sentido discutir a legitimidade de um *nudge* como um gênero de técnica regulatória; as implicações de cada espécie de *nudge*, afinal, seriam muito distintas.

Mais do que isso, a discussão apresentada na Seção 3.1 evidencia que as estratégias prescritas na literatura antes da elaboração da definição de *nudge* eram, de maneira geral, mais próximas à definição estrita de *nudge*. Não se prescreviam, exatamente, técnicas, tais como a simples divulgação de informações factuais ou a simplificação de processos burocráticos. Na verdade, as estratégias mais inovadoras e que mais chamaram a atenção do público exploram nitidamente vieses comportamentais - como a alteração de regras *default* e a utilização de diferentes enquadramentos (*framing*) - ou tomam medidas para evitar o efeito desses vieses – períodos de reflexão (*cooling-off periods*), o que seria mais próximo da definição de *sludge*, que será explorada no Capítulo 4.

Por outro lado, a escolha de uma definição ampla de *nudge* acaba por evitar alguns dos questionamentos que surgem a essa estratégia, quando adotada uma definição estrita. Como apontam Hausman e Welch (2010, p. 134), a defesa dos *nudges* empreendida por Thaler e Sunstein – especialmente por Sunstein - praticamente sugere que *nudges* não têm qualquer

custo. De fato, ao se deparar com críticas que apontam o potencial caráter manipulativo dos *nudges*, por ultrapassarem a deliberação racional, explorarem vieses comportamentais e não serem transparentes, Sunstein (2015b, p. 428) frequentemente aponta que nem todos os nudges geram tais preocupações e, logo, que não há objeções a *nudges* por si só, mas a fundamentos específicos para a utilização de *nudges*, que podem ser ilícitos.

Este caminho, em vez de enfrentar diretamente as críticas e justificar a adoção de *nudges* com base numa comparação entre os seus benefícios e os prejuízos, acaba tentando suavizar essas críticas. Afinal, quando se diz que nem todo *nudge* é problemático, já que alguns *nudges* promovem a deliberação e são transparentes, desconsidera-se que essas estratégias podem sequer estar sendo consideradas por quem adota uma definição restrita de *nudge*. Pode ser, então, que os impactos normativos da crítica não sejam bem avaliados tampouco se defina com clareza de que ônus o regulador deve se desincumbir para justificar a implementação de um *nudge*, especialmente em casos mais polêmicos, quando essa implementação é justificada por fundamentos paternalistas.

Por todas essas razões, se o objetivo desta dissertação é avaliar a legitimidade dos *nudges*, não parece razoável adotar uma definição ampla de *nudge*, como faz Sunstein. Nesse caso, é preciso definir qual das definições estritas é mais adequada a esse empreendimento, isto é, se um *nudge* deve ser ligado à neutralidade racional ou a influências substancialmente não controladoras, sendo que, em ambos os casos, a medida explora fatores supostamente irrelevantes ou, nos termos de Saghai (2013), processos cognitivos rasos.

A noção de *nudge* enquanto medida substancialmente não controladora que explora processos cognitivos rasos, exposta na tentativa de Saghai (2013) de salvar o conceito de *nudge*, não parece adequada às finalidades desta dissertação.

Como já mencionado, Saghai baseia-se expressamente na noção de influências controladoras, exposta por Faden e Beauchamp (1986) como um dos critérios para a avaliação da medida em que um ato é autônomo, conforme já apresentado no Capítulo 2. Para Saghai, a ausência de influências substancialmente não controladoras é uma condição necessária para que uma intervenção seja um *nudge*. Ao mesmo tempo, no entanto, Saghai reconhece que há técnicas muito semelhantes aos *nudges*, que, no entanto, são excessivamente interventivas e acabam exercendo um controle excessivo sobre o agente: os cutucões comportamentais (*behavioral prods*). De acordo com o próprio Saghai (2013, p. 491), como *nudges* e cutucões comportamentais são influências que utilizam os mesmos meios para alterar o comportamento – como a alteração de regras *default*, mudanças de enquadramento (*framing*), âncoras etc. – eles

são frequentemente confundidos. Para Saghai, porém, essa confusão decorre de um foco raso na técnica de influência empregada, que negligencia o nível de controle que a influência exerce.

Esta não parece ser a distinção mais adequada. A definição estrita de *nudge* adotada por Saghai incorpora à definição da técnica de intervenção um juízo de valor ético: *nudges*, por definição, preservariam a autonomia. Nesse caso, toda a discussão normativa que envolve a legitimidade de *nudges* perderia o sentido. Em outras palavras, Saghai busca salvar o conceito de *nudge* para garantir que esse tipo de influência sempre preserve a liberdade de escolhas.

Apesar disso, se a técnica de influência empregada é a mesma, haveria, então, um gênero do qual *nudges* e cutucões comportamentais seriam espécies. Como, então, diferenciar essas duas estratégias supostamente distintas? O que faz com que uma regra *default* específica seja um *nudge* e não um cutucão comportamental? Na visão de Saghai, como já mencionado, essa avaliação depende: a) da capacidade de direcionar a atenção à influência; b) da capacidade de inibir a influência; e c) do nível de esforço necessário para efetivar as capacidades de atenção e de inibição diante da influência, *i.e.*, a influência é não controladora se permite que essas capacidades sejam exercidas sem um custo excessivo. Embora esses aspectos sejam relevantes para a avaliação de uma intervenção específica, eles não parecem adequados para definir essa intervenção conceitualmente. Há, na verdade, uma avaliação normativa que se desloca ao momento de definição de tais técnicas.

A definição estrita que considera *nudges* intervenções não coercitivas e racionalmente neutras, como exposto por Hansen (2016), por outro lado, parece apropriada às finalidades desta dissertação.

Em um primeiro aspecto, essa definição está mais alinhada à literatura das ciências comportamentais que precedeu a elaboração do conceito de *nudge*, assim como às discussões teóricas que apontam críticas a essa medida. Os próprios exemplos de Thaler e Sunstein incluem, na maioria das vezes, intervenções que atuam a partir da exploração de fatores supostamente irrelevantes.

Apesar disso, a principal razão alegada por Hansen para justificar sua abordagem não parece o melhor caminho em defesa de uma definição estrita de *nudge*. Como mencionado na Subseção 3.1.3, Hansen justifica a sua abordagem a partir de um trecho específico de *Nudge*, em que Thaler e Sunstein afirmam que um *nudge* é qualquer influência que afeta o comportamento de humanos e que seria indiferente ao *Homo economicus* (THALER; SUNSTEIN, 2009, p. 8, 2021, p. 11).

Embora se tenha identificado esse elemento como condição necessária à definição original de *nudge* na Seção 3.2 acima, não é clara a medida em que Thaler e Sunstein viam na

exploração de fatores supostamente irrelevantes uma condição necessária para que uma intervenção fosse considerada um *nudge*. Na verdade, pode ser que, ao mencionar que *nudges* influenciam humanos, mas seriam ignorados por *Econs*, Thaler e Sunstein pretendiam apenas sinalizar que *nudges* não envolvem alterações bruscas na estrutura de incentivos dos agentes.

De qualquer forma, há um fundamento alternativo, de caráter metodológico, para se filiar a uma definição estrita alinhada à de Hansen e as de outros autores. A maior parte dos *nudges* que geram algum tipo de preocupação enquadram-se em tal definição estrita. A alteração de uma regra *default* para presumir o consentimento dos agentes, as alterações de enquadramentos (*framing*), a utilização de âncoras, a alteração da ordem de opções e outras alterações da arquitetura de escolhas que exploram vieses comportamentais são as estratégias que costumam ser criticadas como manipulativas, excessivamente interventivas e indevidamente paternalistas. Outras medidas que fazem parte da definição mais ampla de Sunstein – como a simplificação de procedimentos burocráticos, o fornecimento de informações factuais ou *nudges* educativos e que promovem a deliberação – geram poucas preocupações e são menos interventivas.

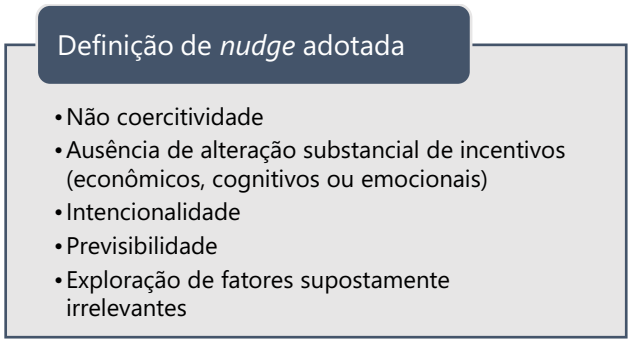
Nesse contexto, em termos metodológicos, parece razoável optar por uma definição estrita de *nudge*. Afinal, se um *nudge* que explora vieses comportamentais for razoável – como se defenderá nesta dissertação –, outras medidas menos interventivas, como o fornecimento de informações factuais, não gerariam qualquer tipo de preocupação. Em outras palavras, se o que se entende por *nudge* nesta dissertação for considerado legítimo, as demais estratégias que fazem parte da definição ampla também serão. Trata-se de um argumento *a majori ad minus*.[212]

Dessa forma, o foco da análise, nessa abordagem, estará voltado às intervenções que exploram vieses comportamentais. Estas serão as estratégias que serão consideradas *nudges* na abordagem adotada nesta dissertação. Em uma definição ampla de *nudge*, outra forma de definir os *nudges* que serão analisados nesta dissertação seria limitar a análise às categorias que se orientam por mecanismos psicológicos automáticos. Por exemplo, os *nudges* de 2º e de 3º grau de Baldwin (2014); os *nudges* 2 de Mongin e Cozic (2018); os *nudges* que desencadeiam heurísticas específicas de Barton e Grüne-Yanoff (2015, p. 343–344); os *nudges* de tipo 1 de Hansen e Jespersen (2013); e, na linha do próprio Sunstein (2015b, p. 426–427), os *nudges* que exploram vieses comportamentais ou *nudges* ligados ao Sistema 1.

[212] Como expõe Carlos Maximiliano (2017, p. 297–298), em sua obra clássica, o argumento *a majori ad minus* é justificado pelo brocardo latino *In eo quod plus est semper inest et minus* ("Quem pode o mais, pode o menos"). A partir desse argumento, aplica-se "as partes a regra feita para o todo, e julga lícito, ou exigível, o menos quando o texto autoriza, ou obriga, ao mais" (MAXIMILIANO, 2017, p. 298).

A imagem abaixo sumariza as condições necessárias para que uma intervenção seja um *nudge* na abordagem adotada nesta dissertação:

Figura 10 - Definição de *nudge* adotada na dissertação



Fonte: Elaboração do Autor.

3.5. Principais conclusões do Capítulo 3

O objetivo deste Capítulo foi expor o desenvolvimento do conceito de *nudge* enquanto técnica de regulação do comportamento privado. Inicialmente, foram mapeados textos prévios à elaboração do conceito de *nudge*, para expor as intervenções que foram reunidas sob essa definição. Em seguida, foram discutidos: a) a elaboração do conceito de *nudge* por Thaler e Sunstein; e b) os refinamentos conceituais e diferentes categorias de *nudges* identificadas na literatura. Por fim, demonstrou-se por que esta dissertação adota uma definição estrita de *nudge* ligada à neutralidade racional. As principais conclusões deste Capítulo foram as seguintes:

1. Antes da elaboração do conceito de *nudge*, as principais propostas de regulação do comportamento privado a partir de *insights* comportamentais envolviam a tentativa de *explorar* traços do comportamento humano, que tradicionalmente poderiam criar prejuízos, para tentar aprimorar o processo de tomada de decisão. Os exemplos mais comuns eram alterações de regras *default* e de enquadramentos (*framing*), e os artigos que discutiam essas medidas elencavam expressamente os fatores comportamentais explorados. Outras medidas bastante discutidas eram períodos de reflexão (*cooling-off periods*), que, em vez de explorar elementos comportamentais, serviam para tentar reduzir a sua influência no processo de decisão e se aproximam dos *sludges*, que serão discutidos no Capítulo 4;

2. Além da não coercitividade, da ausência de alteração substancial na estrutura de incentivos, da previsibilidade e da intencionalidade, a definição original de *nudge* proposta por Thaler e Sunstein parecia conter um componente de neutralidade racional, isto é, não incluía medidas que pudessem ser relevantes para o *Homo economicus*. No desenvolvimento posterior de sua teoria, Sunstein ampliou a definição de *nudge*, e, em sua definição ampla, haveria apenas duas condições necessárias para que uma intervenção fosse um *nudge*: a não coercitividade e a ausência de alteração substancial na estrutura de incentivos;

3. As indefinições em torno do conceito de *nudge* geraram enorme debate na literatura. Duas principais tentativas de refinar essa definição foram identificadas na literatura. De um lado, na linha de Hansen (2016), há quem defenda que *nudges* são intervenções racionalmente neutras, que, além da não coercitividade e da ausência de mudanças de incentivos, são baseados em falhas do comportamento e exploram ativamente essas falhas. Em uma definição ainda mais reduzida, Saghai propõe que *nudges* sejam intervenções substancialmente não controladoras que devem garantir fácil resistibilidade aos sujeitos que pretendem agir de outra forma. Intervenções que não cumprissem esse requisito seriam *cutucões comportamentais* e não mais *nudges*. Por outro lado, há ainda autores que tentam classificar os *nudges* em diferentes categorias, a partir do mecanismo psicológico explorado pela intervenção ou da técnica pela qual o *nudge* específico é implementado;

4. A definição ampla de *nudge* reúne um conjunto bastante heterogêneo de medidas, com implicações normativas muito diversas. As tentativas de defender a utilização de *nudges* com base nessa definição acabam suavizando as críticas, em vez de enfrentá-las diretamente. A definição de Saghai, por outro lado, ao considerar *nudges* como técnicas substancialmente não controladoras, incorpora um juízo de valor que parece inadequado à definição de uma técnica de intervenção, que, por natureza, é valorativamente neutra. Por fim, como os *nudges* que geram preocupações e dilemas éticos exploram fatores supostamente irrelevantes, adotar a definição estrita da racionalidade neutra é o caminho mais adequado a esta dissertação, já que se essas intervenções forem legítimas, as demais estratégias que fazem parte da definição ampla também serão.

No entanto, apresentar uma definição de *nudge* apenas com base em outras tentativas de definir essa técnica de intervenção, ou classificá-la em diferentes categorias, não é suficiente para ter clareza e segurança a fim de empreender uma análise da legitimidade dos *nudges*. Como

já mencionado, inúmeras outras estratégias – como técnicas de desenviesamento, *boosts*, *thinks*, *nudges plus* e *sludges* – foram descritas nos últimos anos. Há, ainda, algumas sobreposições entre algumas dessas técnicas e os *nudges*. Por outro lado, *nudges* são frequentemente confundidos com o paternalismo libertário e separá-los é uma tarefa importante à análise normativa que será empreendida. Por fim, como *nudges* também podem ser implementados por agentes privados, é preciso deixar claro o que se entende por *nudge* enquanto técnica de regulação estatal. Todas essas questões serão discutidas no Capítulo 4.

4. AVANÇOS TEÓRICOS E A DEFINIÇÃO DE *NUDGE*: ESTRATÉGIAS ALTERNATIVAS, RELAÇÃO COM O PATERNALISMO LIBERTÁRIO E ARQUITETURA DE ESCOLHAS PRIVADA

No Capítulo 3, foram apresentadas as razões pelas quais esta dissertação adota uma definição estrita de *nudge*, em que são considerados *nudges* técnicas não coercitivas, que, sem alteração substancial da estrutura de incentivos, direcionam o comportamento de maneira previsível e intencional, a partir da exploração de fatores supostamente irrelevantes. No estágio atual da literatura, no entanto, essa definição, por si só, não é suficiente para ter clareza a respeito de algumas questões adicionais que podem ser importantes à avaliação da legitimidade dos *nudges*.

Inicialmente, como já mencionado, ao lado dos *nudges*, outras técnicas para influenciar escolhas e o comportamento a partir de *insights* das ciências comportamentais vêm sendo discutidas na literatura. É o caso, por exemplo, de *thinks*[213], *boosts*[214], *sludges*[215] e *nudges plus*[216]. Ao mesmo tempo, já desde o início do desenvolvimento da economia comportamental e antes da elaboração do conceito de *nudge*, técnicas de desenviesamento vêm sendo discutidas na literatura[217]. A distinção entre essas técnicas, nem sempre baseadas na arquitetura de escolhas, e os *nudges* nem sempre é clara. Sunstein, inclusive, chega a admitir que algumas dessas técnicas podem ser consideradas *nudges* em sua definição ampla[218].

Como fazer uma avaliação de custos e benefícios de tais estratégias – além, é claro, das estratégias tradicionais, como multas, subsídios, impostos, proibições, obrigações etc. –, sem clareza a respeito dos elementos que as definem? Para complicar ainda mais essa tarefa, parte dos autores que discutem *boosts*, *sludges* e técnicas de desenviesamento não enxerga as mesmas

[213] Cf., por exemplo: (JOHN, 2018; JOHN *et al.*, 2011; JOHN; SMITH; STOKER, 2009; JOHN; STOKER, 2019).
[214] Cf., por exemplo: (GRÜNE-YANOFF; HERTWIG, 2016; HERTWIG, 2017; HERTWIG; GRÜNE-YANOFF, 2017).
[215] Cf., por exemplo: (MILLS, 2020; SUNSTEIN, 2019b, 2020a, 2021a; THALER, 2018a).
[216] Cf., por exemplo: (BANERJEE *et al.*, 2021b; BANERJEE; JOHN, 2021; JOHN, 2018; JOHN; STOKER, 2019).
[217] Cf., por exemplo: (FISCHHOFF, 1982; JOLLS; SUNSTEIN, 2006; STEFFEL; WILLIAMS; POGACAR, 2016).
[218] Como será demonstrado, Sunstein (2020a, p. 7, 2021a, p. 9-10) admite, por exemplo, que, em alguma medida, *sludges* podem ser definidos como fricções não excessivas e, dessa forma, poderiam ser definidos como uma espécie de *nudge*. Em outra oportunidade, Sunstein, juntamente com Reisch (2019, p. 97), menciona que alguns tipos de *nudges*, que tentam aprimorar a capacidade das pessoas para escolherem por conta própria, são às vezes descritos como *boosts*. Sunstein (2018b, p. 64) afirma, ainda, que alguns *nudges* contrapõem-se a vieses comportamentais e, em alguma medida, poderiam ser considerados formas de desenviesamento.

sobreposições vistas por Sunstein[219]. Nesse caso, antes de avaliar a legitimidade dos *nudges* é necessário avaliar cada uma dessas estratégias e compará-las com a definição de *nudge* adotada nesta dissertação.

Em outra perspectiva, *nudges* são frequentemente confundidos com o paternalismo libertário. Como discutido no Capítulo 2, para Thaler e Sunstein, uma das concepções equivocadas do movimento antipaternalista – e que constitui parte do argumento em favor do paternalismo libertário – é a crença de que o paternalismo está sempre atrelado à coerção. Dessa forma, o próprio surgimento do paternalismo libertário está ligado à proposição de técnicas de intervenção não coercitivas, que, em larga medida, confundem-se com *nudges*. Vale ressaltar, mais uma vez, como discutido no Capítulo 3, que o conceito de *nudge* foi sugerido inicialmente por um editor que avaliava o livro que Thaler e Sunstein escreveram para discutir o paternalismo libertário. Há, portanto, uma forte ligação entre o paternalismo libertário e a prescrição de *nudges*.

Apesar disso, *nudges* nem sempre serão paternalistas libertários. Há diversos fundamentos de intervenção que podem ser buscados a partir de *nudges*. Ao mesmo tempo, na definição estrita ligada à neutralidade racional adotada nesta dissertação, há políticas do paternalismo libertário que não são consideradas *nudges*. Diante da relação próxima entre *nudges* e o paternalismo libertário, diferenciá-los é uma tarefa essencial à análise normativa que será empreendida no Capítulo 6. Afinal, uma política implementada por motivos paternalistas gera desafios particulares, que podem não existir diante de outros fundamentos de intervenção.

Por fim, também é necessário deixar claro o que se pretende dizer quando se consideram estímulos comportamentais como técnicas regulatórias. Thaler e Sunstein, ao propor o conceito de *nudge*, deixam claro que essas estratégias podem ser implementadas tanto por agentes privados quanto pelo governo. Sunstein (2015b, p. 424) também ressalta que pode haver uma ordem coercitiva do Estado para que um agente privado implemente um determinado *nudge* e altere a arquitetura de escolhas. Quando, então, *nudges* são puramente privados e quando se está diante de uma técnica de regulação? Esclarecer essa questão é uma etapa fundamental para a definição do escopo de intervenções cuja legitimidade será avaliada no Capítulo 6.

[219] Rebonato (2014), por exemplo, critica Sunstein e Thaler por prestarem pouca atenção à literatura a respeito de técnicas de desenviesamento, isto é, estruturas decisórias que tornam as pessoas menos suscetíveis à manipulação das escolhas. Nitidamente, para Rebonato, *nudges* não são técnicas de desenviesamento. Ralph Hertwig e Grüne-Yanoff (2017), por outro lado, como ainda será discutido em maiores detalhes, até consideram que há alguma sobreposição entre *nudges* educativos e *boosts* de curto prazo; ressaltam, contudo, que *boosts* de longo prazo são bastante distintos de *nudges*. Chris Mills (2020), por seu turno, em sua discussão da simetria entre *nudges* e *sludges*, conforme ainda será discutido, propõe uma distinção conceitual que elimina qualquer sobreposição entre essas duas estratégias.

Todas essas questões, que representam avanços teóricos em torno da definição de *nudge*, serão discutidas neste Capítulo.

4.1. Desenviesamento, *boosts*, *thinks*, *nudges plus* e *sludges*: estratégias alternativas aos *nudges*?

Conforme apontado acima, não só para chegar a uma conclusão a respeito da definição de *nudge* adotada nesta dissertação, como também para a adequada comparação entre as diferentes técnicas de intervenção informadas por *insights* comportamentais, é conveniente tentar extrair os elementos que as definem. Sem pretender esgotar o tema ou conduzir uma ampla revisão da literatura a respeito de tais intervenções, nesta seção, serão apresentadas as principais formulações e características de técnicas de desenviesamento, *thinks*, *boosts*, *sludges* e *nudges plus*. As sobreposições entre todas essas estratégias e as suas distinções com relação aos *nudges* serão apresentadas ao longo desta exposição.

4.1.1. Técnicas de desenviesamento: a tentativa de corrigir os vieses comportamentais

Quando a economia comportamental passou a diagnosticar inúmeros fatores supostamente irrelevantes que interferem no comportamento, como demonstrado no Capítulo 1, é natural que a robustez de seus experimentos passasse a ser questionada. Para testar os limites desses fatores, é comum a utilização de testes destrutivos (*destructive testing*), em que o desenho de um experimento é submetido a determinadas condições a fim de guiar esse desenho aos seus limites de viabilidade (FISCHHOFF, 1982, p. 422). Esse teste permite que se esclareça em que circunstâncias um resultado experimental pode ser confiável, e por que ele funciona quando feito daquela forma específica.

Quando o fenômeno avaliado é um viés comportamental, testes destrutivos são esforços de desenviesamento: ao mostrar quando um desenho experimental que expõe um viés cognitivo falha, o teste destrutivo resulta em um aprimoramento dos julgamentos (FISCHHOFF, 1982, p. 422). Essa é a definição original de desenviesamento e inúmeros experimentos foram feitos desde o início da economia comportamental para tentar reduzir a influência de vieses comportamentais. Assim como o estudo dos vieses esclarece as fontes e os limites da aparente sabedoria das pessoas, o estudo do desenviesamento torna mais claros os limites e as fontes dos erros de comportamento (FISCHHOFF, 1982, p. 423). A redução de vieses comportamentais, como a falácia da taxa-base e a falácia da conjunção, a partir da exposição de problemas em

termos frequentistas, como discutido no Capítulo 1[220], é um exemplo de técnica de desenviesamento.

De acordo com Fischhoff (1982, p. 423), quando um erro de julgamento é identificado, é natural procurar pelo culpado. A responsabilidade pelos vieses pode, em uma primeira perspectiva, ser um resultado das tarefas a que os participantes de um experimento foram submetidos. Pode ser, por exemplo, que essas tarefas sejam injustas para apontar a conclusão feita pelo experimento[221]; ou, simplesmente, os participantes podem não ter compreendido as tarefas que deveriam cumprir. Por outro lado, a origem do viés pode estar no próprio julgador: de fato, pode ser que o sujeito tenha um desempenho ruim em uma determinada tarefa, em razão de seus próprios processos cognitivos ou de sua falta de habilidade. Por fim, existe a possibilidade de que haja um descompasso entre o agente e a tarefa a ser desempenhada; isto é, embora o sujeito tenha as habilidades adequadas, elas não são utilizadas na tarefa a ser julgada.

Cada categoria causal de viés demandará estratégias específicas de desenviesamento. Diante de problemas nas próprias tarefas, o pesquisador pode tentar torná-las mais claras, usar melhores modelos de resposta ou perguntar menos questões. Se o problema está em vieses do próprio julgador, o pesquisador pode alertá-lo desse problema, fornecer *feedbacks* personalizados, treiná-lo, recalibrar suas respostas ou, mesmo, substituir o próprio julgador. Em casos de descompassos entre julgadores e tarefas, em outra perspectiva, o pesquisador pode tentar oferecer formulações alternativas, decompor o problema em múltiplas etapas, considerar situações alternativas ou, em último caso, confiar o problema a um *expert* em solucioná-lo.

Técnicas de desenviesamento nem sempre terão sucesso[222]. De qualquer forma, embora desenvolvidas inicialmente como estratégias puramente experimentais, técnicas de desenviesamento podem contribuir ao aprimoramento de julgamentos e decisões enviesadas. Por essa razão, essas técnicas podem ser incorporadas a políticas públicas que tentam direcionar comportamentos influenciados por fatores supostamente irrelevantes. Afinal, como esses fatores podem causar prejuízos aos decisores, e a opção ótima ao mercado pode ser explorá-los em detrimento dos decisores, intervenções governamentais podem ser necessárias para afastar a influência de fatores supostamente irrelevantes nas escolhas.

---

[220] Nesse sentido, Cf. (COSMIDES; TOOBY, 1996; GIGERENZER, 1991). Cf. nota 54 *supra*.

[221] Se, por exemplo, os sujeitos ficarem confusos com as instruções do experimento, não acreditarem nas premissas dadas pelos pesquisadores ou forem incapazes de expressar o que realmente sabem (FISCHHOFF, 1982, p. 422), a tarefa de julgamento pode não ser um meio adequado para se concluir que há um viés.

[222] Após rever a literatura a respeito de técnicas de desenviesamento relacionadas ao viés retrospectivo e à confiança excessiva, Fischhoff (1982, p. 440) concluiu que, de maneira geral, as tentativas eficazes de reduzir esses vieses envolviam a alteração da natureza psicológica das tarefas e da forma como os sujeitos as abordavam.

Jolls e Sunstein (2006) ponderam que, diante da constatação de traços de racionalidade limitada, são desenvolvidas regras e instituições que restringem ou simplesmente bloqueiam a escolha. Diante de produtos perigosos, cujos verdadeiros riscos não são bem apreendidos pelos consumidores, pode-se, por exemplo, elevar o padrão de responsabilidade dos fornecedores por danos derivados do produto, como ocorre no Brasil, ou, até mesmo, proibir parcial ou integralmente a sua utilização (JOLLS; SUNSTEIN, 2006, p. 207-208). Essas técnicas são conhecidas como isolamento (*insulating*) e buscam evitar o desenvolvimento de atos jurídicos perfeitos a partir da racionalidade limitada das pessoas.

Em vez de isolar as pessoas, uma estratégia mais promissora, para Jolls e Sunstein (2006, p. 200), seria enfrentar problemas de racionalidade limitada explorando diretamente o comportamento racionalmente limitado, para tentar ajudar as pessoas a reduzi-lo ou eliminá-lo. Em outras palavras, em alguns casos, seria desejável que o direito - do ponto de vista material ou procedimental - tentasse desenviesar os agentes que sofrem da racionalidade limitada.

O foco de Jolls e Sunstein (2006) recai sobre o direito material, mas há inúmeros exemplos de regras processuais que poderiam ser elaboradas com fins de desenviesamento.[223] Nesse sentido, são apresentadas intervenções elaboradas para enfrentar erros de julgamento, como o viés do otimismo, e para enfrentar desvios com relação à Teoria da Utilidade Esperada, como o efeito dotação (*endowment effect*).

Para enfrentar erros de julgamento, Jolls e Sunstein apresentam exemplos de desenviesamento no direito do consumidor e no direito comercial.

No âmbito do direito do consumidor, há, por exemplo, tentativas de reduzir o otimismo excessivo de consumidores quanto à materialização dos riscos de um produto, como o cigarro, a partir da exposição desse indivíduo a exemplos concretos dessa materialização. Nesse tipo de intervenção, a heurística da disponibilidade é explorada para conter os efeitos do viés do otimismo. Em outro exemplo, técnicas de *framing* são utilizadas para conter esse mesmo viés, ainda no âmbito de regras destinadas à segurança dos consumidores.[224]

---

[223] Apenas para exemplificar, o viés do autosserviço (*self-serving bias*) é a tendência de autores de ações, em litígios civis, terem expectativas de indenizações sistematicamente mais elevadas do que os réus acreditam (BABCOCK; LOEWENSTEIN; ISSACHAROFF, 1997). Essa tendência acaba criando um impasse para acordos. Linda Babcock, Loewenstein e Issacharoff (1997) demonstraram a possibilidade de reduzir esse viés (desenviesamento), quando os sujeitos analisam conscientemente as fragilidades de seu caso ou as razões pelas quais um juiz poderia decidir em seu desfavor. Nesse caso, como mencionam Jolls e Sunstein (2006, p. 201), uma alteração procedimental seria implementada para fins de desenviesamento.

[224] Se as mulheres são excessivamente otimistas quanto à probabilidade de não contraírem câncer de mama – e, assim, subestimarem o valor de autoexames –, seria recomendável destacar as consequências negativas de não fazer esses exames, em vez de apenas descrever os benefícios correspondentes (JOLLS; SUNSTEIN, 2006, p. 210-211).

No direito comercial, por outro lado, Jolls e Sunstein apontam que a exigência de participação de diretores independentes no conselho de administração de uma companhia é uma medida de desenviesamento: diretores internos são mais otimistas, já que sua autocompreensão e autoestima estão ligadas aos resultados da empresa; ao mesmo tempo, a deliberação com pessoas também internas faria com que esses diretores acabassem em posições ainda mais otimistas, já que a deliberação entre pessoas de mesmo pensamento acaba levando-as a caminhos mais extremos. A inserção de diretores independentes, externos, teria como possível resultado, então, a diminuição do viés do otimismo[225].

Em outra perspectiva, as técnicas para enfrentar desvios da Teoria da Utilidade Esperada estão ligadas ao direito de propriedade e aos relacionamentos de agência e representação[226].

A princípio, Jolls e Sunstein (2006, p. 220) mencionam que a propriedade intelectual pode se organizar sob duas estruturas distintas: a) com regras de propriedade, em que o titular não precisa ceder o uso do seu bem intelectual, a não ser que concorde voluntariamente em fazê-lo - abordagem americana; e b) sob um regime de responsabilidade, em que o titular é obrigado a ceder o bem intelectual, desde que receba o pagamento de uma indenização legalmente especificada – abordagem europeia. Segundo Jolls e Sunstein, a abordagem europeia é, de certa forma, uma medida de desenviesamento: na presença de tais regras, o efeito dotação é diminuído e há uma divergência menor entre o valor que os titulares esperam receber pela utilização do seu bem e o valor que os potenciais usuários estão dispostos a pagar.

Os relacionamentos de agência e representação, por outro lado, reduzem o efeito dotação na perspectiva dos fornecedores: se os produtos são comercializados pessoalmente pelo fornecedor, ele estará disposto a receber, em geral, valores mais elevados do que os consumidores estão dispostos a pagar. O relacionamento de agência modifica essa equação. Quando agem como gerentes de uma empresa no contexto negocial, as pessoas não apresentam a mesma medida de efeito dotação e há um aumento na disposição a aceitar uma oferta (JOLLS; SUNSTEIN, 2006, p. 223).

Apresentados esses exemplos, verifica-se que os limites entre uma medida de desenviesamento e um *nudge* são bastante turvos.

---

[225] No direito brasileiro, é obrigatória a participação de conselheiros independentes na composição do conselho de administração das companhias abertas, nos termos e nos prazos definidos pela Comissão de Valores Mobiliários, como previsto no § 2º, do art. 140 da Lei de Sociedades Anônimas (BRASIL, 1976).

[226] Evidentemente, analisar os traços de diferentes modalidades contratuais ultrapassa o escopo desta dissertação. O objetivo é simplesmente destacar o relacionamento em que um fornecedor vende seus produtos por intermédio de um terceiro.

Nos exemplos citados acima, a utilização de *framing* e a exploração da heurística da disponibilidade, em ambos os casos com o fim de reduzir o otimismo excessivo, parecem se enquadrar na definição de *nudge* adotada nesta dissertação. Por outro lado, exigir a participação de conselheiros independentes no conselho de administração de uma sociedade não parece ser um *nudge*; não há, afinal, a tentativa de influenciar uma escolha específica em uma determinada direção. Mais do que isso, embora Jolls e Sunstein considerem essa intervenção uma técnica de desenviesamento, os fundamentos de uma regra como essa parecem mais ligados à necessidade de evitar conflitos de interesse do que a uma possível redução do viés do otimismo. A redução do viés, nesse caso, é secundária e um subproduto da redução dos conflitos de interesse.

A definição do regime de direito de propriedade intelectual também não deve ser considerada um *nudge*. Nesse caso, as regras são cogentes e o titular, em geral, terá que observá-las. Trata-se de uma escolha normativa mais ligada à finalidade promovida pelo direito de propriedade intelectual – estímulo à inovação contínua – do que a uma falha de mercado comportamental. A mesma afirmação vale para os relacionamentos de agência. Até seria possível imaginar *nudges* que incentivassem a celebração de contratos de agência. No entanto, atuar como representante de um fornecedor não é, por si só, uma intervenção estatal no comportamento privado nem tenta influenciar uma escolha específica. Eventuais limitações de vieses que decorram desse relacionamento contratual são secundárias e indiretas.

De todo modo, parece haver alguma sobreposição entre *nudges* e técnicas de desenviesamento. Essa sobreposição fica ainda mais clara quando Jolls e Sunstein (2006, p. 211-212) afirmam que, em sua visão de desenviesamento – mais conservadora –, a alteração de incentivos não deveria ser considerada uma medida de desenviesamento. Afinal, se um viés comportamental é eliminado com o fornecimento de incentivos financeiros, pode ser que o aparente desvio da racionalidade clássica não seja, na verdade, um verdadeiro desvio; seria o simples resultado de um agir descuidado do agente, que não tinha incentivos para se concentrar. Em sua concepção, o foco de medidas de desenviesamento recai sobre a reestruturação do ambiente sem alterar a motivação dos indivíduos, mas sim o processo pelo qual eles apreendem o seu entorno. Nitidamente, há uma clara semelhança com a alteração da arquitetura de escolhas, um dos elementos definidores de um *nudge*, como exposto no Capítulo 3.

Além disso, Jolls e Sunstein (2006, p. 211-212) reconhecem que, em contraposição às técnicas de desenviesamento por meio do direito, outras técnicas se dedicariam ao enviesamento por meio do direito. Em algumas circunstâncias, ao elaborar uma política pública, o Estado pode desejar, por exemplo, que as pessoas superestimem a probabilidade de os agentes serem detectados ao dirigir acima dos limites de velocidade ou sonegar impostos. Essas

estratégias, para Jolls e Sunstein, não seriam formas de desenviesamento, já que não responderiam a nenhuma forma de racionalidade limitada.

Feitas essas considerações, é possível concluir, a princípio, que *nudges* não se opõem exatamente a técnicas de desenviesamento. Na verdade, estas técnicas não são propriamente uma espécie de técnica de regulação do comportamento privado. O desenviesamento é, simplesmente, um resultado que pode ser buscado por inúmeras técnicas de intervenção distintas, inclusive *nudges*.

Nesse contexto, se o *policy-maker* deseja desenviesar o comportamento de determinados indivíduos, pode optar por adotar um *nudge*, como a mudança de enquadramentos (*framing*) ou a exploração da heurística da disponibilidade. Ao mesmo tempo, pode optar por outras estratégias, como medidas de comando e controle que tenham por efeito indireto a redução dos efeitos de um determinado viés - *e.g.*, exigir que administradores independentes participem do conselho de administração. Há, ainda, como demonstrado por Fischhoff (1982), um amplo leque de medidas que poderiam ser adotadas para o desenviesamento, como a educação e, até mesmo, a substituição do agente que tomaria a decisão. Todas essas estratégias seriam medidas de desenviesamento.

Por outro lado, se os vieses comportamentais enfrentados forem muito penetrantes e o desenviesamento for improvável, medidas de isolamento (*insulating*) podem ser recomendadas. No âmbito do direito do consumidor, por exemplo, ainda que se utilizem técnicas de desenviesamento, técnicas de isolamento, como a responsabilidade objetiva do fornecedor, podem ser necessárias. Técnicas de isolamento, nitidamente, não se confundem com *nudges*.

Ressalta-se, ainda, que há uma particularidade no desenviesamento promovido por *nudges*. De fato, na definição estrita adotada nesta dissertação, *nudges* sempre envolverão a exploração de vieses comportamentais e acabarão desencadeando heurísticas específicas. Se o sujeito é excessivamente otimista, por exemplo, um *nudge* que explora a heurística da disponibilidade pode reduzir o efeito do viés do otimismo; o decisor, no entanto, não estará completamente desenviesado, já que atuará com base na heurística da disponibilidade. Com relação ao viés que *justifica* a intervenção, no entanto, ele poderá ser desenviesado.

Há, ainda, *nudges* que não têm por objetivo o desenviesamento do agente decisor. Atuam, na verdade, enviesando – ou reenviesando – o agente decisor. Considere, por exemplo, a adoção de uma determinada regra *default* para inscrever os agentes automaticamente em determinado plano de previdência. Evidentemente, há fatores supostamente irrelevantes envolvidos na falta de contribuições à previdência que antecede o *nudge*: miopia para o presente, viés do *status quo*, inércia e procrastinação são alguns dos exemplos já amplamente

mencionados nesta dissertação. O *nudge*, no entanto, não corrigirá os vieses – afinal, os sujeitos continuarão enviesados ao presente, a inércia continuará tendo efeitos significativos e as pessoas continuarão a procrastinar. Em vez disso, o *nudge* muda as consequências de todos esses fatores supostamente irrelevantes. Nesses casos, *nudges* não serão técnicas de desenviesamento.[227]

Em suma, *nudges* serão medidas de desenviesamento quando tiverem por objetivo reduzir a influência de um viés comportamental sobre o comportamento do agente. Em todo caso, na definição estrita adotada nesta dissertação, nudges sempre envolverão alguma medida de enviesamento. Ao mesmo tempo, é possível desenviesar sem contar com os *nudges*, a partir de medidas de comando e controle, da educação, da substituição dos agentes decisores e de diversas outras medidas. Finalmente, há *nudges* que, em vez de desenviesar, alteram o resultado de escolhas motivadas por vieses comportamentais e, assim, exploram-nos ativamente.

4.1.2. *Boosts* e o aprimoramento das capacidades dos agentes

Outra estratégia que ganhou popularidade nas ciências comportamentais são os *boosts*. Essas intervenções buscam aumentar a competência na tomada de decisões dos indivíduos, com o objetivo de desenvolver suas habilidades e seu conhecimento, aprimorar o conjunto disponível de ferramentas para tomada de decisões ou reestruturar o ambiente em que as escolhas são feitas, para que essas habilidades e ferramentas possam ser aplicadas com maior efetividade (HERTWIG; GRÜNE-YANOFF, 2017).

Como esclarecem Till Grüne-Yanoff e Hertwig (2016), há pelo menos três classes distintas de intervenções que poderiam ser consideradas *boosts*, baseadas, respectivamente: a) na alteração do ambiente em que escolhas são feitas; b) na extensão do repertório de estratégias, habilidades e conhecimento da tomada de decisão; ou c) na união dos dois elementos anteriores. Todas elas, no entanto, reúnem-se sob um mesmo denominador comum: o objetivo de expandir as competências (*boost*) dos agentes, a fim de que eles possam atingir seus próprios objetivos por conta própria.

---

[227] Mary Steffel, Elanor F. Williams e Ruth Pogacar (2016), inclusive, avaliaram experimentalmente as possibilidades de *desenviesar* uma decisão resultante da alteração de uma regra *default* e constataram que a divulgação da intenção por trás de tais intervenções não é suficiente para afastar sua influência, ainda que os *defaults* sejam vistos como antiéticos, como será discutido no Capítulo 6. Apesar disso, os autores constataram que uma intervenção ativa pode ter efeito de desenviesamento: quando, além da divulgação das intenções por trás da intervenção, os sujeitos articulavam suas preferências prévias ativamente, a influência do *default* diminuía e os participantes conseguiam decidir de maneira menos enviesada. Essa intervenção ativa se assemelha aos *nudges plus* que serão discutidos abaixo.

Para exemplificar o primeiro conjunto de estratégias, Grüne-Yanoff e Hertwig (2016, p. 10) mencionam o aprimoramento das capacidades de inferências bayesianas, a partir da representação por meio de frequências naturais, tema já discutido no Capítulo 1.[228] Essa técnica, amplamente explorada pelo programa de heurísticas rápidas e frugais de Gigerenzer, embora possa parecer com uma mudança de enquadramento (*framing*), permite que as pessoas entendam o raciocínio estatístico e tomem, assim, decisões bem-informadas.

Considere, por exemplo, que a eficiência de um tratamento pode ser representada em termos de redução absoluta ou relativa de riscos[229], e que representá-la em termos relativos faz com que mais agentes passem pelo tratamento. Um *nudge* nesse contexto, segundo Grüne-Yanoff e Hertwig, tentaria explorar esse fator supostamente irrelevante, para fazer com que as pessoas superestimassem a redução de riscos gerada pelo procedimento. *Boosts*, em contrapartida, tentariam aprimorar a alfabetização estatística desses mesmos agentes, para que eles pudessem compreender representações confusas e enganosas, tornando-as menos manipulativas e menos exigentes computacionalmente. Em vez de direcionar as pessoas a um determinado tratamento, *boosts* tentariam ajudá-las a entender as informações estatísticas e tomar uma decisão mais alinhada às suas próprias preferências.

Por outro lado, em vez de alterar a arquitetura de escolhas para estimular maior compreensão, como no exemplo da utilização de representações frequentistas em vez de percentuais, *boosts* também podem focar em identificar e corrigir déficits de habilidades ou conhecimentos específicos que tenham consequências abrangentes para a saúde, a riqueza e a felicidade dos agentes (GRÜNE-YANOFF; HERTWIG, 2016, p. 11). O objetivo, nesse caso, não é educar ou informar os agentes de forma mais ampla, mas fornecer conhecimentos elementares que podem trazer um alto retorno a partir de uma pequena quantidade de informações. Quando se ensinam às pessoas os sintomas básicos de ataques cardíacos ou derrames, por exemplo, é possível diminuir consideravelmente os índices de morte ou de perda na qualidade de vida.

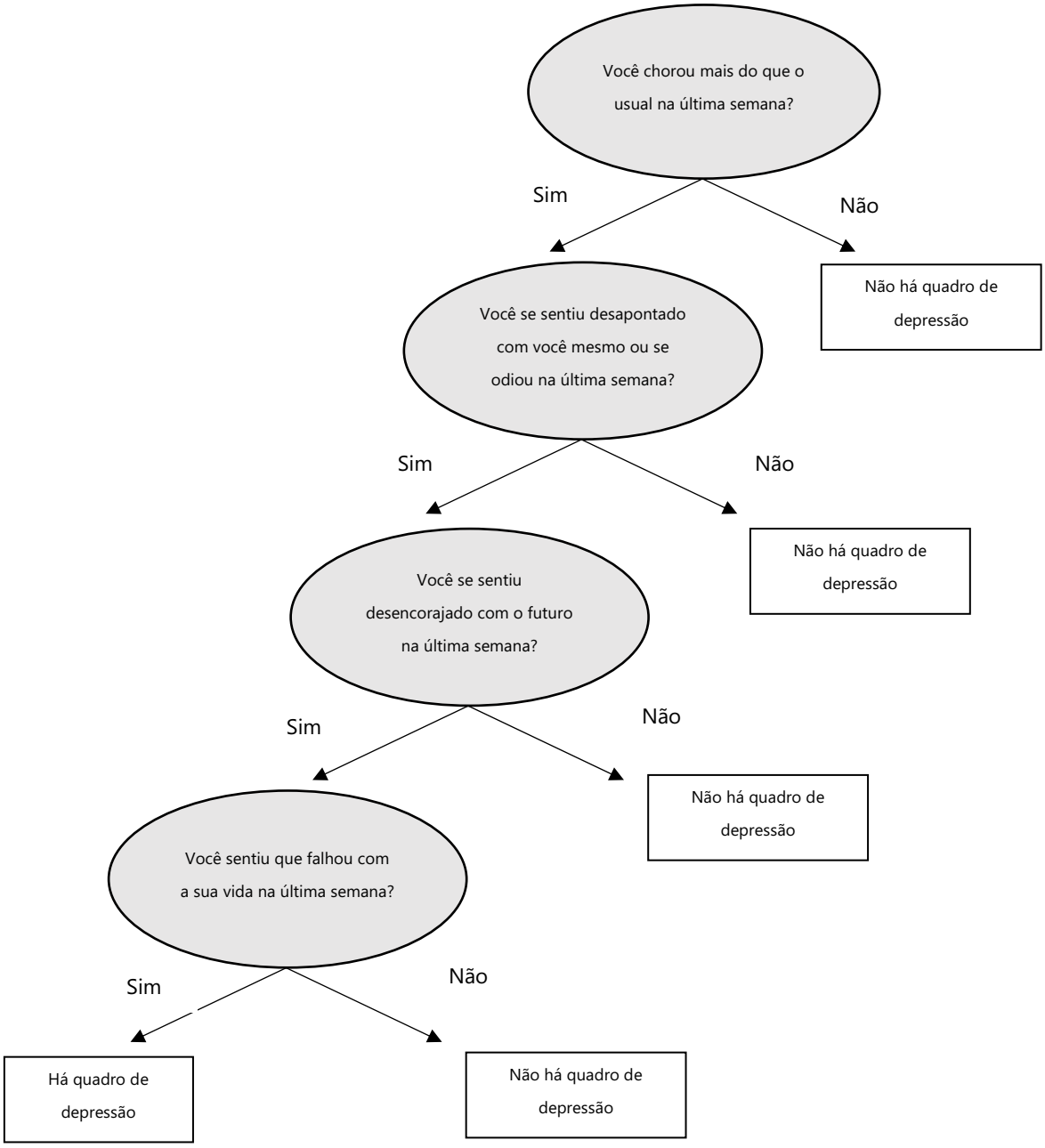
Por fim, no terceiro grupo de *boosts*, reunindo a alteração da arquitetura de escolhas e o fornecimento de conhecimentos elementares, há estratégias que buscam construir e ensinar heurísticas simples, intuitivas e eficientes para auxiliar na tomada de decisões em inúmeras situações em que o conhecimento sobre riscos específicos é incompleto e incerto (GRÜNE-

[228] Cf. a Subseção 1.2.2 acima.

[229] Em outras palavras, seria possível dizer que fazer um teste a cada dois anos reduziria suas chances de morrer de câncer em aproximadamente 1/3 ao longo dos próximos 10 anos (redução relativa de risco) ou dizer que essa mesma escolha reduziria suas chances de morrer de câncer de 3 em 1000 para 2 em 1000 ao longo dos próximos 10 anos (redução absoluta de risco).

YANOFF; HERTWIG, 2016, p. 14). A utilização de árvores rápidas e frugais – árvores de decisão que utilizam heurísticas para a tomada de decisão, em que, após cada nó de decisão, há apenas um nó terminal - é capaz de simplificar e acelerar o processo decisório. A imagem abaixo apresenta um exemplo de árvore rápida e frugal:

Figura 11 - Árvore de decisão rápida e frugal



Fonte: Grüne-Yanoff e Hertwig (2016, p. 13, tradução livre)

Outra classificação dos *boosts* divide-os em conformidade com as habilidades que estão sendo impulsionadas (HERTWIG; GRÜNE-YANOFF, 2017, p. 979). A princípio, *boosts* de

alfabetização de risco tentam estabelecer ou aprimorar as competências no entendimento de informações estatísticas, em domínios como a saúde, a meteorologia e as finanças. O exemplo da representação de dados estatísticos em frequências relativas enquadra-se nessa categoria. Por outro lado, também há *boosts* voltados ao gerenciamento da incerteza, que tentam estabelecer ou criar regras procedimentais para previsões, avaliações e tomada de decisões sob condições de incerteza. A árvore de decisão rápida e frugal é um exemplo de *boost* dessa categoria. Finalmente, *boosts* motivacionais tentam aprimorar a competência de controle cognitivo, autocontrole e ajustes autônomos da própria motivação. Aqui, enquadram-se estratégias como treinamento de atenção e de estados de atenção, estratégias de pré-compromisso e escrita expressiva.

Uma outra distinção entre *boosts* é essencial para diferenciá-los de *nudges*. Como já mencionado, embora *boosts* tenham como foco o aprimoramento das capacidades dos agentes, parte de suas estratégias envolvem alterações na arquitetura de escolhas. Em outras palavras, há de fato uma semelhança entre *nudges* e *boosts*, especialmente se adotada a definição ampla de *nudge*. Não por outra razão, como já mencionado, Sunstein e Reich (2019, p. 97) tratam *boosts*, em geral, como sinônimo de *nudges* educativos. Em outra oportunidade, Sunstein (2016a, p. 30) chega a dizer que *nudges* que tentam fortalecer o Sistema 2 são às vezes descritos como *boosts* e também incluiriam estratégias de desenviesamento. *Nudges* educativos seriam, então, tentativas de informar as pessoas ou torná-las competentes para que pudessem tomar decisões adequadas por conta própria (SUNSTEIN, 2016a, p. 32).

De qualquer forma, essa sobreposição não parece ser integral. Para evidenciar esse ponto, é útil distinguir *boosts* de curto prazo e *boosts* de longo prazo (HERTWIG; GRÜNE-YANOFF, 2017, p. 977). No primeiro caso, o aprimoramento de competências decisórias é limitado a um contexto específico. No caso da representação de estatísticas em frequências naturais, por exemplo, não há qualquer aspiração de aprimorar as capacidades dos agentes para além das tarefas específicas em que essa alteração é feita. Para Hertwig e Grüne-Yanoff, haveria uma sobreposição entre esses *boosts* de curto prazo e os *nudges* educativos. *Boosts* de longo prazo, no entanto, apresentam características distintas: no caso do raciocínio Bayesiano, para desenvolvê-lo em longo prazo, *boosts* tentariam aprimorar as capacidades das pessoas para traduzir ativamente quaisquer probabilidades que encontrassem para frequências e, assim, simplificar cálculos Bayesianos.

Essa distinção fica ainda mais clara se um *nudge* clássico, na linha da definição estrita adotada nesta dissertação, é levado em consideração e comparado a um *boost*. A inscrição

automática de indivíduos a planos de previdência, no *Save More Tomorrow*[230], por exemplo, claramente explora fatores supostamente irrelevantes, como o viés do presente e a inércia. Um *boost* com a mesma finalidade, ao invés disso, tentaria treinar os sujeitos para utilizarem regras de bolso simples que levassem a maior poupança, em vez de lhes ensinar conhecimentos teóricos, como juros compostos, inflação e diversificação de risco. Mais do que isso, um *boost* tentaria ainda conectar os sujeitos aos seus *selves* futuros e, assim, ensiná-los a evitar o viés do presente (HERTWIG; GRÜNE-YANOFF, 2017, p. 978).

Feitas essas considerações, verifica-se que há distinções entre as duas abordagens em algumas dimensões distintas.

Inicialmente, como Hertwig e Grüne-Yanoff (2017, p. 978) ponderam, *nudges* tentam alterar o comportamento, sem o objetivo primário de aperfeiçoar competências específicas de um agente, embora esse aprimoramento possa surgir como efeito secundário de um determinado *nudge*. Assim, diante de uma falha de mercado comportamental ou de outro motivo para intervenção estatal, um *nudge* buscaria alterar a arquitetura de escolhas, para explorar falhas cognitivas e deficiências motivacionais, e, assim, alterar o comportamento dos agentes. Esse comportamento, enquanto resultado final do processo decisório, é o elemento a ser influenciado pelo *nudge*. Um *boost*, por seu turno, investiria na construção e no desenvolvimento de habilidades para solucionar o mesmo problema.

Em outros termos, *nudges* e *boosts* exploram caminhos causais distintos para gerar mudança de comportamento. Para Hertwig e Grüne-Yanoff (2017, p. 979), *nudges* exploram os processos motivacionais e cognitivos internos do sujeito e, a partir disso, desenvolvem a arquitetura de escolhas externa, para que haja uma mudança de comportamento, mantidos esses processos funcionais intactos. Nesse caso, um *nudge* pode até ser educativo e gerar um aumento de competências, mas esse efeito é derivado de uma alteração prévia do comportamento. Um *boost*, por outro lado, tenta estimular o desenvolvimento de competências, a partir da reelaboração de aspectos do ambiente dos indivíduos ou de técnicas para ensiná-los a readequar esse ambiente. Ainda que também possa haver alteração na arquitetura de escolhas, a mudança de comportamento é o reflexo de um prévio aprimoramento das capacidades decisórias.

Por outro lado, enquanto *nudges* partem do paradigma de heurísticas e vieses desenvolvido por Kahneman e Tversky, *boosts* foram inicialmente associados ao programa de heurísticas simples e racionalidade ecológica (GRÜNE-YANOFF; HERTWIG, 2016),

[230] Vide nota 149 *supra*.

elaborado por Gerd Gigerenzer e seus seguidores[231]. No desenvolvimento posterior da teoria, Hertwig e Grüne-Yanoff (2017, p. 980) registraram que os proponentes dos *boosts* não adotam uma única visão de arquitetura cognitiva humana. Embora incluam a teoria das heurísticas simples e da racionalidade ecológica, *boosts* vão além, incluindo outras visões cuja análise ultrapassa os objetivos desta dissertação. Em linha com essas matrizes teóricas, *nudges* presumem uma arquitetura cognitiva baseada no sistema dual defendido por Kahneman – como já demonstrado no Capítulo 1; *boosts*, por outro lado, presumem processos cognitivos e motivacionais maleáveis, que podem ser desenvolvidos e aprimorados.

Há, ainda, uma distinção empírica entre *nudges* (não educativos) e *boosts* (HERTWIG; GRÜNE-YANOFF, 2017, p. 981). Considerando os traços definidores de cada uma das espécies de intervenção, haverá um grau diferente de reversibilidade nos efeitos comportamentais induzidos por cada uma delas. Se, eliminada a intervenção, o comportamento retornar ao estado anterior, para Hertwig e Grüne-Yanoff, é provável que a política seja um *nudge*. Caso contrário, se o comportamento for mantido, é provável que a intervenção seja um *boost*. Essa é uma questão empírica e que deve, evidentemente, ser testada. A comparação entre os efeitos – de curto e de longo prazo – dessas intervenções será discutida no Capítulo 5.

Por fim, Hertwig e Grüne-Yanoff (2017, p. 981) salientam que, embora possa haver alguma sobreposição entre *boosts* e *nudges* educativos, os objetivos mais amplos de cada uma dessas técnicas são diversos. *Boosts* teriam como objetivo final equipar os indivíduos com ferramentas estatísticas que podem ser aplicadas a circunstâncias muito diversas, que vão além dos casos englobados por técnicas de obrigação de divulgação de informações, avisos e rótulos (*labels*). O objetivo é dar poder a esses agentes, para permitir que eles decidam melhor em arquiteturas de escolhas comercialmente construídas, que nem sempre contarão com intervenções benevolentes do Estado. A ambição dos *nudges*, por outro lado, seria corrigir erros momentâneos em momentos específicos.

4.1.3. De *thinks* a *nudges plus*: incorporando elementos de deliberação e reflexão na tomada de decisões

Em contraste com os *nudges*, outra tentativa de influenciar o comportamento parte da premissa de que, dados o contexto e o enquadramento adequado, as pessoas podem chegar a

[231] O próprio Gigerenzer (2015), ao criticar a utilização de *nudges*, como será discutido no Capítulo 6, prescreve, em seu lugar, a utilização de medidas para aprimorar a educação estatística, que, em geral, estão associadas aos *boosts*.

um melhor entendimento dos problemas e obter soluções coletivas mais efetivas (JOHN; SMITH; STOKER, 2009, p. 2-3). A ideia é estimular a deliberação e o diálogo entre cidadãos, para que sejam tomadas decisões informadas sobre escolhas coletivas e sobre a direção da política pública. As intervenções, nesse contexto, são denominadas *thinks*, como desenvolvido por Peter John e outros pesquisadores.

Essa estratégia alternativa para transformação do comportamento humano emerge da ideia de que a legitimidade dos políticos se origina da deliberação pública entre cidadãos livres e iguais (JOHN *et al*., 2011, p. 11). Enquanto os *nudges* foram propostos a partir dos pressupostos da economia comportamental, os *thinks* são baseados no giro deliberativo que dominou a teoria democrática nas últimas décadas. Nessa abordagem, a legitimidade depende do fluxo livre de discussão e troca de ideias, num ambiente de respeito mútuo e de entendimento (JOHN; SMITH; STOKER, 2009, p. 8). O valor da deliberação, como expõem John, Graham Smith e Gerry Stoker (2009), está mais ligado à sua natureza pública do que à troca de informações; ao defender suas visões em público, os agentes conseguiriam eliminar suas preferências irracionais, baseadas em crenças empiricamente falsas, moralmente repugnantes, que não seriam defendidas por ninguém no debate público, ou puramente autointeressadas.

O orçamento participativo de Porto Alegre, segundo John, Smith e Stoker, seria um exemplo bem-sucedido de *think*. Nessa proposta, milhares de cidadãos participariam anualmente das decisões a respeito da forma e da alocação de investimentos do orçamento municipal. Essa política teria o efeito de contribuir à redistribuição da renda em benefício das comunidades mais pobres da cidade. Outro exemplo mencionado pelos autores é a Assembleia dos Cidadãos de British Columbia, no Canadá, que foi designada para fazer recomendações na reforma eleitoral da província. Essa assembleia teria gerado um nível enorme de motivação e compromisso por parte dos cidadãos selecionados.

A partir desse exemplo, John, Smith e Stoker (2009, p. 11) concluíram que as normas e os procedimentos de deliberação precisam ser alimentados e podem não emergir naturalmente, isto é, a falta de tais procedimentos não significa a ausência de interesse dos cidadãos em participar das deliberações públicas. Na verdade, parte da motivação dos sujeitos é simplesmente ter uma influência em decisões políticas significativas. Para pôr em prática essa motivação, no entanto, é preciso elaborar o *design* institucional para trazer ao mesmo espaço de debate pessoas de origens distintas e fazê-las debater questões diversas no mesmo fórum deliberativo. Inovações democráticas, como o uso de tecnologia da informação, poderiam ser usadas para promover essa orientação cívica.

Desde a proposta original, admite-se que *thinks* exigem um maior esforço para motivar os indivíduos do que *nudges* (JOHN; SMITH; STOKER, 2009, p. 14). Afinal, para adquirir informações e, em seguida, debater as propostas com os pares, os indivíduos acabam arcando com custos significativos. Apesar disso, para John, Smith e Stoker, diferentemente de *nudges*, *thinks* conseguiriam enfrentar a raiz dos problemas e levar a alterações mais duradouras do comportamento humano. Assim, em vez de uma postura individualista e tecnocrática, seria possível pensar em influenciar o comportamento de maneira mais atenta ao contexto coletivo da vida pública em que preferências podem ser formadas ou corrigidas.

Na prática, porém, é difícil implementar essas iniciativas, garantir que os *policy-makers* as utilizem ou, mesmo, executar as recomendações de exercícios deliberativos (JOHN, 2018, p. 11). De fato, como constata John, as pessoas que se propõem a participar de exercícios deliberativos acabam compondo uma amostra não representativa da população envolvida e incapaz de ser suficiente para gerar os resultados políticos necessários para uma mudança de comportamento massiva. Além de tudo, estratégias deliberativas também são caras. Por isso, não é possível afirmar automaticamente que *thinks* poderiam ser adotados como uma ampla estratégia para aprimorar o comportamento humano e a tomada de decisões.[232]

Mais do que isso, como demonstrado empiricamente na própria evolução desse programa de pesquisas, enquanto os *nudges* parecem entregar resultados razoáveis, é muito mais difícil concluir que *thinks* seriam efetivos.[233] Em um experimento, por exemplo, os pesquisadores compararam os efeitos de um *nudge* (fornecimento de informações), um *think* (discussão) e um placebo, para estimular a doação de órgãos entre estudantes. Suas conclusões indicaram que *nudges* tiveram o melhor resultado, seguidos dos placebos, enquanto os *thinks* tiveram o pior desempenho (JOHN, 2018, p. 126).

Em contraposição, como já mencionado, *nudges* são de fácil implantação, baratos e não envolvem muito esforço. Seus efeitos no comportamento, além disso, podem ser imediatos. Por isso, John (2018, p. 11) pensa que *thinks* parecem mais adequados para questões mais espinhosas, como a definição do local em que uma usina nuclear será instalada.

---

[232] Christopher H. Achen e Larry M. Bartels (2016), nesse sentido, demonstraram que o aumento da democratização direta, com a participação dos cidadãos em decisões conjuntas, em estratégias que poderiam ser consideradas *thinks*, como primárias diretas para a seleção de candidatos em eleições ou referendos, não gerava os resultados benéficos esperados. Um exemplo que os autores mencionam, com base em um experimento conduzido por Jeff Tessin (2009), envolve a proteção a incêndios no estado do Illinois, nos Estados Unidos. Nos anos 1990, alguns condados desse estado começaram a exigir a aprovação popular (referendos) para aumentos nos impostos sobre a propriedade para além da inflação. Em tais condados, o tempo de resposta a incêndios era aproximadamente 7% superior aos distritos que não exigiam referendos, embora as economias em termos de impostos fossem muito baixas. Em outras palavras, a participação popular levou a potenciais perdas de bem-estar por meio de um risco maior de incêndios.

[233] Nesse sentido, Cf.: (JOHN *et al.*, 2011).

Apesar disso, quando analisado o seu funcionamento prático, *nudges* seriam mais próximos de *thinks* do que as definições dão a entender. Segundo John (2018, p. 11–12), *nudges* iriam além de enviar mensagens ou criar regras *default*. Na verdade, para que *nudges* funcionassem de maneira adequada, os cidadãos precisariam compreender as escolhas implícitas dos *nudges* em termos de políticas públicas. Se, por exemplo, um *nudge* sugere que um sujeito pague os seus impostos tempestivamente, a intervenção funcionará melhor se as pessoas compreenderem de maneira adequada a função do pagamento de impostos e como ele funciona para garantir a prestação de serviços públicos. Esse tipo de reflexão, para John (2018, p. 12), pode simplesmente estar implícito na forma como os *nudges* funcionam na prática, ainda que os cidadãos não estejam o tempo todo pensando nos motivos e premissas de uma intervenção que lhes é transmitida.

A partir dessas premissas, John e outros autores ligados ao seu programa de pesquisas (BANERJEE; JOHN, 2021; JOHN, 2018; JOHN; STOKER, 2019) passaram a defender o rompimento da dicotomia entre *nudges* e *thinks*. Nessa abordagem, embora *nudges* costumem estar ligados ao funcionamento do Sistema 1, pode ser que outras partes do cérebro, incluindo os processos deliberativos do Sistema 2, também estejam envolvidas quando um *nudge* é posto em prática. Ao mesmo tempo, *nudges* exigiriam um entendimento prévio da situação de escolha, o que implica um grau de engajamento cognitivo e habilidade de entender nexos causais, o que exige um processamento cognitivo que se aproxima do Sistema 2 (JOHN, 2018, p. 127).

Garantir que algum grau de reflexão ou pensamento seja integrado à elaboração e à implementação de intervenções comportamentais seria o desafio dos *policy-makers* para garantir maior responsividade dos cidadãos (JOHN, 2018, p. 12). Essas estratégias, que seriam uma expansão da definição original de *nudge*, com o objetivo de incorporar elementos reflexivos e deliberativos por parte do indivíduo que será influenciado pelo *nudge*, receberam a denominação de *nudge plus* (JOHN, 2018, p. 13; JOHN; STOKER, 2019, p. 4). O objetivo seria amplificar as possibilidades de uso de *nudges*, para situações que levassem a mudanças verdadeiras e duradouras de comportamento, além de torná-los mais democráticos e promover a autonomia. Uma combinação de *nudges* e pensamento reflexivo poderia auxiliar na criação de um sistema automático e autorregulado, em que as pessoas atingem suas metas e em que há sinergia entre os objetivos sociais e os individuais (JOHN, 2018, p. 133).

Um *nudge plus*, então, é uma estratégia híbrida entre um *nudge* e um *think*, que combina um *nudge*, como uma regra *default*, com um mecanismo ativo, como um compromisso (*pledge*), que permite a reflexão e a alteração dos rumos da decisão, se assim o sujeito desejar

(BANERJEE; JOHN, 2021, p. 3). Pode-se pensar, por exemplo, num *default* direcionado à aquisição de alimentos mais saudáveis que, antes de ser efetivado, incorpora um mecanismo de compromisso a respeito das escolhas de alimentação futura, incluindo a possibilidade de optar por alimentos não saudáveis em um determinado momento. Antes de fazer a escolha, nesse caso, o sujeito poderia refletir a respeito da medida em que ela realmente se conforma às suas preferências.

Além disso, *nudges plus* não se confundem com *boosts*. Como já mencionado, *boosts* aprimoram as capacidades dos agentes decisores, ao permitir que eles utilizem suas heurísticas de forma mais inteligente. Ainda no contexto da alimentação saudável, uma regra rápida, como usar pratos menores ou combinar alimentos saudáveis com *junk food*, poderia ser vista como um *boost* (BANERJEE; JOHN, 2021, p. 4). Essa estratégia, no entanto, não envolve qualquer tipo de reflexão, o que deve acontecer para que se esteja diante de um *nudge plus*.

Para a elaboração e implementação de *nudges plus*, como explicam Sanchayan Banerjee e John (2021, p. 6-9), dois fatores podem ser alterados: o momento de incorporação do *plus* ao *nudge* e a estratégia pela qual esses elementos são combinados. O elemento reflexivo pode ser incorporado de maneira simultânea ao *nudge* ou de forma sequencial, isto é, antes ou depois da influência do *nudge*. Em ambos os casos, essa incorporação pode ser combinada em uma só intervenção ou em intervenções distintas, como uma combinação de ferramentas.

Inicialmente, o *nudge* e o *plus* podem ser combinados em uma só estratégia, de maneira sequencial ou simultânea. No caso de um GPS, por exemplo, o *plus* pode ser incorporado na própria intervenção, de forma simultânea: o sistema pode encorajar o sujeito a utilizar o mapa de maneira heurística, optando por ignorar os conselhos conscientemente, a partir de sua experiência do passado. Uma estratégia de autocompromisso duplo, por outro lado, envolve um compromisso para o presente, considerando o cenário atual e objetivos de curto prazo, assim como um compromisso futuro, que o sujeito planeja ao seu *self* do futuro. Um monitorador de atividades físicas, por exemplo, dá *feedbacks* do nível de atividade diário ou semanal do sujeito, enquanto pode, ainda, fazer com que o sujeito estabeleça objetivos futuros de atividade física. Com o tempo, novos *feedbacks* são oferecidos, permitindo que o sujeito ajuste os seus esforços, se quiser continuar seguindo o seu planejamento de longo prazo. Logo, em um primeiro momento, o sujeito deve deliberar para fazer o seu compromisso futuro; em seguida, ao comparar os objetivos planejados e aquilo que realmente foi efetivado, o sujeito pode fazer um novo compromisso, mais uma vez com um elemento reflexivo, que pode ser, inclusive, o aumento de seus esforços para cumprir o compromisso prévio. Os elementos reflexivos, então, são aplicados de maneira sequencial, mas como parte da mesma ferramenta.

Em outra perspectiva, um *nudge plus* pode ser desenvolvido como uma combinação de ferramentas – mais uma vez, de maneira simultânea ou sequencial. Um *nudge* clássico, nesse contexto, pode ser combinado com *feedbacks* visuais ativos, que aumentam a saliência para a utilização das informações. Um *nudge* como um sinal verde associado a uma determinada opção mais saudável, para estimular a sua escolha, por exemplo, pode ser combinado com alertas de informação que explicam explicitamente o que a cor significa, o que, portanto, estimularia a reflexão do agente. Esse seria um exemplo de *nudge plus* baseado na combinação simultânea de duas intervenções distintas. Essa combinação, por outro lado, também pode ser feita de forma sequencial. De fato, um *nudge* pode ser precedido ou seguido de um compromisso ativo do sujeito. Se, por exemplo, antes ou depois de visualizar as opções padrão do cardápio de um restaurante, o sujeito faz um compromisso de seguir uma dieta saudável, as características das opções padrão se tornarão mais salientes e as escolhas seriam motivadas pela própria reflexão.

Considerando que os *nudges plus* são um desenvolvimento recente da literatura - o termo foi cunhado em um texto de John e Stoker em 2017, que viria a ser publicado em 2019 (JOHN; STOKER, 2019) -, ainda não há clareza da medida em que eles podem ser mais efetivos e mais eficientes que *nudges*. De todo modo, de acordo com seus proponentes, *nudges plus* têm o benefício de restaurar a autonomia do consumidor e a sua agência (BANERJEE; JOHN, 2021, p. 4).

4.1.4. *Sludges* e a criação de fricções que dificultam determinadas escolhas

Em "Ninguém Escreve ao Coronel", Gabriel Garcia Márquez conta a história de um Coronel eternamente à espera da carta que lhe traria a aposentadoria a que tinha direito, enquanto a burocracia estatal da Colômbia - ou de Macondo - não lhe dava qualquer retorno. Além de viver na penúria com sua esposa, depositando as fichas de sua esperança no galo de briga que o filho lhe deixara, em todas as sextas-feiras, o Coronel ia ao correio em busca de sua carta e tinha que se contentar com a resposta do carteiro, de que ninguém escrevia ao coronel. Havia, portanto, uma enorme burocracia que impedia o Coronel de ter acesso ao seu benefício. Essa burocracia, assim como outras fricções que dificultam a tomada de decisões, é o que a literatura tem denominado *sludges*[234].

[234] O significado tradicional de *sludge* é algo como uma substância suja, molhada e pegajosa, similar à lama. Nitidamente, o termo foi escolhido para ter uma conotação negativa às fricções que os *sludges* representam.

O conceito foi introduzido, inicialmente, para representar características do ambiente decisório ou do próprio decisor que poderiam limitar a efetividade dos *nudges* (LAMBERTON, 2016). Essa não é, no entanto, a acepção exata que tem sido utilizada atualmente. Ao retomar o conceito e popularizá-lo, Thaler (2018a, p. 431) mencionou que algumas atividades seriam essencialmente *nudges* para o mal: os *sludges*. São atividades que prejudicam o ambiente de escolhas e tornam mais difícil a realização de boas escolhas e de atividades benéficas para a sociedade. Na linha de Thaler, *sludges* podem ocorrer tanto ao desencorajar comportamentos que estão no melhor interesse do decisor, quanto ao encorajar comportamentos autodestrutivos.

Em termos mais técnicos, como demonstrado por Sunstein (2020a, p. 4, 2021a, p. 4-5), *sludges* seriam fricções excessivas que criam dificuldades à tomada de decisões, normalmente pelo acréscimo de etapas complexas a processos necessários para que uma escolha seja feita. Assim como no exemplo do Coronel de Garcia Márquez, são exemplos de *sludges*: o tempo de espera; o preenchimento inesgotável de papéis antes de ser apto para auferir um benefício ou participar de um programa; instruções confusas para tarefas administrativas; exigências de certificação prévia antes de se exercer determinada profissão etc. Em outros termos, *sludges* dificultam a seleção de uma determina opção de escolha, ao criar fricções para que essa seleção seja efetivada.

Além disso, de acordo com Sunstein (2021a, p. 5-6), *sludges* seriam simplesmente como um tipo de custo de transação. Nem todo custo de transação, no entanto, seria um *sludge* em sua visão. Taxas exigidas para que uma determinada tarefa seja desempenhada, por exemplo, não são *sludges*. Embora reconheça que definir as condições necessárias e suficientes para *sludges* seja uma tarefa difícil, Sunstein (2020a, p. 5, 2021a, p. 6) não inclui incentivos ou desincentivos estritamente monetários dentro do conceito. Em sua visão, ainda que *sludges* possam representar um custo significativo aos agentes, seria importante ter uma distinção qualitativa entre um imposto ou uma multa, de um lado, e algo como uma exigência de preenchimento de formulário, de outro lado.

Do mesmo modo, Sunstein (2020a, p. 5, 2021a, p. 6) acrescenta que uma proibição não é considerada um *sludge*. Uma imposição, por outro lado, pode ou não ser um *sludge*, a depender do que está sendo exigido. Se um processo desnecessariamente complexo é estabelecido para que um paciente receba um tratamento psiquiátrico contra depressão ou ansiedade, um *sludge* parece estar presente. Em outras palavras, se o que está sendo exigido são fricções no processo decisório, a intervenção específica é um *sludge.* A intervenção é semelhante ao estabelecimento de restrições procedimentais mencionado por Thaler e Sunstein (2003b) como uma das estratégias compatíveis com o paternalismo libertário, como discutido

no Capítulo 3. Por outro lado, obrigações mais amplas, que determinam uma opção escolha específica, não seriam *sludges*.

Além disso, em sua primeira tentativa de expor de forma mais analítica os aspectos conceituais de um *sludge*, Sunstein (2020a, p. 5) ressaltou que *sludges* seriam ruins por definição – o que, inclusive, deriva da própria expressão utilizada para representar essa intervenção. *Sludges* seriam fricções *excessivas*. Para Sunstein (2020a, p. 7), seria útil reservar a expressão para imposições com valência negativa. Fricções que trouxessem efeitos benéficos, como evitar um determinado viés comportamental, deveriam ser vistas como *nudges* que promovem a deliberação, como ainda será demonstrado abaixo. Em outros termos, em um primeiro momento, a definição de *sludge* estava atrelada a um juízo de valor acerca de tal intervenção.

Em sua obra mais recente, no entanto, Sunstein (2021a, p. 7-8) afirma que *sludges* certamente não são sempre negativos. Eles podem ser excessivos, insuficientes ou ótimos. Para uma decisão bastante impactante na vida privada de um agente, como um divórcio, *sludges* podem ter efeitos positivos e garantir que a decisão não seja afetada, por exemplo, pela empatia quente-frio. *Sludges* também seriam importantes para a exclusão de um arquivo importante do computador, a renúncia a direitos ou qualquer outra ação cujo resultado pode ser imprudente.

Sunstein (2019b, p. 1865-1873, 2020a, p. 12-15, 2020b, p. 163-172, 2021a), inclusive, dedica-se de forma mais analítica aos motivos que podem justificar a adoção de *sludges*. Embora ultrapasse o objetivo desta dissertação discuti-los em detalhe, Sunstein menciona razões como: a) a garantia de elegibilidade e de integridade de um programa estatal; b) a reação a problemas de autocontrole; c) a proteção da privacidade; d) a proteção da segurança; e) o direcionamento de benefícios aos agentes que mais precisam ou que mais merecem – e que, portanto, estariam mais dispostos a superar *sludges*; e f) a reunião de dados importantes ou essenciais sobre as pessoas, para fins de políticas públicas.

Quais seriam, então, as distinções entre *nudges* e *sludges*? Sunstein (2020a, p. 7-8, 2021a, p. 9-10) pondera que, se *sludges* forem definidos simplesmente como fricções, sem a qualificação *excessivas*, eles poderiam ser um tipo de *nudge*, que poderia ser implementado para finalidades positivas ou negativas. Evidentemente, seriam uma categoria muito particular de *nudges*: o aumento não substancial de fricções. São *nudges* que buscam promover a deliberação, com o objetivo de permitir que as pessoas pensem com mais cuidado em suas decisões, como no clássico exemplo já mencionado dos períodos de reflexão.

Outros *nudges* buscam apenas reduzir fricções e, assim, simplificar decisões, *i.e.*, reduzir os *sludges*. O *Save More Tomorrow* é um clássico exemplo nesse sentido. Em outra

perspectiva, estratégias que buscam direcionar ou enganar as pessoas para fazê-las tomar decisões não intencionais e potencialmente danosas, como *dark patterns*[235], não são *sludges*: são *nudges* utilizados para finalidades negativas. Há, por fim, os casos que envolvem excessiva fricção, sem qualquer benefício ao agente, e que, portanto, seriam o domínio *exclusivo* dos *sludges*.

A tabela abaixo resume a sobreposição e a distinção entre *nudges* e *sludges*, segundo Suns:

Tabela 4 - *Nudges* e *sludges*: sobreposições e distinções segundo Sunstein

	Baixa fricção	Alta fricção
Finalidades positivas	*Nudge* útil que facilita uma decisão (*e.g.*, simplificação, mapas de aeroportos, inscrição automática em bons planos de previdência).	*Nudges* ou *sludges* que promovem deliberação (*e.g.*, períodos de reflexão, períodos de espera etc.) **Sobreposição entre os conceitos**.
Finalidades negativas	*Nudge* prejudicial que facilita uma decisão (*e.g.*, inscrição automática em algum programa custoso e inútil).	*Sludge stricto sensu* (*e.g.*, preenchimentos de formulários, longos períodos de espera para obter uma carteira de habilitação ou um visto).

Fonte: Sunstein (2021, p. 10)

Considerando que *sludges* não envolvem incentivos puramente econômicos, mas são, ao mesmo tempo, espécies de custos de transação por impor *fricções*, pode-se questionar o que exatamente seriam essas fricções. Sunstein, como mencionado, exclui do seu escopo medidas de comando e controle clássicas, como proibições ou obrigações que não se limitam ao acréscimo de fricções.

A partir de uma análise da definição de *nudge* e de quais são as suas condições necessárias, Stuart Mills (2020) conclui que essas fricções seriam a alteração dos custos: a) hedônicos, isto é, do prazer/conforto ligado a uma opção específica de escolha; b) sociais, ou seja, dos custos sociais e morais relacionados a uma opção de escolha; ou c) obscurecedores (*obscurants*), isto é, a alteração da carga psicológica/cognitiva ligada a uma determinada opção.

[235] Padrões obscuros (*dark patterns*) são interfaces elaboradas para confundir as pessoas e dificultar a expressão de suas próprias preferências na internet ou, ainda, manipular esses usuários a tomar decisões que eles não fariam sem essas interfaces. Essencialmente, padrões obscuros exploram vieses comportamentais e empurram as pessoas à compra de produtos e serviços que elas não desejavam ou a divulgar informações privadas que não seriam necessárias (MARKS, 2021, p. 535).

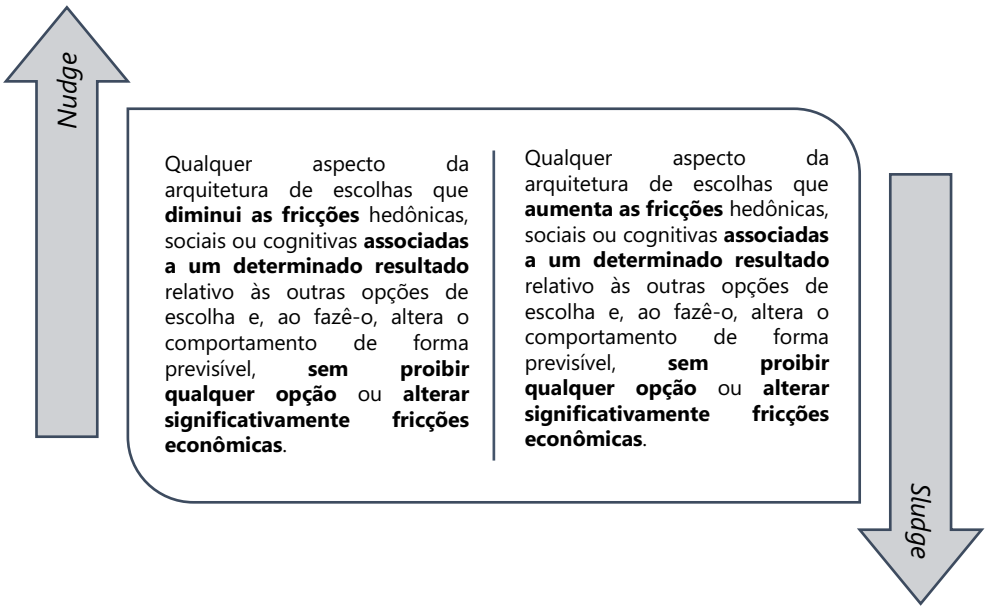
Mills pondera que alterações em todos esses custos (fricções) seriam admissíveis na teoria de um *nudge*, isto é, não afastariam as condições necessárias para sua definição.

Além disso, como Mills (2020, p. 8) pondera, *nudges* como inscrições automáticas em programas como o *Save More Tomorrow* são prescritos por Sunstein (2020a) como ferramentas para redução de *sludges*. Ao mesmo tempo, enquanto a burocracia excessiva, como incontáveis formulários e papéis a serem preenchidos, é frequentemente mencionada como exemplo de *sludge*, Sunstein (2014d, 2015c, 2019a, 2019c) já ressaltou, em inúmeras oportunidades, que o seu conceito amplo de *nudge* inclui medidas de simplificação de processos decisórios desnecessariamente complexos. Nesse caso, em linha com o que Mills (2020, p. 8) defende, se *nudges* podem reduzir *sludges* e *sludges*, por sua vez, são aumentos de fricções ou fricções excessivas, parece que *nudges* são formas de reduzir fricções. Por essa razão, Mills conclui que: *nudges* reduzem as fricções associadas a uma opção específica, enquanto *sludges* aumentam as fricções associadas a uma determinada opção de escolha. Esse elemento deve ser incorporado à definição estrita de *nudge* adotada nesta dissertação no Capítulo 3.

Mills (2020, p. 9) também apresenta um conceito que, além de útil para a compreensão da definição de *sludge*, é ainda essencial na discussão normativa que será empreendida no Capítulo 6: a simetria *nudge*/*sludge*. A ideia é simples: se *nudges* reduzem as fricções associadas a uma determinada opção, é possível concluir, em sentido contrário, que as fricções relativas das opções alternativas serão ampliadas. Em outros termos, um *nudge* acabaria por impor *sludges* com relação às opções que não são estimuladas. Uma regra *default* para incluir automaticamente trabalhadores em um plano de previdência privada, por exemplo, aumentaria as fricções relacionadas à solução de planos alternativos ou, mesmo, à opção de não poupar quantias significativas para o futuro. Analogamente, quando um *sludge* cria fricções para determinadas opções de escolha, pode haver a redução das fricções correspondentes às opções alternativas. Em outros termos, um *sludge* acabaria sendo um *nudge* com relação às opções para as quais fricções não são ampliadas.

Com base nessa simetria, Mills (2020, p. 10) sugere uma alteração modesta à clássica definição de *nudge*, apresentada na Subseção 3.1.3 acima, cumulada a uma alteração na definição de *sludge* para diferenciar as duas estratégias, tal como exposto abaixo:

Figura 12 - *Nudge versus Sludge*

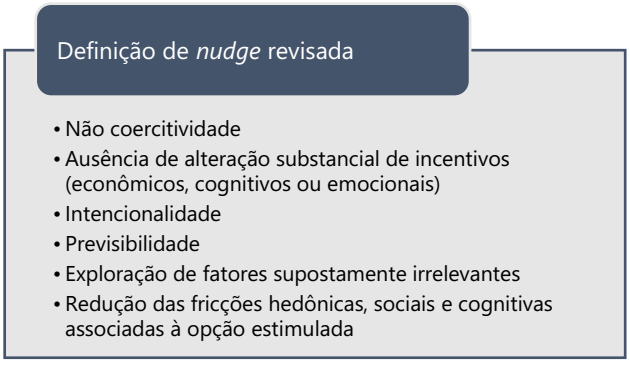


Fonte: Elaboração do Autor.

Mills (2020, p. 11) reconhece que, ao apresentar essa distinção, assim como o conceito de simetria *nudge/sludge*, um problema ontológico parece ser evidenciado. Se todo *nudge* produz *sludge* e todo *sludge* produz *nudge*, como definir a linguagem que deve ser adotada ao discutir uma intervenção? Em vez de prender-se em questões linguísticas, Mills (2020, p. 11–12) propõe focar na forma pela qual a arquitetura de escolhas é alterada para distinguir *nudges* e *sludges*. Nesse caso, um *nudge* é classificado dessa forma por reduzir as fricções associadas a um resultado específico com a intenção de estimular a escolha desse resultado; eventuais *sludges* que sejam impostos pela intervenção são apenas um subproduto do *nudge*. *Sludges*, em contrapartida, aumentam as fricções associadas a um resultado específico, com a intenção de encorajar um resultado alternativo; *nudges* para o resultado alternativo, assim, também surgem como um subproduto dos *sludges*.

Com base em tais distinções, é possível chegar a uma definição estrita de *nudge* revisada, para distinguir essas intervenções de *sludges*, acrescentando o elemento da redução de fricções hedônicas, sociais ou cognitivas à sua definição. A imagem abaixo apresenta essa definição revisada:

Figura 13 - Definição de *nudge* revisada com base na distinção *nudge*/*sludge*



Fonte: Elaboração do Autor.

Os aspectos normativos associados a essas distinções serão retomados no Capítulo 6. De qualquer forma - e, para finalizar esta seção -, convém ressaltar que *sludges* podem causar prejuízos muito diversos. Sunstein (2019b, 2020a, 2021a) enumera questões como a redução da efetividade de direitos fundamentais, no caso, por exemplo, da criação de fricções para o exercício de direitos como a liberdade de manifestação, ou, mesmo, efeitos regressivos em termos de redistribuição, ao impor custos excessivos para pessoas mais pobres.[236] Nesse caso, é essencial pensar na redução dos *sludges* e nas auditorias de *sludge* que Sunstein tanto recomenda.

4.2. *Nudges versus* paternalismo libertário: separando a técnica de intervenção das razões para se intervir[237]

Não é difícil perceber que pensar em técnicas de intervenção sem discutir razões para intervir na vida privada é uma tarefa difícil. Há casos, inclusive, em que as razões para intervenção são incorporadas como condições necessárias à definição de uma determinada técnica. Apenas para exemplificar, Thaler (2018a) afirma que *sludges* seriam essencialmente *nudges* para o mal; Sunstein (2020a, p. 5), do mesmo modo, chega a afirmar que seria recomendável limitar a expressão *sludges* a fricções criadas para finalidades negativas ou que são prejudiciais aos agentes. Analogamente, Sunstein (2014a, 2015b, 2016a) afirma, em diversas oportunidades, que *nudges* não são problemáticos porque tentam influenciar as pessoas de maneira a melhorar o seu bem-estar segundo os seus próprios critérios (*as judged by*

[236] Sobre esses efeitos negativos e sobre os prejuízos cognitivos da pobreza, Cf.: (HAUSHOFER; FEHR, 2014; MULLAINATHAN; SHAFIR, 2016; SHAH; MULLAINATHAN; SHAFIR, 2012; SILVA *et al.*, 2018)

[237] Agradeço à Profa. Dra. Rute Saraiva por chamar a minha atenção aos pontos discutidos nesta seção.

themselves). A premissa dessa afirmação é, nitidamente, a incorporação de uma razão para intervenção - a promoção dos melhores interesses do sujeito - na definição de uma técnica para intervir - o próprio *nudge*.

Apesar disso, fundamentos para intervenção não se confundem com técnicas de intervir. No paternalismo comportamental coercitivo de Conly (2013), por exemplo, como já exposto no Capítulo 2, embora os fundamentos para intervir pareçam ser idênticos aos de Sunstein, Thaler e de outros paternalistas comportamentais, as técnicas de intervenção recomendadas são completamente distintas. Do mesmo modo, uma técnica como um *nudge* pode ser utilizada para que um arquiteto de escolhas possa se beneficiar, em detrimento do agente - tema que será mais bem abordado na Seção 4.3[238].

A partir de tal distinção, não parece uma boa ideia apresentar definições de técnicas de intervenção em termos normativos, atreladas a um determinado juízo de valor ou a razões para intervenção. Por esse motivo, optou-se, nesta dissertação, por separar a análise dos fundamentos à regulação e das técnicas de intervenção.

Como a discussão apresentada nos parágrafos anteriores evidencia, em muitos dos paternalismos comportamentais, especialmente na obra de Thaler e Sunstein, há uma certa confusão entre o que justifica uma intervenção e como ela é implementada. Deixar essa separação clara facilita a avaliação normativa tanto dos *nudges*, quanto, de forma mais específica, de *nudges* orientados pelo paternalismo libertário.

Nem sempre, é verdade, será fácil identificar quais são as razões específicas que orientam uma determinada intervenção.

Neste ponto, convém retomar a distinção apresentada no Capítulo 2 entre o paternalismo misto (*mixed paternalism*), em que motivos para intervir se somam, de modo que haveria múltiplos fundamentos de natureza diversa - não só paternalista - para uma determinada intervenção, e o paternalismo não misto (*unmixed paternalismo*), no qual as razões para intervir são exclusivamente paternalistas (FEINBERG, 1986, p. 8). A maior parte das regras paternalistas, como aponta Feinberg (1986, p. 16), são mistas. Quando há razões plausíveis em diferentes frentes para uma determinada intervenção, elas podem se reforçar, e críticas normativas a um determinado tipo de fundamento - *i.e.*, motivos paternalistas - acabam por não abalar a legitimidade da intervenção específica.

Feinberg (1986, p. 16-18) destaca, ainda, a dificuldade que existe para se identificar os reais fundamentos de uma determinada intervenção. Se, por exemplo, o Congresso aprova uma

[238] Cf., ainda, (AKERLOF; SHILLER, 2015).

lei que, supostamente, teria sido motivada por uma relevante razão de ordem pública e, anos depois, descobre-se que o fundamento era outro, que parece ilegítimo, quais são os verdadeiros fundamentos das regras aprovadas?

Para identificar esses verdadeiros motivos, como Feinberg (1986, p. 17) acrescenta, é necessário distinguir "razões conscientes", “motivações profundas”, “justificativas implícitas” e “verdadeiras justificativas”. Pode ser, por exemplo, que o legislador vote em um projeto de lei e cite um determinado fator como o fundamento do seu voto - isto é, a sua razão consciente - , mas, na verdade, as suas motivações profundas estão ligadas a outros fatores que ultrapassam a sua consciência ou que seriam moralmente repugnantes.

Se pensarmos, por exemplo, na regulação de *fake news* no Brasil, que, atualmente, está em discussão na Câmara dos Deputados[239], as motivações profundas para se posicionar de maneira contrária ou favorável ao projeto parecem estar ligadas a interesses eleitorais. As razões conscientes, por seu turno, baseiam-se na defesa da liberdade de expressão, de um lado, ou na necessidade de combater os prejuízos trazidos por notícias falsas ou reduzir o controle informacional de grandes plataformas.

A definição das verdadeiras justificativas, por outro lado, não é clara nas distinções apresentadas por Feinberg. De qualquer forma, identificá-las parece uma tarefa essencialmente contingente e pouco promissora em termos de resultados práticos. Os verdadeiros motivos que embasam uma determinada norma, além de refletirem problemas a serem enfrentados ao tempo em que essa norma foi proposta, também são influenciados pelos valores do momento em que a análise é feita. Ao mesmo tempo, essa análise sempre será influenciada pelos efeitos do viés retrospectivo. Se uma determinada regra se mostrou inefetiva, por exemplo, pode ser que a análise retrospectiva do desenho normativo tenha outra visão dos fundamentos que o motivaram. Não parece possível, portanto, que se identifique um único e imutável fundamento de uma determinada regra.

Independentemente disso, ao analisar uma determinada intervenção, é possível construir justificativas implícitas para sua proposição, ainda que essas justificativas não coincidam com as verdadeiras razões ou com as motivações profundas que motivaram o apoio a ela. Apesar disso, as justificativas implícitas apresentam uma reconstrução racional plausivelmente coerente (FEINBERG, 1986, p. 17) e podem servir como um norte na tentativa de avaliar os fundamentos de uma determinada regra.

---

[239] Projeto de Lei n.º 2630, de 2020. Vide: https://www.camara.leg.br/propostas-legislativas/2256735. Acesso em: 28 maio 2022.

Feitas essas considerações, Sunstein (2014a, p. 80–81) afirma que a sua definição de paternalismo não inclui esforços do governo para prevenir danos a terceiros – em linha com o princípio do dano – nem objetivos sociais comuns, como a proteção da privacidade, a proteção de espécies ameaçadas ou contra a discriminação. Por outro lado, incluem-se na sua definição medidas voltadas para superar ou influenciar os julgamentos das pessoas a respeito da medida em que devem beber álcool ou dirigir enquanto usam os seus celulares. Sunstein reconhece, contudo, que alguns desses últimos casos podem gerar danos a terceiros, isto é, pode ser que os fundamentos para intervir nestes casos sejam mistos.

Apesar disso, uma definição de *nudge* atrelada ao paternalismo limita consideravelmente o universo de *nudges* legítimos, já que muitas intervenções comportamentais podem ter o objetivo de promover o bem-estar social (welfarismo), sem qualquer preocupação com o bem-estar individual das pessoas que são empurradas, ou mesmo políticas paternalistas tradicionais, como a obrigação de utilizar cintos de segurança, em que se promove o bem-estar em conformidade com as preferências do próprio elaborador da política (TOR, 2016, p. 5-6).

Como afirma Hansen (2016), a definição de *nudge* não faz exigências quanto às motivações à elaboração da política. Essa é uma definição técnica de *nudge*, que não está atrelada às razões para intervenção, e ela parece ser mais adequada. Afinal, enquanto um *nudge* é uma das técnicas de intervenção disponíveis no cardápio do regulador, o paternalismo libertário é um movimento que busca justificar políticas públicas específicas que tentam proteger as pessoas de si mesmas, sem coerção ou mudança significativa de incentivos. Os conceitos não se confundem, embora existam muitos *nudges* justificáveis de acordo com os critérios do paternalismo libertário, como será demonstrado no Capítulo 6.

Mais do que isso, na própria obra de Thaler e Sunstein (2009), há exemplos de *nudges* com fundamento welfarista. Aqui, o bem-estar do indivíduo fica em segundo plano; o fundamento da intervenção é o aumento do bem-estar coletivo. *Nudges* que buscam estimular doações de órgãos são um exemplo claro[240]. Nesse caso, pode ser que sujeitos que se opõem à doação de órgãos, ou que não a consideram como um elemento que aumente o seu próprio bem-estar sejam ainda assim empurrados (TOR, 2016, p. 7). Em outras palavras, essas pessoas podem ser influenciadas a serem doadoras, ainda que essa opção não melhore o seu bem-estar segundo seus próprios critérios. O *policy-maker*, nesse caso, não se importará com esse descompasso: a finalidade da intervenção é o aumento do bem-estar coletivo.

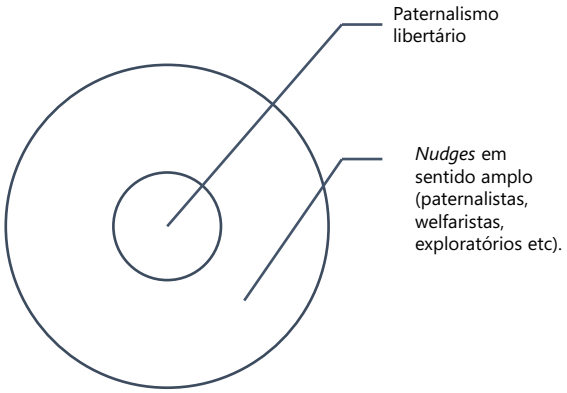
[240] Sobre essa distinção, Cf.: (SARAIVA, 2020, p. 21-24).

Em outra perspectiva, *nudges* também podem ser exploratórios, quando são usados com objetivos negativos ao agente e à própria sociedade. Esse aspecto será retomado na próxima seção, que discutirá os *nudges* privados. De qualquer forma, apenas para exemplificar, consumidores podem ser empurrados para uma regra *default* que lhes atribui um plano de saúde caro ou um seguro para os seus telefones celulares que não é de seu interesse. Do mesmo modo, o governante incumbente pode tentar empurrar as pessoas para não votarem, se constatar que as suas chances de eleição são maiores em caso de alta abstenção. *Sludges*, aqui, também poderiam ser amplamente utilizados. Essas intervenções teriam propósitos negativos e, mesmo que sem coerção, sua legitimidade seria certamente questionável.

Dessa forma, os *nudges* não se confundem com o paternalismo libertário. Na verdade, os fundamentos para a utilização de um *nudge* são muito diversos, assim como para a implementação de qualquer técnica de regulação do comportamento privado.

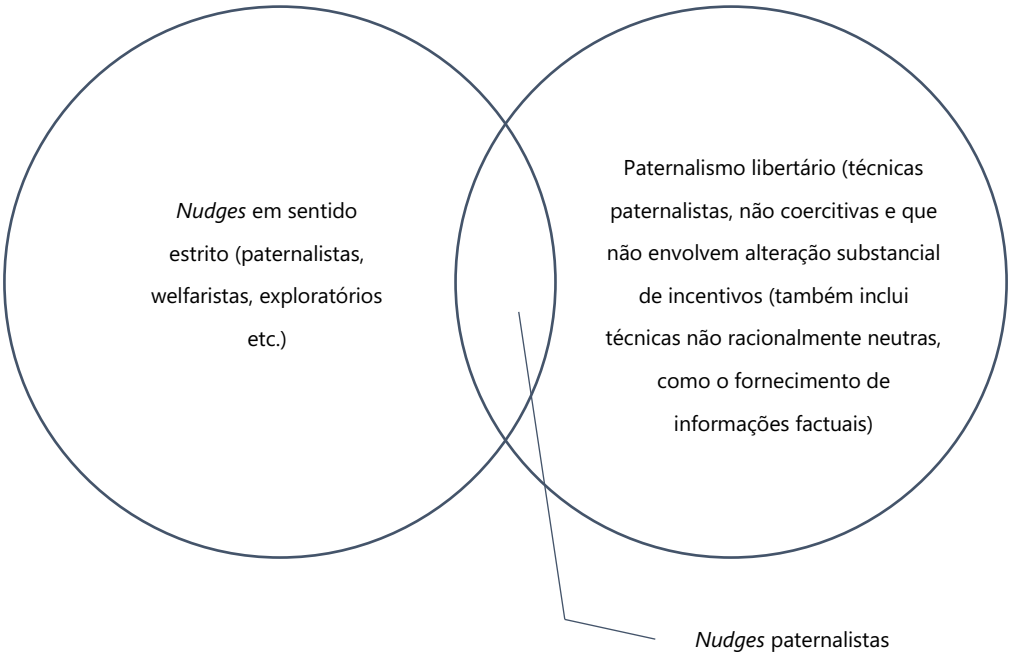
Em uma definição ampla de *nudge*, como a adotada por Sunstein, as medidas do paternalismo libertário até poderiam ser, em geral, definidas como *nudges*. Na definição restrita, por outro lado, em que *nudges* são medidas racionalmente neutras, o paternalismo libertário engloba medidas que vão além dos *nudges* em sentido estrito, como o fornecimento de informações factuais. As imagens abaixo esclarecem essa distinção:

Figura 14 - Relação entre o paternalismo libertário e *nudges* em acepção ampla



Fonte: Elaboração do Autor.

Figura 15 - Relação entre o paternalismo libertário e *nudges* em sentido estrito



Fonte: Elaboração do Autor.

Por fim, em justiça à obra de Thaler e Sunstein, convém mencionar que essa distinção entre técnica e fundamento de intervenção não lhes passa despercebida. De fato, na proposição original do paternalismo libertário, Thaler e Sunstein (2003b, p. 26) argumentaram que as mesmas estratégias que o implementam poderiam ser usadas em muitos domínios. Se, para além do paternalismo, essas técnicas fossem implementadas para proteção de terceiros, haveria uma espécie de benevolência libertária (*libertarian benevolence*). É o caso, por exemplo, do empregador que estimula doações para caridade entre os seus empregados. São políticas nitidamente welfaristas. Naquela oportunidade, inclusive, Thaler e Sunstein (2003b, p. 33) usam o exemplo de regras *default* para ampliar doações de órgãos como uma medida de benevolência libertária.

Em outra oportunidade, Sunstein (2015b, p. 426-427) até chega a diferenciar *nudges* paternalistas de *nudges* que enfrentam falhas de mercado. Evidentemente, se o fundamento de intervenção é enfrentar uma falha de mercado, não há paternalismo nessa hipótese. Um *nudge* como esse poderia ser defendido com fundamentos welfaristas e, até mesmo, libertários. Mas, não há paternalismo. Sunstein (2020a, p. 7) admite, ainda, embora, muitas vezes, isso não ganhe tanta relevância em sua obra, que *nudges* podem ser utilizados para propósitos positivos e

negativos. Em outras palavras, mais uma vez, evidencia-se que *nudges* não se confundem com o paternalismo libertário.

Apesar disso - e esse é o mote desta seção -, Sunstein continua defendendo a utilização de *nudges* com base no critério como julgado por si próprios (*as judged by themselves*) – o que, na verdade, é um critério de defesa do paternalismo libertário. Essa defesa só é válida se a crítica for direcionada, exclusivamente, a *nudges* paternalistas libertários. Um *nudge* para doação de órgãos não seria afetado por uma crítica como essa. De qualquer forma, preocupações ligadas à manipulação e ao caráter sub-reptício dos *nudges* afetariam essa intervenção.

Essa confusão entre técnica de intervenção e os seus fundamentos dificulta consideravelmente a sua avaliação normativa - tema que será abordado no Capítulo 6. Afinal, críticas que são direcionadas especificamente ao paternalismo libertário podem acabar impactando a legitimidade de aplicação de *nudges*. Em outros termos, a crítica não teria qualquer objeção a um *nudge* que enfrenta uma falha de mercado, por exemplo; como os conceitos se confundem, essa crítica pode acabar rejeitando um *nudge* que poderia ter sua legitimidade incontroversa. Do mesmo modo, críticas aos *nudges*, ligadas a um temor de que eles venham a ser utilizados para reforçar o poder estatal e ao fato de eles explorarem limitações cognitivas, acabam impactando políticas legítimas que têm fundamento paternalista libertário - se adotada a definição estrita de *nudge*, tal como exposto acima. Nesse caso, talvez não houvesse por parte de tais críticos uma objeção ao fornecimento de informações factuais; mas, como os dois conceitos são apresentados de forma quase indissociável, a crítica pode acabar direcionada a uma intervenção como essa.

4.3. *Nudges* enquanto técnica de regulação estatal e o rearranjo da arquitetura de escolhas por agentes privados[241]

Para justificar a adoção de *nudges* e do paternalismo libertário, Thaler e Sunstein (2009, 2021) frequentemente argumentam que a arquitetura de escolhas é inevitável; assim, ainda que se decida não intervir, uma disposição aleatória dos elementos que a compõem também influenciaria o comportamento das pessoas. Para reforçar esse argumento, Sunstein (2015b, 2019a) chega a mencionar que a natureza e o clima empurram as pessoas a todo momento. Nesse contexto, em meio a influências aleatórias, seria melhor que o arquiteto de escolhas tentasse influenciar as pessoas para direções que melhorassem o seu bem-estar.

[241] Agradeço à Professora Dra. Rute Saraiva e ao Professor Dr. Leonardo Martins Wykrota pelas provocações que deram origem a essa seção.

Uma influência acidental do contexto, no entanto, não pode ser comparada a uma influência intencional e previsível. A intencionalidade é um elemento essencial para a justificativa normativa de uma intervenção (HANSEN, 2016). Não por outra razão, como demonstrado discutido Capítulo 3, enquanto intervenções que alteram o comportamento de forma previsível, *nudges* presumem a intencionalidade do arquiteto de escolhas responsável por elaborar a política. Influências do clima e da natureza são, de fato, inevitáveis; uma medida regulatória do Estado, por outro lado, além de evitável, pode ser elaborada de inúmeras formas.

Além disso, Thaler e Sunstein ressaltam, a todo momento, que os arquitetos de escolha podem ser tanto entidades governamentais, quanto agentes privados.[242] Ao justificar a utilização de *nudges* enquanto técnicas regulatórias, Sunstein (2019a, p. 116), inclusive, fundamenta sua abordagem na ampla e inevitável utilização de *nudges* pelo mercado; em outros termos, se os agentes do mercado empurram o consumidor para opções que lhes podem gerar mais lucro, o governo também deveria intervir para promover o seu bem-estar.

Mais uma vez por razões metodológicas, unir essas duas formas diversas de utilizar *nudges* não parece ser a opção adequada. Os desafios normativos que surgem da utilização de *nudges* privados - em contratos ou na própria elaboração da arquitetura de escolhas do ambiente em que um produto ou um serviço é utilizado – são completamente diversos daqueles que surgem quando o Estado resolve utilizar *nudges*. O objetivo desta dissertação é discutir a legitimidade de *nudges* enquanto técnica de regulação estatal; *nudges* privados não fazem parte do seu escopo. De qualquer forma, como agentes privados são frequentemente arquitetos de escolha em *nudges* elaborados pelo governo, algumas distinções conceituais são essenciais à discussão normativa que será feita no Capítulo 6.

Inicialmente, em termos jurídicos, para além da intencionalidade, a arquitetura de escolhas pressupõe uma relação jurídica entre, no mínimo, dois agentes: a) o arquiteto de escolhas, que será responsável por elaborar o contexto em que escolhas serão tomadas; e b) um decisor, que será influenciado pela arquitetura de escolhas. Alguns exemplos deixam essa afirmação mais clara: uma empresa organiza o ambiente em que os seus inúmeros empregados tomarão decisões diversas; um condomínio edilício é responsável por organizar o contexto em que os condôminos farão escolhas; o Estado, enquanto prestador de serviços públicos, organiza o contexto em que os usuários desses serviços tomarão decisões; um fornecedor de produtos ou serviços elabora a arquitetura de escolhas dos consumidores; e assim sucessivamente. São

[242] Thaler e Sunstein (2021, p. 8) afirmam, por exemplo, que muitas políticas que eles recomendam podem e têm sido implementadas no setor privado (com ou sem um *nudge* elaborado pelo governo). Empregadores, por exemplo, podem criar *nudges* para redução da poluição do ar.

relações jurídicas específicas que criam a oportunidade para que um agente interfira no contexto em que um terceiro tomará decisões.

No âmbito dessas relações jurídicas, a arquitetura de escolhas intermedeia a relação entre arquiteto de escolhas e decisor. A partir dela, o arquiteto de escolhas define elementos como: a) as opções que poderão ser escolhidas pelo decisor; b) quais destas opções ficarão mais disponíveis visualmente (salientes) para o decisor; c) que tipo de escolhas devem ser feitas de maneira ativa e quais escolhas serão presumidas; d) o custo econômico associado a cada uma das opções; e) o nível de fricções – sociais, hedônicas ou psicológicas/cognitivas – associado a cada opção; f) o procedimento para que escolhas sejam feitas; g) as informações textuais e visuais que acompanharão as opções disponíveis ao decisor; etc. Algumas dessas estratégias são, nitidamente, *nudges*.

A natureza dessas relações jurídicas definirá como a arquitetura de escolhas poderá ser (re)organizada e quais os limites de tal (re)organização.

Nesse sentido, se a relação entre arquiteto de escolhas e decisor for de natureza puramente privada, haverá um espaço maior para reorganização, que, em geral, será implementada contratualmente ou como parte do próprio *design* do produto ou do serviço. Em regra, essas decisões serão livremente tomadas pelo arquiteto de escolhas e nudges implementados neste contexto são puramente privados. Uma ferramenta de busca, por exemplo, terá que definir um critério para classificação dos resultados - *i.e.*, quais serão os primeiros resultados a aparecerem em resposta a uma determinada pesquisa -; os custos - em termos monetários ou de dados - associados à utilização dos serviços; a forma como é exercido o consentimento com relação ao tratamento de dados; os espaços que serão destinados a produtos integrados à plataforma etc.

Há toda uma literatura direcionada à aplicação de *nudges* privados. Explorá-la ultrapassa o escopo desta dissertação[243]. De qualquer forma, algumas considerações a respeito de *nudges* privados auxiliam a distinguir os dilemas que surgem de sua implementação daqueles que surgem diante de *nudges* enquanto medidas regulatórias.

Como mencionado acima, *nudges* privados são adotados de maneira essencialmente livre por agentes privados. Tanto quando adotados em contrato quanto como parte do *design* de um produto ou serviço, os *nudges* privados, em geral, contam com o consentimento do decisor. Em tese, se ele não estivesse de acordo com o *nudge* específico, poderia optar por encerrar ou sequer iniciar sua relação jurídica com o arquiteto de escolhas. Um consumidor poderia

[243] Para um apanhado de alguns textos relevantes dessa literatura, Cf.: (ABDUKADIROV, 2016).

consumir de concorrentes que não utilizassem *nudges* que lhe desagradassem; um empregado poderia trabalhar em outra companhia; o morador de um condomínio edilício poderia encontrar outro lugar para morar, e assim sucessivamente.

De qualquer forma, esses *nudges* nem sempre funcionam de forma transparente e, por explorarem limitações comportamentais, eles podem passar despercebidos. Como já mencionado, Akerlof e Shiller (2015) demonstram que, como as pessoas têm falhas – nem sempre estão completamente informadas e podem ter dificuldades para saber o que realmente querem –, elas podem ser enganadas (*tricked*); o livre mercado, então, em equilíbrio, além de fornecer o que as pessoas querem e precisam, será um campo aberto à pescaria de tolos (*phishing for phools*), isto é, à exploração das falhas comportamentais dessas pessoas. Nesse caso, "se temos uma fraqueza – se temos uma maneira pela qual somos pescáveis (phishable) – o pescador (phisherman) estará lá esperando" (AKERLOF; SHILLER, 2015, p. 171)[244]. Afinal, se não estiver, os seus concorrentes estariam lá e conseguiriam ter vantagens competitivas.

Embora Akerlof e Shiller não discutam apenas *nudges* - a pesca de tolos pode ser efetivada a partir de toda sorte de técnicas comportamentais, incluindo *nudges* e *sludges* -, toda a sua exposição evidencia que *nudges* privados nem sempre serão benéficos. Mais do que isso, é esperado que esses *nudges* sejam elaborados para beneficiar os arquitetos de escolhas em detrimento dos decisores.

Nesse sentido, Jodi N. Beggs (2016), ao discutir *nudges* do setor privado, faz uma distinção entre *nudges* de Pareto e *nudges rent-seeking*. Ao implementar *nudges* de Pareto, os fornecedores direcionam o comportamento dos consumidores para melhor alinhar as suas decisões aos seus próprios interesses de longo prazo. Beggs menciona exemplos como a inscrição automática de empregados em programas de previdência 401(k). Por outro lado, *nudges rent-seeking* se posicionam de forma contrária aos interesses dos consumidores: são casos em que empresas tentam extrair renda dos consumidores a partir de *nudges*. Em outros termos, são casos de *phishing for phools*. Academias de ginástica, por exemplo, podem criar modelos de negócio que, com combinações de *nudges* e *sludges*, estimulam, ao mesmo tempo, a adesão de consumidores e a não utilização dos equipamentos das respectivas academias (BEGGS, 2016, p. 139-140)[245].

---

[244] Tradução livre. Texto original: "[...] *if we have a weakness – if we have a way in which we are phishable - the phisherman will be there in waiting*".

[245] Mason Marks (2021) explora o que seriam *nudges rent seeking* no âmbito de plataformas digitais. Em sua visão, a partir da arquitetura de escolhas, plataformas digitais buscam, manipular e regular o comportamento. Padrões obscuros (*dark patterns*) e outras técnicas comportamentais seriam utilizados para manter a atenção das pessoas, empurrá-las para agir de forma contrária aos seus próprios interesses, conformar-se às normas estabelecidas pelos

Nesse caso, se a reorganização privada da arquitetura de escolhas pode trazer prejuízos de bem-estar, é natural que se questione se há necessidade de alguma intervenção estatal. Ainda que este não seja o objeto desta dissertação, convém pontuar que, de fato, intervenções podem ser necessárias para proibir ou alterar certas formas de *nudges* privados. Essas intervenções são comuns em diversos setores do direito, como o direito do consumidor, o direito da concorrência, o direito regulatório etc.

No âmbito do direito privado, por exemplo, Ludmilla Junqueira Duarte Oliveira e Brunello Souza Stancioli (2021, p. 13) argumentam que "[r]econhecer que o agente capaz não deve ser tido por autossuficiente e racional deve levar o Direito a criar mecanismos para que as decisões com repercussões jurídicas sejam tomadas da melhor forma possível", levando-se em consideração a condição subjetiva do sujeito, assim como os elementos do contexto. O Código de Defesa do Consumidor, nesse sentido, incorpora previsões que "sugerem que, no contexto do mercado de consumo, a tomada de decisões pelo consumidor mereceu maior atenção do legislador" (OLIVEIRA; STANCIOLI, 2021, p. 13). Nesse contexto, há regras que reconhecem a vulnerabilidade do consumidor (agente decisor) e incentivam o fornecimento de informações, além de vedações a publicidades e práticas abusivas enganosas, como nos casos em que essas publicidades possam "induzir em erro o consumidor a respeito dos produtos e serviços comercializados ou de estimulá-los a se comportar de forma prejudicial ou perigosa à sua saúde ou segurança" (OLIVEIRA; STANCIOLI, 2021, p. 13).

A arquitetura de escolhas pode, ainda, ser objeto de intervenção no âmbito do direito concorrencial[246]. Em um dos processos mais famosos do direito antitruste europeu, por exemplo, a Direção-Geral da Concorrência da Comissão Europeia condenou o Google pelo fato de a sua ferramenta de busca genérica supostamente favorecer o Google Shopping, em detrimento de outros serviços de busca especializada. O suposto favorecimento seria configurado a partir da estruturação da arquitetura de escolhas da ferramenta de busca do Google, que estaria oferecendo espaços privilegiados ao Google Shopping, de modo que outros serviços de busca de preços estariam menos à mostra aos consumidores. A mesma prática foi investigada no Brasil, mas o Conselho Administrativo de Defesa Econômica acabou arquivando o processo, sem identificar qualquer infração à ordem econômica[247]. De todo modo, constata-

desenvolvedores das plataformas e impedir que usuários utilizem plataformas concorrentes (MARKS, 2021, p. 531-532).

[246] A *Competition & Markets Authority*, autoridade de defesa da concorrência do Reino Unido, por exemplo, publicou recentemente um estudo a respeito dos prejuízos que a arquitetura de escolhas *online* pode causar aos consumidores e à concorrência. Cf., nesse sentido: (COMPETITION & MARKET AUTHORITY, 2022).

[247] Para uma avaliação do caso Google Shopping com ênfase aos resultados na Europa e nos Estados Unidos, Cf.: (CARVALHO, 2022).

se que *nudges* e a arquitetura de escolhas podem ser submetidos ao escrutínio da defesa da concorrência.

O Banco Central do Brasil (BCB), por sua vez, disciplina a arquitetura de escolhas da relação entre instituições financeiras e consumidores de diversas formas. A Resolução n.º 3.919 do BCB, por exemplo, prevê, em seu art. 13, os elementos que devem constar em faturas de cartão de crédito. De maneira semelhante, nos Estados Unidos, o *Credit Card Accountability Responsability of Disclosure Act of 2009* criou parâmetros de transparência que devem constar em faturas de cartão de crédito, como projeções de amortização dos débitos já existentes e esclarecimentos quanto a encargos aplicáveis em cada hipótese de pagamento da fatura[248].

Todos esses exemplos representam limitações no escopo de legitimidade da arquitetura de escolhas privada. Diante dessas regras, alterações da arquitetura de escolhas que ultrapassem esses limites podem acabar sendo consideradas ilícitas. Nesse caso, quanto mais intervenções forem feitas, mais a arquitetura de escolhas deixa de ser puramente privada e passa a ser minimamente disciplinada por regras de ordem pública. Pode ser, inclusive, que, especialmente diante de regras prescritivas, a alteração da arquitetura de escolhas seja determinada pelo Estado; um *nudge*, nessa perspectiva, seria uma medida regulatória e a sua análise está dentro do escopo desta dissertação.

Nesse contexto, uma segunda distinção é relevante e tem implicações à análise normativa que será empreendida nos próximos Capítulos.

No Capítulo 2, uma das classificações de paternalismo que foi apresentada diferenciava o paternalismo puro - medida imposta apenas ao próprio sujeito que por ela será protegido – do paternalismo impuro, em que a classe de pessoas afetadas pela medida é maior do que a classe de pessoas que está sendo protegida (DWORKIN, 2020). Do mesmo modo, *nudges* podem afetar apenas os decisores (*nudges* puros), isto é, a classe de pessoas cujo comportamento é influenciado, ou podem ser imposições estatais para arquitetos de escolhas privados em sua relação com os decisores (*nudges* impuros). Como Sunstein (2015b, p. 424) explica, embora os *nudges* preservem a liberdade de escolhas para a população que será afetada pela política, pode haver, paralelamente, obrigações para outros agentes da população, isto é, os arquitetos de escolha.

Por serem cumulados com medidas coercitivas, *nudges* impuros geram maior controvérsia e devem ser tomados com maior cuidado. De maneira geral, esses *nudges* podem ser implementados: a) por determinações legais, como a presunção de adesão ao Cadastro

[248] Para uma discussão de tais medidas regulatórias enquanto técnicas de simplificação (*nudges*), Cf.: (SILVA *et al.*, 2019, p. 89-92).

Positivo, prevista na Lei Complementar n.º 166, de 8 de abril de 2019; bi) no âmbito do poder normativo das agências reguladoras[249], como as já mencionadas regras do BCB; c) em editais de licitações promovidas pelo Poder Público, para delegações ou permissões de serviços públicos; e e) como sanções ao descumprimento de alguma obrigação legal, como no exemplo da condenação do Google por abuso de posição dominante na União Europeia, o que pode ocorrer tanto no âmbito do contencioso administrativo quanto em decisões judiciais.

Nudges puros, por outro lado, são adotados em situações em que o arquiteto de escolhas é a própria administração pública. Em outros termos, há uma relação direta entre Estado e decisor, que não é intermediada por qualquer entidade privada; é o caso, por exemplo, da prestação direta de serviços públicos pelo Poder Público[250] ou da relação entre a administração e os seus servidores e empregados públicos. De forma geral, esses *nudges* podem ser implementados por: a) determinações legais, como a presunção de adesão dos servidores da União aos planos de previdência complementar do FUNPRESP, nos termos da Lei n.º 12.618, de 30 de abril de 2012; e b) por atos administrativos típicos, como no famoso exemplo da mosca dos mictórios do aeroporto de Amsterdam[251].

Em ambos os casos, *nudges* puros e impuros, ainda que no âmbito de sua própria discricionariedade, o Estado deve motivar as suas alterações da arquitetura de escolhas, em linha, inclusive, com os princípios que guiam a atividade da Administração Pública. A depender das maneiras como essas alterações são implementadas, serão distintas as técnicas de controle dos seus limites, isto é, do escopo de legitimidade normativa de *nudges* enquanto medidas regulatórias. Esse aspecto será retomado no Capítulo 6.

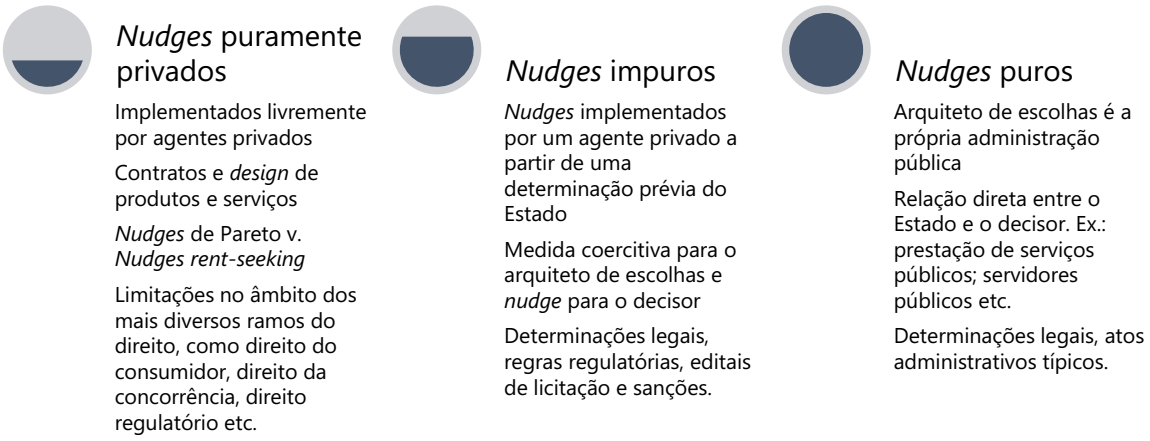
Os principais traços de *nudges* puramente privados, *nudges* impuros e *nudges* puros, encontram-se descritos na imagem abaixo:

[249] De acordo com o art. 174 da Constituição da República, o Estado, enquanto agente normativo e regulador da atividade econômica, exercerá as funções de fiscalização, incentivo e planejamento (BRASIL, 1988). Essa função é desempenhada pelas agências reguladoras.

[250] Como previsto no art. 175 da Constituição da República, "[i]ncumbe ao Poder Público, na forma da lei, diretamente ou sob regime de concessão ou permissão, sempre através de licitação, a prestação de serviços públicos. (BRASIL, 1988)"

[251] Vide: https://worksthatwork.com/1/urinal-fly. Acesso em: 17 mai. 2022.

Figura 16 - *Nudges* puramente privados, *nudges* impuros e *nudges* puros



Fonte: Elaboração do Autor.

4.4. Principais conclusões do Capítulo 4

Neste Capítulo, algumas questões adicionais do estágio atual da literatura a respeito dos *nudges* foram apresentadas. Inicialmente, foram indicados os traços definidores de estratégias como técnicas de desenviesamento, *boosts*, *thinks*, *nudges plus* e *sludges*, assim como as diferenças entre essas técnicas e os *nudges*. Em seguida, para deixar a distinção entre técnica de intervenção e fundamentos para intervir mais clara, *nudges* foram diferenciados do paternalismo libertário. Por fim, a distinção entre *nudges* adotados por agentes privados e *nudges* implementados como técnica regulatória foi esclarecida. As principais conclusões deste Capítulo foram as seguintes:

1. Embora algumas intervenções comportamentais tenham sobreposições com *nudges*, as diferenças entre todas essas técnicas são claras e podem ser sumarizadas nos seguintes pontos:
    a) Técnicas de desenviesamento não são uma espécie de técnica de regulação, mas um resultado que pode ser buscado por inúmeras técnicas. *Nudges* podem ser medidas de desenviesamento quando tiverem por objetivo reduzir a influência de um viés comportamental. Em todo caso, na definição estrita adotada nesta dissertação, *nudges* sempre envolverão alguma medida de enviesamento. Há, ainda, *nudges* que, em vez de desenviesar, alteram o resultado de escolhas motivadas por vieses comportamentais, explorando-os ativamente;

b) em comparação aos *boosts*, *nudges* tentam influenciar o comportamento, a partir de alterações da arquitetura de escolhas, com o objetivo de explorar vieses comportamentais. Um *nudge* pode até ser educativo e gerar um aumento de competências, mas esse efeito é derivado de uma alteração prévia do comportamento, promovida por uma alteração da arquitetura de escolhas que reduz as fricções associadas a uma determinada opção. Um *boost*, por outro lado, tenta estimular o desenvolvimento de competências, a partir da reelaboração de aspectos do ambiente dos indivíduos ou de técnicas para ensiná-los a readequar esse ambiente. Ainda que também possa haver alteração na arquitetura de escolhas, a mudança de comportamento, nesse caso, é reflexo de um prévio aprimoramento das capacidades decisórias;

c) *Thinks* são alterações procedimentais que exigem a deliberação pública, como pré-requisito a uma determinada decisão. *Thinks* costumam se dedicar a decisões mais importantes e pontuais, enquanto *nudges* se voltam a decisões cotidianas. As dificuldades para implementar *thinks* levaram ao desenvolvimento dos *nudges plus*, que nada mais são do que a incorporação de um elemento deliberativo a um *nudge* específico que altera a arquitetura de escolhas para reduzir as fricções associadas a uma determinada opção. A intenção, nitidamente, é reduzir as preocupações ligadas aos *nudges* e promover a agência dos decisores;

d) *Nudges* reduzem as fricções – hedônicas, sociais ou cognitivas – associadas a uma determinada opção de escolha. É o caso de regras *default* que presumem o consentimento e da alteração da ordem de opções para tornar uma opção mais saliente. *Sludges*, em contrapartida, ampliam essas fricções. São exemplos a imposição de períodos de reflexão (*cooling-off periods*) e a criação de etapas a serem cumpridas para que uma determinada decisão possa ser tomada. Há, ainda, uma simetria entre *nudges* e *sludges*: como subproduto de um *nudge*, haverá um aumento das fricções relativas das opções não influenciadas (*sludge*); *sludges*, por sua vez, reduzirão indiretamente as fricções relativas das opções alternativas àquela que é influenciada pelo *sludge*.

2. *Nudges* não fazem exigências quanto às motivações que justificam a sua implementação. Enquanto um *nudge* é uma das estratégias disponíveis ao regulador, o paternalismo libertário é um movimento que busca justificar políticas públicas que tentam proteger as pessoas de si mesmas, sem coerção ou mudança significativa de incentivos. *Nudges* podem ser implementados por diversos fundamentos, inclusive

welfaristas, isto é, com o objetivo de ampliar o bem-estar coletivo, ou exploratórios, quando são utilizados para promover o bem-estar do arquiteto de escolhas em detrimento do próprio agente decisor. O paternalismo libertário, além disso, pode englobar medidas que vão além dos *nudges*, já que não tentam explorar fatores supostamente irrelevantes, como o fornecimento de informações factuais.

3. A arquitetura de escolhas pressupõe uma relação jurídica prévia entre no mínimo um arquiteto de escolhas e um decisor. A natureza dessa relação influenciará nos limites e nas formas de controle da arquitetura de escolhas. *Nudges* implementados livremente por arquitetos de escolha no âmbito de uma relação privada são puramente privados e encontrarão limites nas regras específicas que cada ramo do direito – como o direito do consumidor, o direito da concorrência, o direito regulatório etc. – impõe aos agentes privados. Quando os limites se tornam obrigações positivas para organizar a arquitetura de escolhas e elaborar *nudges* de uma maneira específica, esses *nudges* são *nudges* impuros, que estão associados a uma medida coercitiva a um arquiteto de escolhas privados. *Nudges* puros, por outro lado, são identificados quando o arquiteto de escolhas é a própria administração pública, como na prestação de serviços públicos e na relação entre administração e servidores públicos.

Estabelecidas essas questões conceituais, é conveniente ter uma boa apreensão de como *nudges* funcionam na prática. Considerando as questões conceituais já enfrentadas nesta dissertação, o Capítulo 5 tentará demonstrar o verdadeiro alcance dos *nudges* à luz da literatura. Mais do que isso, serão apontados desafios que arquitetos de escolha devem superar para maximizar a efetividade dos *nudges*. Essa é uma tarefa prévia imprescindível à análise da legitimidade dos *nudges*, que será discutida no Capítulo 6.

5. *NUDGES* E EFETIVIDADE: FERRAMENTAS ÚTEIS, COMPLEMENTARES E COM BOM CUSTO-BENEFÍCIO PARA PROMOVER A MUDANÇA DE COMPORTAMENTO

Um primeiro passo essencial à discussão da legitimidade dos *nudges* – que será empreendida no Capítulo 6 – é compreender, empiricamente, qual é o seu impacto. Há quem afirme que *nudges* não são uma ferramenta útil em busca da solução de verdadeiros problemas em termos de políticas públicas. Selinger e Whyte (2012, p. 30), por exemplo, afirmam que não há evidências sólidas de que *nudges* poderiam realmente contribuir de maneira significativa para políticas públicas. Segundo esses autores, o *hype* em torno dessa abordagem se deve a uma retórica política sempre inflada. Em sua visão, *nudges* não conseguiriam resolver problemas complexos de maneira confiável e sem controvérsia (SELINGER; WHYTE, 2012, p. 31).

Essa crítica de Selinger e Whyte é expressamente desenvolvida a partir da obra de Adam Burgess (2012), que se debruçou sobre a aplicação de *nudges* a questões relacionadas à saúde pública. Além de considerações a respeito da introdução dos *nudges* no Reino Unido, Burgess faz consideráveis ressalvas à efetividade dessa técnica de intervenção. A princípio, Burgess afirma que *nudges* só são baseados em evidências no sentido de se originarem de equívocos comportamentais identificados em experimentos (BURGESS, 2012, p. 6). É por essa razão, segundo o seu argumento, que o exemplo da mosca desenhada em mictórios do aeroporto de Amsterdam[252] seria tão difundido: não haveria muitos outros exemplos de sucesso claro.

Feitas essas considerações, Burgess (2012, p. 13) defende que *nudges* só fariam sentido nas situações em que o problema atacado é claramente comportamental, em vez de estar relacionados à falta de recursos ou de tecnologia. Desse modo, *nudges* seriam aplicáveis apenas a contextos bastante particulares, já que "raramente há algo como um problema exclusivamente comportamental" (BURGESS, 2012, p. 13)[253]. A partir de tal perspectiva, Burgess (2012, p. 15-16) conclui argumentando que, para problemas complexos, como questões ligadas à saúde pública, *nudges* não seriam soluções vantajosas. Na verdade, no contexto analisado por Burgess, *nudges* não seriam uma alternativa às políticas tradicionais, mas apenas a última adição ao portfólio de abordagens intervencionistas de política pública.

Nesse contexto, uma série de estudos buscou estimar os verdadeiros efeitos de diversos tipos de *nudges*, em cenários variados. Compreender esses resultados é necessário, inicialmente, para se identificar o alcance da utilização de *nudges*, isto é, os tipos de problemas

[252] Vide nota 251 *supra*.
[253] Tradução livre. Texto original: "[...] *there is rarely such a thing as an exclusively behavioral issue,* […]".

cuja solução pode ser fomentada por *nudges*. Ao mesmo tempo, essa demonstração também permite que se avalie de maneira mais precisa os potenciais prejuízos e riscos que podem surgir da utilização de *nudges*. Afinal, se *nudges* forem muito efetivos e provocarem uma mudança ampla de comportamento, *nudges* utilizados com finalidades ilegítimas geram ainda mais preocupações.

O objetivo deste Capítulo é tentar enfrentar a crítica posta à efetividade dos *nudges*. Afinal, *nudges* são uma ferramenta *efetiva* na elaboração de políticas públicas ou são fruto de uma retórica política inflada? Para tentar responder essa questão, serão discutidos: a) os dados identificados na literatura a respeito da efetividade dos *nudges*; b) os desafios à escalabilidade de qualquer política pública e a avaliação dos efeitos *nudges* em longo prazo; c) algumas situações específicas em que se pode esperar que *nudges* sejam, notadamente, quando as pessoas tiverem fortes preferências antecedentes aos *nudges*, quando agentes do mercado se opuserem ao resultado do *nudge*, e quando essa intervenção desencadear a reatância psicológica; e d) a posição dos *nudges* como ferramenta complementar no arcabouço de técnicas regulatórias.

5.1. Revisão de literatura: *nudges* são efetivos como ferramenta para mudança do comportamento

A crítica posta por Burgess (2012) e por Sellinger e Whyte (2012) é essencialmente empírica. Para respondê-la, então, é necessário avaliar os dados existentes na literatura e verificar se *nudges* são, de fato, efetivos e têm um custo-benefício razoável. Para enfrentar essa questão, serão apresentados dados de meta-análises e revisões sistemáticas ou quase sistemáticas que se dedicaram a essa questão, que permitem uma visão geral dos tamanhos de efeitos de *nudges* em experimentos publicados em periódicos acadêmicos[254].

Inicialmente, em uma abordagem cognitiva que separa os processos afetados por *nudges* e *sludges*, Yu Luo e outros autores (2021) conduziram uma meta-análise de estudos randomizados controlados (*randomized controlled trials – RCTs*), envolvendo *nudges* e *sludges*. Foram analisados 188 experimentos, e, com relação aos *nudges*[255], especificamente,

[254] Não há, aqui, a intenção de esgotar todas as revisões sistemáticas e meta-análises já conduzidas na literatura. Serão apresentados alguns exemplos significativos. Uma meta-análise recente voltada à comparação de experimentos acadêmicos com *nudges* implementados na prática será apresentada na Seção 5.2, já que ela é fundamental à avaliação de *nudges* em larga escala.

[255] Luo e outros autores (2021) definiram *nudges* como intervenções que reduzem fricções na arquitetura de escolhas, o que se alinha à definição adotada nesta dissertação.

Luo e seus colegas identificaram um tamanho de efeitos médio equivalente e 0,33 no índice *d* de Cohen[256]. Assim, segundo essa meta-análise, *nudges* produzem efeitos significativos em termos estatísticos, com tamanho de efeito pequeno para médio.

Ao comparar os efeitos das intervenções de acordo com os processos cognitivos envolvidos - em linha com os processos descritos para diferenciar a forma como *nudges* e *sludges* atuam, conforme exposto no Capítulo 4 -, Luo e seus colaboradores identificaram efeitos mais significativos com *nudges* que reduziam o esforço necessário para se escolher uma determinada opção (0,62 no índice *d* de Cohen, isto é, efeitos médios para grandes). A alteração de regras *default* assim como o aumento de acessibilidade e conveniência, ou a simplificação, são exemplos de técnicas que reduzem o nível de esforço associado a determinadas decisões. Esse tamanho de efeito era superior - de maneira estatisticamente confiável – ao de intervenções que tentavam influenciar a motivação intrínseca do sujeito para facilitar determinados comportamentos[257].

Em outra meta-análise recente, Stephanie Mertens e outros autores (2022) analisaram mais de 200 estudos envolvendo a aplicação de *nudges*. Seus resultados indicaram que essas alterações da arquitetura de escolhas promovem a mudança do comportamento com tamanho de efeito pequeno para médio (índice *d* de Cohen equivalente a 0,45).

Além disso, a partir de uma definição ampla de *nudge*, Mertens e seus colaboradores dividiram os *nudges* em três categorias:

a) *nudges* que alteram as informações relativas à decisão, como a tradução de informações, o aumento da visibilidade das informações ou o fornecimento de um ponto de referência social;
b) nudges que alteram a estrutura das decisões, como a alteração de regras *default*, as mudanças do esforço associado a determinada opção (redução de fricções), a adaptação das categorias ou o agrupamento de opções de escolhas e a alteração das consequências das decisões;

[256] O índice de Cohen é uma medida de magnitude de efeitos utilizada para descrever a diferença média padronizada de um determinado efeito, isto é, o número de desvios padrão entre os resultados do grupo experimental e do grupo controle, para testar uma determinada intervenção. De acordo com esse índice, representado pela letra *d*, um tamanho de efeitos equivalente a 0,2 seria considerado pequeno, enquanto tamanhos de efeitos equivalentes a 0,5 e a 0,8 seriam considerados, respectivamente, médios e grandes. Cf., nesse sentido: (LAKENS, 2013; LENHARD; LENHARD, 2016).

[257] Luo e seus coautores também avaliaram se haveria uma distinção de efetividade dos *nudges* entre diferentes domínios – educação, meio-ambiente, finanças, saúde e política. Não foram identificados efeitos significativos desses domínios no resultado dos experimentos.

c) *nudges* que auxiliam as decisões, como o fornecimento de lembretes ou o estímulo ao compromisso.

De todas essas categorias, *nudges* que alteram a estrutura das decisões - e que se aproximam da definição estrita adotada nesta dissertação – são consistentemente mais efetivos[258] que *nudges* que alteram as informações relativas à decisão[259] e que *nudges* que auxiliam a tomada de decisões[260]. A alteração de regras *default*, inclusive, é, em meio aos *nudges*, a técnica que promove a alteração do comportamento de maneira mais efetiva[261].

Ao especular as razões para essa diferença, Mertens e seus coautores (2022, p. 7) afirmam que a informação decisória e o auxílio à tomada de decisões são técnicas que tentam facilitar processos decisórios deliberativos; a alteração da estrutura decisória, por seu turno, busca estimular processos deliberativos automáticos. Nesse caso, as primeiras categorias de *nudge* podem apresentar resultados mais modestos por excederem os limites cognitivos das pessoas para tomada de decisões de maneira mais frequente, especialmente em situações de alta carga cognitiva ou pressão temporal.

Por outro lado, Mertens e seus colaboradores também avaliaram a efetividade de *nudges* de acordo com o domínio em que a técnica é utilizada – saúde, alimentação, meio ambiente, finanças, pró-sociais e outros. Os dados de sua meta-análise indicam que intervenções na arquitetura de escolhas no domínio da alimentação são mais efetivas do que em outros domínios - tamanho de efeito médio para alto, equivalente a 0,72 no índice *d* de Cohen, o que equivale a cerca de 2,5 vezes os tamanhos de efeitos médios de setores como saúde, meio-ambiente, finanças, pró-sociais e outros. Esse valor é ainda maior se consideradas apenas alterações da estrutura de decisões no setor da alimentação – tamanho de efeito alto, equivalente a 0,86 no índice *d* de Cohen.

John Beshears e Harry Kosowsky (2020), por outro lado, documentaram efeitos estatísticos significantes em *nudges* aplicados em diferentes disciplinas[262], setores de aplicação, estruturas de pesquisa e tipos de *nudges* distintos. Como Beshears e Kosowky não fizeram uma meta-análise, os autores não chegaram a um tamanho de efeito médio em comum a todo e

[258] Tamanho de efeito médio – índice *d* de Cohen equivalente a 0,55.
[259] Tamanho de efeito pequeno para médio – índice *d* de Cohen equivalente a 0,38.
[260] Tamanho de efeito pequeno para médio – índice *d* de Cohen equivalente a 0,31.
[261] Tamanho de efeito médio – índice de Cohen *d* equivalente a 0,62.
[262] Beshears e Kosowsky identificaram a disciplina acadêmica a que cada um dos estudos selecionados estava vinculada – economia e finanças; ciência ambiental; marketing e comportamento do consumidor; medicina; psicologia e ciências cognitivas; saúde pública; e outras. A análise de tais resultados ultrapassa o escopo desta dissertação.

qualquer tipo de *nudge*. Beshears e Kosowsky também não identificaram os tamanhos de efeitos de todos os experimentos analisados, por falta de informações.

De qualquer forma, um primeiro aspecto analisado por Beshears e Koswoski foi a efetividade dos *nudges* a depender dos seus respectivos domínios de aplicação. Os dados coletados indicavam que os *nudges* mais significativos parecem estar ligados ao meio-ambiente e à saúde - índices *d* de Cohen de 0,421 e 0,416, respectivamente, sendo 0,504 nos experimentos ligados exclusivamente à alimentação saudável. Ao mesmo tempo, *nudges* que estimulam um comportamento pró-social e *nudges* aplicados ao setor financeiro têm efeitos mais modestos – índice *d* de Cohen de 0,213 e 0,316, respectivamente[263].

Além disso, Beshears e Kosowski associaram a efetividade dos *nudges* a uma variável de seu funcionamento: a medida em que um *nudge* envolvia processos automáticos[264]. Ao controlar essa variável, Beshears e Kosowski destacaram que *nudges* que envolvem processos automáticos apresentam tamanhos de efeito aproximadamente 50% superiores aos de *nudges* que não envolvem tais processos - a automaticidade se traduzia em um aumento de 0,193 no índice *d* de Cohen. Com relação às categorias específicas, de acordo com a análise de Beshears e Kosowski, *nudges* que circundam os Sistemas 1 e 2 têm efeitos mais significativos do que *nudges* que desencadeiam o Sistema 1 e que promovem o Sistema 2, mas as diferenças não foram significativas em termos estatísticos.

Em outra pesquisa relevante, Dennis Hummel e Alexander Maedche (2019), em uma extensa revisão de literatura, demonstraram que as magnitudes de efeito identificadas em experimentos envolvendo *nudges* são muito diversas. Dos 308 efeitos analisados por Hummel e Maedche, 190 ou 62% dos efeitos apresentaram relevância estatística (p valor igual ou inferior a 0,05). Em outras palavras, mais de um terço dos efeitos analisados eram insignificantes em termos estatísticos. Os efeitos significativos em termos estatísticos apresentavam um tamanho de efeito relativo mediano equivalente a 39%[265]. De todo modo, Hummel e Maedche

---

[263] Beshears e Kosowski diferenciaram os *nudges* segundo três categorias: a) *nudges* que desencadeiam processos do Sistema 1; b) *nudges* que promovem os processos cognitivos reflexivos do Sistema 2; e c) *nudges* que circundam esses sistemas e removem o indivíduo de algum aspecto do processo de tomada de decisões. Beshears e Kosowski ressaltaram, ainda, que um *nudge* poderia se enquadrar em mais de uma categoria ao mesmo tempo, por operar por múltiplos mecanismos simultaneamente. Embora a pesquisa tenha identificado distinções de tamanhos de efeito segundo a categoria de *nudge*, essas diferenças não foram estatisticamente significativas.

[264] Todas as categorias de *nudges* definidas por Beshears e Kosowski poderiam se basear em processos automáticos do comportamento. Em outras palavras, na sua visão, a exploração de processos automáticos pode ser utilizada para desencadear processos do Sistema 1, promover processos do Sistema 2 ou circundar esses sistemas.

[265] Em outras palavras, diante de um *nudge* específico, havia um acréscimo de 39% na incidência do comportamento promovido, em comparação com o grupo controle. Hummel e Maedche não conseguiram reunir todos os dados necessários para calcular o índice *d* de Cohen, já que muitos experimentos não indicavam o desvio padrão, o que é necessário para esse cálculo. Ao reunir apenas os 49 efeitos em que esse dado estava disponível,

identificaram distinções a depender da categoria de *nudge* analisada e do contexto em que eles se aplicavam.

Partindo de uma definição ampla de *nudge* e utilizando categorias previamente definidas por Sunstein (2014d, 2015c, 2019a, 2019c) e mencionadas no Capítulo 3, Hummel e Maedche identificaram que, de todas as categorias, as regras *default* são, de longe, os *nudges* com maior efetividade - em geral, a magnitude mediana dos efeitos das regras *default* em relação ao comportamento observado era de 50%. As demais categorias - simplificação, referência social, mudança de esforços, divulgação de informações, avisos/gráficos, pré-compromisso, lembretes, desencadeamento de intenções de implementação e *feedbacks* - apresentavam magnitudes de efeitos medianas entre 7% (pré-compromisso) e 39% (desencadeamento de intenções de implementação).

Além disso, Hummel e Maedche verificaram um tamanho de efeitos mediano de 21%, se considerados todos os contextos decisórios em que essas intervenções poderiam ser aplicadas. Ao avaliar contextos específicos, Hummel e Maedche identificaram que *nudges* apresentavam efeitos mais robustos nos contextos da proteção ambiental (39%) e da privacidade (44%). Na elaboração de políticas e no consumo de energia, por sua vez, os efeitos eram mais modestos (respectivamente, 6% e 13%). Efeitos intermediários foram identificados no contexto da saúde (21%) e das finanças (28%).

Benartzi e seus colegas (2017), por sua vez, calcularam – de maneira seletiva, mas sistemática – o impacto de *nudges* em comparação com ferramentas tradicionais, com base na razão entre o efeito causal de uma intervenção no resultado de determinada variável e o seu custo de implementação, ajustado pela inflação. Os autores investigaram políticas distintas em quatro cenários: a) poupança para a aposentadoria; b) inscrições para a universidade; c) conservação de energia; e d) vacinação contra a gripe. Foram comparados *nudges* com medidas tradicionais como incentivos fiscais, fornecimento de informações, educação etc.

Com relação à poupança para aposentadoria, Benartzi e seus coautores concluíram que um *nudge* que promovia decisões ativas aumentava as contribuições por ano em 100 dólares para cada dólar gasto. Essa medida era muito mais efetiva do que técnicas como incentivos fiscais (2,77 dólares por dólar gasto na Dinamarca e 1,24 dólares por dólar gasto nos Estados Unidos), fornecimento de informações (14,58 dólares por dólar gasto) e medidas em que o empregador deveria fazer contribuições correspondentes a até 20% ou 50% das contribuições feitas pelos empregados (respectivamente, 5,59 e 2,97 dólares por dólar gasto).

---

extraídos de 11 experimentos, Hummel e Maedche identificaram um índice *d* de Cohen mediano de 0,41 (tamanho de efeito pequeno para médio).

Já quanto à inscrição de alunos em faculdades, um *nudge* que simplificava o processo de solicitação de auxílio estudantil (*Free application for Federal Student Aid – FAFSA*) tinha como resultado um aumento de 1,53 estudantes inscritos a cada 1.000 dólares gastos. Outras medidas tinham um custo-benefício muito limitado, como o fornecimento de uma quantia mensal para estudantes (0,0351 estudantes inscritos a cada 1.000 dólares gastos), o fornecimento de subsídios monetários (0,0051 estudantes inscritos a cada 1.000 dólares gastos) e a concessão de créditos tributários (efeitos negligenciáveis a cada 1.000 dólares gastos).

A conservação de energia, segundo o estudo de Benartzi e seus colegas, também era promovida com maior eficiência a partir de *nudges*. Nesse sentido, um *nudge* que explorava normas sociais promovia uma redução no consumo de energia equivalente a 27,3 kWh a cada dólar gasto. Mais uma vez, a efetividade de medidas alternativas era inferior, como se constatou para um *nudge* que ligava o uso de energia a questões de saúde e questões ambientais (redução de 0,050 kWh por dólar gasto), um *nudge* que fornecia informações na conta (redução negligenciável de kWh por dólar gasto), descontos na conta de energia elétrica (redução de 3,41 kWh por dólar gasto) e incentivos e educação (redução de 14 kWh por dólar gasto).

Por fim, Benartzi e seus coautores também demonstraram que *nudges* tinham bom custo-benefício para aumentar a vacinação contra a gripe. Um *nudge* que estimulava o planejamento para vacinação, por exemplo, aumentou o número de adultos vacinados em 12,8 a cada 100 dólares gastos. O agendamento automático (regra *default*) de uma data para vacinação ampliava esse número em 3,65 adultos a cada 100 dólares gastos. Uma campanha educacional para estimular a vacinação apresentava efeitos intermediários - aumento em 8,85 adultos vacinados a cada 100 dólares gastos –, enquanto outras estratégias, como incentivos monetários e vacinação gratuita no ambiente de trabalho, apresentaram efeitos mais modestos (respectivamente, aumento de 1,78 e 1,07 adultos vacinados a cada 100 dólares gastos.

Assim, de maneira geral, embora ressaltem que mais estudos são necessários para definir a questão, Benartzi e seus coautores concluíram que o impacto dos *nudges* é frequentemente maior do que o de ferramentas tradicionais, se analisados os seus custos-benefícios. Em sua visão, as vantagens comparativas de *nudges* são maiores quando o objetivo do *policy-maker* é modificar o comportamento cotidiano de indivíduos que estão tomando decisões imperfeitas, enviesadas e apressadas. Em outras palavras, embora possam ter efeitos modestos, considerando os baixos custos para sua implementação, *nudges* podem ser eficientes e ter um impacto por unidade de custo superior ao de outras estratégias.

Romain Cadario e Pierre Chandon (2020), por sua vez, conduziram uma meta-análise de experimentos de campo envolvendo *nudges* que têm por objetivo aumentar o consumo de

alimentos saudáveis.[266] Cadario e Chandon controlaram o tipo de intervenção segundo três critérios diferentes:

a) *Nudges* orientados para a cognição tentam influenciar o que os consumidores *sabem*. Envolvem, por exemplo, rótulos de nutrição descritivos, rótulos de nutrição avaliativos e aprimoramentos visuais, isto é, estratégias que informam aos consumidores a disponibilidade de opções saudáveis ao ampliar sua visibilidade;

b) *Nudges* orientados para o afeto tentam influenciar como os consumidores se sentem, sem necessariamente mudar o que eles sabem. Os exemplos de Cadario e Chandon são aprimoramentos hedônicos, que buscam ampliar o apelo hedônico (emocional) de opções saudáveis a partir de descrições hedônicas vívidas ou de disposições de alimentos de forma atrativa, e ligações telefônicas em prol da alimentação saudável, que encorajam diretamente as pessoas a melhorarem a sua alimentação;

c) *Nudges* orientados para o comportamento tentam impactar os comportamentos das pessoas sem necessariamente alterar o que elas sabem ou como elas se sentem, às vezes sem a sua própria consciência. Aprimoramentos de conveniência, que tornam fisicamente mais fácil o consumo ou a seleção de opções saudáveis ou dificultam o consumo de opções não saudáveis, e mudanças no tamanho de pratos, tigelas e copos são exemplos dessa modalidade de *nudge*.

Inicialmente, ao avaliar 299 tamanhos de efeitos derivados de 90 artigos e 96 experimentos, Cadario e Chandon constaram um tamanho de efeito médio de *nudges* para alimentação saudável equivalente a 0,23 no índice *d* de Cohen, o que equivale a um tamanho de efeitos pequeno. Em uma estimativa de consumo de alimentos diário equivalente a 1.727 kcal, esse tamanho de efeitos equivaleria a uma redução de 124 kcal ou pouco menos de 8 colheres de sopa de açúcar no consumo diário de um adulto.

Há, ainda, diferenças substanciais a depender do tipo de *nudge* adotado. Intervenções orientadas para a cognição - que, em grande medida, não se enquadram na definição de *nudge*

[266] Cadario e Chandon utilizam uma série de experimentos do cientista comportamental Brian Wansink em sua meta-análise. Wansink era Professor da Universidade de Cornell e coordenava o *Food and Brand Lab*, tendo sido reconhecido como um dos mais famosos cientistas comportamentais das últimas décadas. Apesar disso, reconheceu-se posteriormente que Wansink estava falsificando resultados, e inúmeros de seus estudos foram objeto de retratação (LIST, 2022, p. 38-39). Atentos a essas críticas, Cadario e Chandon incluíram uma variável binária adicional para controlar as 51 observações que se originaram do laboratório de Wansink. Essa variável, no entanto, não indicou alterações significativas estatisticamente, o que valida os resultados obtidos por Cadario e Chandon.

adotada nesta dissertação - apresentaram efeitos limitados - 0,12 no índice *d* de Cohen, isto é, efeitos quase insignificantes em termos estatísticos. A redução no consumo promovida por essas intervenções equivalia a 64 kcal ou 4 colheres de sopa de açúcar. *Nudges* orientados para o afeto, por sua vez, apresentaram tamanhos de efeitos pequenos - 0,24 no índice *d* de Cohen ou uma redução de consumo equivalente a 129 kcal ou 8,1 colheres de sopa de açúcar por dia. Finalmente, *nudges* orientados para o comportamento apresentaram tamanhos de efeitos pequenos para médios - 0,39 no índice *d* de Cohen ou uma redução de 209 kcal ou 13,1 colheres de sopa de açúcar por dia. Dentre essas intervenções, o efeito de mudanças no tamanho de pratos e outros utensílios foi a intervenção mais efetiva, segundo Cadario e Chandon: 0,59 no índice *d* de Cohen, isto é, um tamanho de efeitos médio para grande.

Finalmente, Cadario e Chandon identificaram alguns dados interessantes a partir do controle de variáveis ligadas aos métodos experimentais utilizados, à população envolvida e aos diferentes tipos de comportamento alimentar que foram influenciados.

A princípio, técnicas voltadas à redução de alimentos não saudáveis apresentaram um tamanho de efeito 30% superior ao de técnicas ligadas ao estímulo do consumo de alimentos saudáveis. Ao mesmo tempo, intervenções voltadas à redução do volume de alimentação como um todo são menos efetivas do que técnicas voltadas à redução de medidas específicas de alimentação. Em outras palavras, é mais fácil fazer com que as pessoas deixem de comer frituras do que estimular o consumo de verduras; fazer com que elas simplesmente comam menos, por sua vez, é ainda mais difícil do que as duas tarefas anteriores.

Além disso, Cadario e Chandon verificaram que os efeitos em estudos conduzidos nos Estados Unidos foram 47% maiores do que aqueles identificados em outros países.[267] Na interpretação de Cadario e Chandon, isso pode ser explicado pelo menor foco de americanos na experiência de alimentar-se, em comparação aos efeitos à saúde, ou pode ser que americanos se baseiem mais em sinais externos do que nos internos. Também se levantou a hipótese de que esse número tenha sido causado por uma maior proporção de sobrepeso nos Estados Unidos, assim como um maior tamanho das porções.

Há, ainda, distinções marcantes quanto ao local em que as intervenções foram aplicadas: *nudges* em mercearias e supermercados foram, em média, 146% menos efetivos do que aqueles aplicados em restaurantes e cafeterias. Na hipótese de Cadario e Chandon, ainda que mais pesquisas sejam necessárias para validar esse ponto, é mais fácil responder a uma intervenção quando se escolhe para si próprio, a partir de um volume limitado de opções e para consumo

[267] Foram investigados experimentos nos seguintes países: Bélgica, Canadá, França, Irlanda, Israel, Japão, Países Baixos e Reino Unido (CADARIO; CHANDON, 2020).

imediato, do que quando se escolhe para toda a família, a partir de uma miríade de opções e para múltiplas ocasiões de consumo.

Por fim, Jon M. Jachimowicz e seus coautores (2019) fizeram uma meta-análise de efeitos de regras *default* que também merece ser mencionada. A partir de 58 conjuntos de dados extraídos de 56 estudos publicados em 35 artigos, os autores identificaram que regras *default* que presumem o consentimento (*opt-out defaults*) ampliam consideravelmente a decisão pré-selecionada em comparação a regras que exigem consentimento explícito (*opt-in defaults*) - índice *d* de Cohen equivalente a 0,68, isto é, tamanho de efeitos médio para grande. Ao avaliar apenas variáveis medidas em escala binária - isto é, opções ligadas a escolhas do tipo "sim" ou "não", em vez de opções contínuas (por exemplo, a quantia de dinheiro doada ou investida) -, Jachimowicz e seus colegas identificaram um aumento de 27,24% na opção pré-selecionada a partir da alteração da regra *default*.

Apesar disso, Jachimowicz e seus coautores também identificaram bastante variabilidade nos tamanhos de efeito dos experimentos que foram objeto de sua meta-análise. Diante de tal heterogeneidade, "o verdadeiro efeito de qualquer estudo de default pode ser substancialmente menor ou maior do que a média meta-analítica" (JACHIMOWICZ *et al*., 2019, p. 14)[268]. Nesse cenário, Jachimowicz e seus colegas tentaram evidenciar os fatores que explicam o sucesso de uma determinada regra *default*, a partir dos atributos dos estudos analisados, a fim de reduzir a heterogeneidade de seus resultados.

Em uma primeira conclusão, os autores verificaram que *defaults* são mais efetivos em domínios consumeristas do que ambientais ou de saúde. Em seguida, Jachimowicz e seus coautores tentaram verificar se os canais psicológicos pelos quais uma regra *default* influencia decisões auxiliam a explicar a variabilidade de tamanhos de efeito que foi encontrada.

De acordo com Isaac Dinner e outros autores (2011), em estudo expressamente citado na meta-análise de Jachimowicz e seus colegas, *defaults* influenciam as decisões a partir de três canais psicológicos distintos. Em primeiro lugar, essa influência pode acontecer pela aprovação (*endorsement*), quando os decisores acreditam que as intenções do arquiteto de escolhas, materializadas na arquitetura de escolhas e na regra *default*, podem ser benéficas. Ao mesmo tempo, os decisores podem ser influenciados pela facilidade (*ease*), quando manter a opção pré-selecionada permite a decisão que exige o menor esforço dos decisores. Finalmente, a influência também pode ser explicada pelo efeito dotação (*endowment*), já mencionado no Capítulo 1, uma vez que o sujeito passa a avaliar outras opções em comparação à opção pré-selecionada

[268] Tradução livre. Texto original: "[...] *the true effect of any given default study could be substantially lower or higher than the meta-analytic average*".

que já estaria ligada a ele e passa a refletir o *status quo*. Perder essa opção, nesse caso, gera uma reação negativa no sujeito, em razão da aversão à perda.

Nesse contexto, Jachimowicz e seus coautores analisaram a sua amostragem de estudos com base nos mecanismos explorados pelos *defaults*. Seus resultados indicaram que tanto a aprovação quanto o efeito dotação são moderadores dos efeitos de um *default*, enquanto a facilidade não apresentou resultados estatisticamente significativos. Nesse caso, haveria uma expectativa de que um *default* que ativasse canais de aprovação e de dotação tivesse um tamanho de efeitos médio para grande.

Tabela 5 - Revisão bibliográfica da efetividade de *nudges*[269]

Artigo	Tamanho de efeito médio	Tipo de *nudge* mais efetivo	Domínio com efeitos mais significativos	Domínio com efeitos menos significativos	Tipos de experimentos analisados
Yu Luo et al., 2021	0,33 no índice *d* de Cohen (pequeno para médio)	Redução do esforço necessário para escolher uma opção	Não houve diferença significativa	Não houve diferença significativa	RCTs
Mertens et al., 2022	0,45 no índice *d* de Cohen (pequeno para médio)	Alteração da estrutura das decisões	Alimentação	Finanças	Variados
Beshears e Kosowsky, 2020	Não calculado	Exploração de processos automáticos	Meio-ambiente e saúde (incluindo alimentação)	Comportamento pró-social e financeiro	Variados
Hummel e Maedche, 2019	21%	Regras *default*	Meio-ambiente e privacidade	Elaboração de políticas e consumo de energia	Variados
Cadario e Chandon, 2020	0,23 no índice *d* de Cohen (pequeno)	Orientados para o comportamento	N/A[270]	N/A	Experimentos de campo
Jachimowicz et al., 2019	0,68 no índice *d* de Cohen (médio para grande)	N/A[271]	Consumo	Meio-ambiente e saúde	Variados

Fonte: Elaboração do Autor.

A partir de todas as meta-análises e revisões bibliográficas analisadas acima, é possível concluir, em primeiro lugar, que *nudges* têm efeitos estatisticamente significativos. De maneira geral, seus tamanhos de efeito são pequenos para médios, com a ressalva de que determinados tipos de *nudge* – especialmente aqueles ligados à definição estrita adotada nesta dissertação,

[269] A revisão sistemática de Benartzi e outros autores (2017) não foi incluída por avaliar custo-benefício, em vez da efetividade propriamente dita.
[270] A meta-análise reuniu apenas dados de experimentos ligados à alimentação saudável.
[271] A meta-análise reuniu apenas alterações de regras *default*.

associada à exploração de fatores supostamente irrelevantes baseados em processos intuitivos - têm efeitos mais significativos.

Do mesmo modo, há domínios mais propensos à utilização de *nudges*, como a alimentação, a saúde, o consumo e o meio ambiente; em outros casos, como nas finanças, *nudges* parecem ter efeitos mais modestos. A natureza e as consequências das decisões em cada um desses domínios pode explicar essa distinção. Afinal, decisões financeiras frequentemente envolvem desembolso, o que pode gerar um desconforto para os agentes – a já mencionada dor do pagamento[272] -, o que pode agir de maneira contrária aos *nudges*.

Os seus efeitos podem, ainda, se prolongar no tempo e impactar decisões futuras dos agentes. De fato, ter menos recursos financeiros por investir uma determinada parcela de seu salário em conformidade com um *nudge*, por exemplo, é algo que trará impacto para todas as despesas que o sujeito terá, especialmente em casos de maior escassez – tema que será retomado no Capítulo 6. Em outros casos, além de não haver propriamente um envolvimento de despesas para os indivíduos, as decisões podem se esgotar imediatamente. Ao consumir um alimento saudável, por exemplo, por mais que a decisão tenha impactos no futuro - na saúde, por exemplo –, o ato de alimentar-se esgota naquele mesmo momento. Suas consequências não se traduzem imediatamente às escolhas futuras do indivíduo, embora possam afetá-las.

O custo-benefício dos *nudges*, por outro lado, parece ser superior ao de outras estratégias, embora mais estudos a esse respeito sejam necessários em busca de conclusões definitivas. De fato, depois de demonstrar que o custo-benefício dos *nudges* parece ser excelente, Benartzi e seus coautores (2017) ressaltam que não estão comparando maçãs-com-maçãs. As estratégias avaliadas envolvem populações de tamanho diverso, e os autores não identificaram estudos suficientes que comparassem múltiplas intervenções ao mesmo tempo em populações similares. Essa é uma agenda de pesquisa que deve ser promovida, para uma compreensão mais ampla do custo-benefício dos *nudges*.

De qualquer forma, a revisão bibliográfica a respeito da efetividade dos *nudges* apresentada acima indica que, diferentemente do que alegam Burgess (2012) e Sellinger e Whyte (2012), a literatura apresenta evidências robustas de que nudges podem ser utilizados como ferramenta regulatória efetiva para alterar o comportamento humano. Não por outra razão, *nudges* vem sendo adotados em políticas públicas de inúmeros países, como já mencionado ao longo desta dissertação.

[272] Cf. nota 210 *supra*.

5.2. *Nudges* persistentes e em larga escala: caminhos para evitar quedas de tensão

A Seção 5.1 demonstrou que *nudges* têm efeitos estatisticamente relevantes e, em muitos contextos, são um exemplo de política baseada em evidências científicas. A transformação de tais evidências em políticas públicas efetivas, no entanto, envolve uma série de desafios. Afinal, em um cenário experimental, com número limitado de participantes, *nudges* podem ser efetivos, com excelente custo-benefício. Políticas públicas reais, no entanto, não são feitas em placas de Pétri, e a realidade é mais complexa que cenários experimentais. Transformações em maior escala não são imediatamente realizáveis e inúmeros desafios podem impedir a tradução de evidências científicas em mudança de comportamento em larga escala. Esta Seção se dedica a expor esses desafios.

5.2.1. Evitando quedas de tensão: armadilhas que podem limitar a escalabilidade de *nudges*

O economista americano John A. List (2022) publicou recentemente um livro a respeito da ciência da escalabilidade (*scaling*), em que o efeito tensão (*the voltage effect*) é minuciosamente descrito para explicar por que algumas ideias obtêm ganhos de escala e outras acabam por falhar ao serem expandidas. List aponta, inicialmente, cinco fatores que levam a quedas de tensão (*voltage drops*), isto é, fatores que impedem a implementação de uma ideia – aqui, incluídas políticas públicas – em grande escala. Compreender o efeito tensão e evitar quedas de tensão são etapas necessárias à tentativa de maximizar os efeitos de *nudges*.

O primeiro fator que leva a quedas de voltagem são falsos positivos, isto é, a interpretação de evidências como prova de que algo é verdadeiro, quando essa não é, de fato, a realidade (LIST, 2022, p. 23). List menciona que dados podem "mentir" em várias situações. Judea Pearl e Dana Mackenzie (2018) indicam, por exemplo, que há viés de confusão (*counfounding bias*), quando a verdadeira relação causal entre duas variáveis X e Y é atrapalhada por uma terceira variável Z que gera uma correlação espúria entre X e Y. Se, por exemplo, uma droga é testada em pacientes que são mais jovens, em média, que as pessoas do grupo controle, a idade se torna um elemento de confusão, e, se não forem coletados dados a respeito das idades, não será possível separar efeitos verdadeiros de espúrios (PEARL; MACKENZIE, 2018, p. 137-138).

No caso da implementação de um *nudge*, o arquiteto de escolhas deve avaliar minuciosamente a evidência analisada, para tentar afastar correlações espúrias que estejam ofuscando a verdadeira relação causal entre a intervenção e os resultados pretendidos. O ideal

é que falsos positivos sejam detectados antes de implementar uma política pública em larga escala. Afinal, como List acrescenta, se a descoberta ocorrer apenas depois de a ideia ter sido escalada, haverá queda de voltagem em larga escala; detectar um eventual erro cedo permite que o *policy-maker* tente outras estratégias para promover o objetivo buscado com sua política[273].

Nesse contexto, List (2022, p. 34-37) sugere que a replicação[274] dos resultados encontrados em determinado experimento é o caminho ideal para evitar falsos positivos. É importante, ainda, que essa replicação seja feita de maneira independente, sem que haja interesse em confirmar os resultados encontrados. Quanto maior for o conjunto de dados que embasa um determinado *nudge*, por exemplo, menor será a probabilidade de se estar diante de um falso positivo. Diante de resultados experimentais de que um determinado *nudge* pode ser efetivo, é recomendável que a estratégia proposta seja testada mais vezes no mesmo tipo de população e com a mesma configuração experimental para se ter mais confiança nos resultados. Contar com revisões sistemáticas e meta-análises de experimentos semelhantes, ao mesmo tempo, é um caminho para se garantir que os efeitos buscados foram confirmados em estudos semelhantes realizados posteriormente[275].

---

[273] Há, ainda, vieses cognitivos - fatores supostamente irrelevantes – que dificultam a visualização de falsos positivos. O viés da confirmação, já exposto no Capítulo 2, por exemplo, faz com que as evidências sejam interpretadas de maneira que se conforma às crenças prévias dos pesquisadores - e dos arquitetos de escolha. Outro viés que List (2022, p. 29-31) menciona é o efeito manada (*bandwagon bias*), decorrente de influências sociais em nossos processos mentais. No clássico experimento de Solomon Asch (1955), por exemplo, verificou-se que as decisões de outras pessoas afetavam as escolhas que seriam feitas por um indivíduo – as pessoas seguiam a opinião da maioria ainda que soubessem que as respostas estavam equivocadas. Esse viés pode fazer com que uma ideia seja indiscriminadamente adotada por toda uma equipe, o que pode levar à propagação de um falso positivo. A maldição do vencedor é o último fator supostamente irrelevante que List (2022, p. 33) menciona: sempre que houver concorrência por um bem de valor incerto, o vencedor do leilão costuma pagar mais do que o bem verdadeiramente vale. Boas ideias também funcionam dessa forma: o investimento para colocá-las em prática acaba sendo maior do que elas realmente valem; se detectado um falso positivo, nesse caso, é provável que já tenha sido despendida uma quantidade considerável de recursos, o que leva os agentes a insistirem tentando fazer a política dar certo, ainda que ela não seja escalável. A falácia dos custos afundados (*sunk costs fallacy*) corrobora esse efeito, já que, diferentemente do *Homo economicus*, as pessoas levam custos afundados em consideração ao tomar decisões (THALER, 2015, p. 64-66).

[274] A partir dos anos 2010, a replicação de experimentos se tornou um grave problema científico. List (2022, p. 37) menciona que um psicólogo organizou um trabalho para replicar 100 experimentos cujos resultados foram reputados em periódicos de grande prestígio acadêmico. Apenas 39 desses experimentos produziram os resultados que haviam sido reportados anteriormente.

[275] A forma como os incentivos acadêmicos está estruturada pode, ainda, estimular a falsificação de resultados, isto é, a corrupção científica (LIST, 2022, p. 39). Afinal, há um estímulo constante à produção acadêmica que valoriza a quantidade de publicações em periódicos relevantes para mensurar o sucesso dos cientistas. O exemplo já mencionado de Brian Wansink ilustra essa afirmação. Recentemente, um célebre experimento a respeito da desonestidade conduzido pelo célebre economista Dan Ariely e por outros autores (SHU *et al.*, 2012) foi objeto de retratação, por utilizar dados supostamente falsos - Cf., nesse sentido: (MARCUS, 2021). Para evitar esses problemas, é recomendável consultar a base de dados da *Retraction Watch* – http://retractiondatabase.org/RetractionSearch.aspx? - que reúne estudos e experimentos que foram objeto de retratação.

O segundo fator que pode levar a quedas de tensão é a falta de conhecimento do público que será afetado pela intervenção. Se o arquiteto de escolhas não compreender os diferentes tipos de pessoas que serão submetidos a uma determinada intervenção e os seus respectivos comportamentos, ele não conseguirá prever de forma precisa quem irá responder a um determinado *nudge* em larga escala nem entender a amplitude que sua proposta poderá obter (LIST, 2022, p. 54). As peculiaridades locais e culturais são extremamente relevantes nesse sentido[276]. Afinal, pode ser que uma determinada parcela da população tenha preferências específicas a respeito da direção à qual o *nudge* tenta empurrá-la, e, nesse caso, a política poderá ser inefetiva[277].

Se demandas e necessidades específicas de alguns segmentos da população forem ignoradas pelo arquiteto de escolhas, quedas de tensão inevitavelmente acontecerão ao expandir um *nudge* em larga escala (LIST, 2022, p. 55). Para evitar que isso aconteça, é essencial que, diante de dados que embasem a utilização de um determinado *nudge*, o arquiteto de escolhas verifique a representatividade da população afetada. A meta-análise conduzida por Mertens e outros autores (2022) não identificou diferenças estatisticamente significativas em estudos conduzidos dentro ou fora dos Estados Unidos; Cadario e Chandon (2020), por seu turno, ao avaliar *nudges* voltados à alimentação saudável, identificaram efeitos mais robustos nos Estados Unidos do que em outros países.

De qualquer forma, em cada experimento específico, essa é uma questão à qual os pesquisadores devem estar atentos: se, acidentalmente, a amostragem selecionada não for representativa da população, haverá viés de seleção[278]. Os resultados da pesquisa, então, não serão escaláveis. Para evitar esse problema, é recomendável, sempre que possível, que *nudges* sejam testados em estudos randomizados controlados (*randomized controlled trials* – RCTs), em que um tratamento A é atribuído aleatoriamente a alguns indivíduos e não para outros, e as

[276] Nesse contexto, é importante ter em vista que a maior parte dos estudos científicos que embasam a utilização de *nudges* é conduzida nos Estados Unidos. Joseph Henrich, Steven J. Heine e Ara Norenzayan (2010) argumentam que, em muitos casos, dados científicos obtidos em experimentos conduzidos em sociedades ocidentais, educadas, industrializadas, ricas e democráticas - representadas pelo acrônimo WEIRD (*Western, educated, industrialized, rich, and democratic*) - não são aplicáveis para contextos culturais diferentes. Populações WEIRD utilizam processos cognitivos diferentes de outras populações, o que pode reduzir a força de evidências científicas aplicadas apenas a essas sociedades. A experimentação no contexto cultural e local em que um determinado *nudge* será aplicado, nesse caso, é fundamental para o sucesso da intervenção.

[277] Burgess (2012, p. 12-13) pondera, nesse sentido, que um significativo grupo de pessoas não estaria disposto a ser "empurrado". A rejeição à adoção de um *nudge*, por motivos políticos, por exemplo, poderia ser a derrocada de uma política pública que nele estivesse baseada. Esse aspecto será retomado na Seção 5.3.

[278] Pearl e Mackenzie (2018, p. 325) afirmam que o viés de seleção ocorre quando o grupo da amostra que está a ser estudado é diferente da população-alvo de alguma forma relevante. Se, por exemplo, busca-se estimar a relação causal entre cigarros e o câncer, a população de pacientes hospitalizados com câncer não será representativa da população ou, mesmo, da população de fumantes (PEARL; MACKENZIE, 2018, p. 168). A incidência de câncer seria, assim, superestimada.

mudanças observadas em uma variável B são comparadas (PEARL; MACKENZIE, 2018, p. 132-133)[279].

O terceiro elemento que pode gerar quedas de tensão é a falta de compreensão das condições relevantes que criam o sucesso em escala e que devem permanecer para que a alta tensão continue (LIST, 2022, p. 74). Ao identificar um *nudge* potencialmente efetivo, o arquiteto de escolhas precisa identificar quais são os elementos que guiam o alto desempenho e que deverão ser mantidos para se obter escala.

List (2022, p. 75) aponta que esse desempenho pode ser motivado tanto pelas pessoas que implementam uma ideia quanto pelos ingredientes que a compõem. Se o sucesso por trás de um *nudge* é motivado pela personalidade do arquiteto de escolhas – que pode, por exemplo, ser um político popular –, haverá limitações para implementar esse *nudge* em grande escala. Afinal, a intervenção exigiria a presença desse político específico, o que pode limitar a sua abrangência. Se o sucesso for explicado pelos ingredientes que compõem o *nudge*, o arquiteto de escolhas deve entender quais ingredientes são negociáveis e quais não são negociáveis, por serem indispensáveis ao sucesso da política, além de verificar se estes últimos ingredientes são, de fato, escaláveis.

Nesse contexto, um primeiro ponto a que o arquiteto de escolhas deve se atentar é a manutenção da fidelidade nas fontes não negociáveis do seu sucesso. A princípio, é preciso que, em larga escala, a população afetada tenha aderência ao *nudge*; caso contrário, a política não conseguirá ter sucesso em modificar o comportamento (LIST, 2022, p. 77). Esse ponto será retomado na Subseção 5.2.2 abaixo. Ao mesmo tempo, deve-se evitar desvios do programa (*program drift*), quando as ideias propostas pelo arquiteto de escolhas não podem ser aplicadas da mesma forma em larga escala, seja por restrições organizacionais que não existiam em menor escala, seja pela impossibilidade de replicar um *nudge* em larga escala de maneira fidedigna (LIST, 2022, p. 80-81).

Os elementos não negociáveis por trás do sucesso de um *nudge* podem envolver tecnologias específicas. Nesse caso, deve-se verificar se as pessoas realmente utilizam os meios tecnológicos que serão utilizados para implementar um *nudge*, a fim de evitar quedas de tensão. List (2022, p. 84-86) alerta que a incompetência, a preguiça e o desperdício humanos são frequentemente um empecilho à obtenção de escala. Em outras palavras, como Thaler e Sunstein sempre alertaram, esperar o erro das pessoas é uma cautela que todo arquiteto de

[279] Em 2019, os economistas Abhijit Banerjee, Esther Duflo e Michael Kremer foram laureados com o prêmio Nobel em economia, em razão do seu trabalho com RCTs em prol da redução da pobreza. Cf, nesse sentido: (CALLAWAY, 2019).

escolhas deve ter. Ao elaborar um *nudge* ou incorporá-lo em alguma tecnologia específica, deve-se atentar à forma pela qual a população-alvo se comporta no momento específico em que ele será implementado.

Para superar esses desafios e garantir a aderência a um *nudge* específico, os arquitetos de escolha devem tornar essa aderência simples. Como List (2022, p. 87) sugere, para evitar quedas de tensão, o arquiteto de escolhas deve, desde o início, imaginar como a intervenção funcionaria em larga escala, se aplicada a toda a população, com suas situações variadas, por um longo período de tempo. O primeiro passo, nesse sentido, é compreender as variáveis não negociáveis da intervenção desde a sua concepção. Em grande escala, esses elementos devem estar disponíveis ou a ideia não será escalável. Se, por exemplo, a adoção de um determinado *nudge* exigir a presença constante de cientistas comportamentais, muitos municípios teriam dificuldade para adotá-lo; se essa presença não puder ser substituída por recursos tecnológicos, o *nudge* em questão dificilmente seria escalável.

O quarto desafio apontado por List (2022, p. 91) para que quedas de tensão sejam evitadas é a atenção aos efeitos de transbordamento (*spillovers effects*), isto é, impactos indesejados que um evento ou resultado pode ter em outro evento ou resultado. Um exemplo clássico de tais efeitos foi identificado por Sam Peltzman (1975), que identificou que a exigência de utilização de cintos de segurança em automóveis acabava aumentando o número de acidentes. Esse efeito – que ficou conhecido como efeito Peltzman – decorre de uma compensação de risco: como os motoristas se sentem mais seguros com medidas regulatórias que os protegem, eles acabam assumindo mais riscos, o que acaba causando mais acidentes (LIST, 2022, p. 90).

Efeitos de transbordamento ficam mais visíveis em grande escala. Os efeitos gerais de equilíbrio de mercado são, segundo List (2022, p. 92-94), um primeiro aspecto que favorece essa melhor visualização. Esses efeitos descrevem, basicamente, a relação autoajustável entre oferta e demanda, que é moldada por alterações no mercado. Quando o equilíbrio é alterado em alguma área, o sistema passa por ajustes automáticos até atingir um novo equilíbrio. Se, por exemplo, um governo utiliza *nudges* para fazer com que mais pessoas participem de programas de ensino técnico, como o Pronatec, acima do que a demanda por emprego é capaz de absorver, a política poderá gerar efeitos de transbordamento: a redução dos salários de tais empregados.

Além disso, o convívio social pode criar efeitos de transbordamento - positivos ou negativos - que podem inviabilizar a escalabilidade de um *nudge* (LIST, 2022, p. 101). Schultz e outros autores (2018), por exemplo, testaram *nudges* voltados à redução do consumo de energia, em que uma mensagem normativa nas contas de energia elétrica detalhava o uso de

energia médio da vizinhança. A política apresentou efeitos ambíguos: se o sujeito estivesse consumindo menos energia do que a média de sua vizinhança, o *nudge* provocava um aumento do consumo; caso contrário, a técnica conseguia produzir os efeitos esperados, *i.e.*, reduzir o consumo de eletricidade. Diante de tal efeito de transbordamento, Schultz e seus coautores introduziram um simples componente visual ao *nudge*: *emoticons* que evocavam emoções positivas ou negativas a depender do nível de consumo dos indivíduos foram introduzidos nas contas. O efeito de transbordamento, com isso, foi eliminado[280].

O arquiteto de escolhas também precisa estar atento a efeitos de transbordamento em rede (*network effect spillovers*), isto é, às situações em que a utilização de um produto ou a adoção de alguma política amplifica os benefícios ou os custos para toda a população envolvida (LIST, 2022, p. 106). Se, por exemplo, *sludges* dificultam a vacinação contra a COVID-19 entre as pessoas de uma determinada região, o comportamento das pessoas que deixarem de se vacinar ampliará os riscos à população como um todo. Analogamente, *nudges* que estimulem a vacinação causarão efeitos colaterais positivos em rede, já que contribuirão à imunidade de rebanho (*herd immunity*).

O último fator que pode conduzir a quedas de tensão são os custos à obtenção de escala. Como explica List (2022, p. 110), economias de escala existem quando o custo total de produção de uma unidade diminui quando o número de unidades a ser produzidas cresce. No caso de uma política baseada em *nudges*, haveria ganhos de escala se o custo para implementar essa política diminuísse ao passo que sua escala de implementação aumentasse. Em sentido contrário, se o custo médio de produzir algo aumenta quanto maior o crescimento de sua operação, há deseconomias de escala. Se, por exemplo, um elemento-chave para a implementação de um *nudge* é escasso ou de difícil obtenção - como o acompanhamento e o monitoramento por cientistas comportamentais renomados -, dificilmente ele conseguiria ser implementado em escala mais ampla.

A atenção aos custos envolvidos na aplicação de um *nudge* é necessária, já que os seus resultados devem justificar as despesas incorridas para implementá-lo. Caso contrário, com a ampliação da escala do *nudge*, haverá quedas de tensão e a ideia não será escalável (2022, p.

[280] Em interpretação dos seus resultados, Schultz e seus colegas (2018) afirmam que mensagens normativas apresentam efeitos mistos em contextos de campo. Embora campanhas baseadas em normas sociais sejam voltadas àqueles que se comportam de maneira indesejada, como elas são implementadas em grande escala, aqueles que apresentarem bom comportamento também acabarão recebendo as mensagens. Nesse caso, por mais que uma mensagem normativa descritiva possa ser útil para estimular um comportamento construtivo àqueles que agem de forma destrutiva, ela pode estimular um comportamento destrutivo àqueles que já agem de forma construtiva. Compreender os tamanhos dessas duas populações, assim, é essencial para garantir o sucesso de uma determinada política.

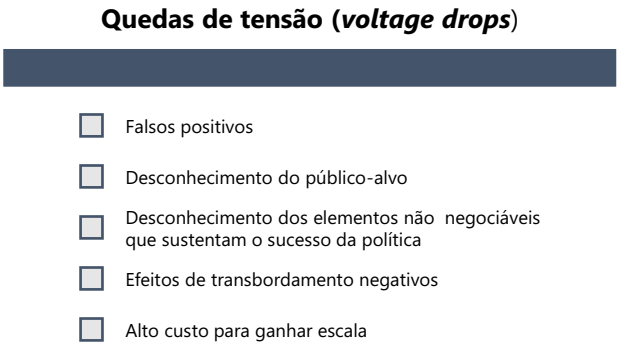
110). Nesse caso, é preciso entender de fato o custo-benefício da política proposta e verificar qual seria o grau de efetividade do *nudge* em implementar a mudança de comportamento ao longo do tempo, por unidade despendida.

Ainda que um *nudge* não tenha por objetivo auferir lucros - afinal, esta dissertação se dedica à avaliação de *nudges* utilizados enquanto técnica regulatória e não se debruça propriamente sobre *nudges* puramente privados -, é necessário ter alguma métrica para mensurar a utilidade gerada pela política (LIST, 2022, p. 121). Afinal, em um cenário de recursos limitados e inúmeros problemas urgentes que exigem soluções escaláveis, compreender o custo-benefício de uma determinada intervenção é essencial para definir se há justificativas para continuar investindo esses recursos nessa política específica ou se eles não deveriam ser redirecionados.

Para não ter surpresas com os custos necessários para escalar um determinado *nudge*, o arquiteto de escolhas - ainda de acordo com as propostas de List (2022, p. 124) - deve fazer um exercício de indução reversa (*backward induction*): retroativamente, o arquiteto de escolhas deve estimar qual seria o melhor resultado possível para o problema que se quer resolver e então ajustar a sua estratégia para seguir as etapas necessárias para atingir esse resultado. O arquiteto de escolhas pode constatar, por exemplo, que um *nudge* específico teria maior efetividade se fosse acompanhado por uma equipe de cientistas comportamentais de renome, com o uso de tecnologias avançadas e custosas. No entanto, como esse cenário não seria escalável, o arquiteto de escolhas trabalharia, desde o início, com uma solução que não fosse tão efetiva, mas que pudesse ser escalável. A mensagem de List (2022, p. 125-126) é clara: a perfeição é inimiga da escala.

Os cinco desafios que devem ser superados para que um *nudge* seja estabelecido em larga escala, sem quedas de tensão, estão enumerados na imagem abaixo:

Figura 17 - Fatores que levam a quedas de tensão



Fonte: Elaboração do Autor com base em List (2022).

Estabelecidas essas premissas e demonstrados os fatores que podem gerar quedas de tensão, é necessário, então, explorar exemplos concretos de *nudges* que foram adotados em larga escala com sucesso. A ideia não é esgotar esse assunto, mas trazer alguns estudos exemplificativos, além de ressaltar que caminhos de pesquisas que devem ser explorados pela literatura dedicada à efetividade dos *nudges*.

5.2.2. *Nudges* em longo prazo: avaliando a persistência dos efeitos de *nudges*

Uma primeira questão, essencial para qualquer intervenção em grande escala, envolve a persistência de seus efeitos no tempo.[281] Como discutido acima, um dos fatores que pode gerar quedas de tensão é a falta de aderência da população-alvo a uma determinada medida; sem aderência, não há escalabilidade. Nesse sentido, considerando a não coercibilidade dos *nudges*, um primeiro questionamento envolve a medida em que eles induzem mudanças permanentes de comportamento; isto é, sem o *nudge*, o sujeito continuaria a se comportar na direção estimulada pela política?

[281] Algumas das meta-análises mencionadas na Seção 5.1 trazem ponderações e dados a esse respeito. Beshears e Kosowski (2020, p. 16), por exemplo, embora tenham ressaltado a efetividade dos *nudges*, afirmaram que alguns *nudges* poderiam ter impacto em longo prazo menor do que o esperado, o que exigiria a exposição a novos *nudges*, a fim de se buscar uma mudança definitiva de comportamento. Por outro lado, em sua meta-análise dos efeitos de experimentos de campo envolvendo *nudges* para promover a alimentação saudável, Cadario e Chandon (2020, p. 34) verificaram que os tamanhos de efeito não eram afetados pela duração dos estudos – de 1 a 15 semanas. De todo modo, os autores ressaltaram que o tempo de duração dos estudos pode não evidenciar a distinção entre efeitos de curto ou longo prazo, já que, em algumas configurações, como restaurantes, mesmo que em um longo período de testes, os consumidores podem estar frequentando o estabelecimento pela primeira vez. O período máximo de duração dos estudos considerados por Cadario e Chandon, ainda, não é suficiente para indicar se os *nudges*, de fato, induziram mudanças de comportamento em longo prazo. Afinal, 15 semanas ou pouco mais de 4 meses não é um período longo o suficiente para comprovar que os efeitos são persistentes.

Inicialmente, um exemplo descrito por Thaler e Sunstein (2021) e por Henrik Cronqvist, Thaler e Frank Yu (2018) parece indicar que ao menos alguns *nudges* têm o potencial de gerar efeitos de longo prazo. A política de previdência da Suécia, implementada em 2000, envolve alterações da arquitetura de escolhas, de modo que os resultados dessa política funcionam quase como um experimento natural a respeito da persistência dos efeitos de *nudges*. A contribuição à previdência na Suécia é obrigatória e definida em 16% da renda (THALER; SUNSTEIN, 2021, p. 199). A reforma de 2000 pretendia dissociar parte desse tributo para criar planos de contribuição definida no que ficou conhecido como Plano Sueco de Previdência Premium (*Swedish Premium Pension Plan*).

Esse plano apresentava as seguintes características: a) um fundo específico era definido como o fundo padrão (*default*) e seria atribuído àqueles que não tivessem tomado qualquer decisão; b) campanhas massivas de publicidade incentivavam as pessoas a escolherem ativamente seus próprios portfólios; c) havia liberdade para adesão de qualquer fundo que cumprisse determinados requisitos fiduciários, de modo que o número de fundos disponíveis era bastante alto - 456; d) todos os participantes recebiam informações sobre os fundos em um folheto; e e) à exceção do fundo padrão, todos os demais poderiam promover publicidade para auferir lucro.

Como Thaler e Sunstein (2021, p. 203), assim como Cronqvist, Thaler e Yu (2018, p. 4-5), ponderam, nessa proposta, havia uma batalha de *nudges*: enquanto a regra *default* explorava o viés do *status quo*, a inércia e a procrastinação para promover o fundo padrão, campanhas publicitárias - do governo e dos próprios fundos - estimulavam a escolha ativa. No fim das contas, no cenário inicial, apenas um terço da população de participantes optou pela opção padrão; ou seja, dois terços da população escolheu seus planos de maneira ativa.

De todo modo, a evolução da estratégia sueca para a previdência permite que se reflita a respeito da persistência dos efeitos de *nudges*. Depois da implementação da reforma previdenciária, as campanhas publicitárias – públicas e privadas – em prol da escolha ativa perderam sua força. Em 2003, de todos os novos aderentes à previdência sueca, apenas 1 de cada 10 (9,4%) optava por escolher ativamente o seu portfólio; em 2010, esse percentual havia sido reduzido para apenas 3%, e, em anos recentes, ele é inferior a 1% (CRONQVIST; THALER; YU, 2018, p. 5; THALER; SUNSTEIN, 2021, p. 210). Em outras palavras, com a redução da intensidade da publicidade – *nudge* que promovia escolhas ativas –, o efeito da regra *default* foi potencializado e mais pessoas passaram a utilizá-la.

Além disso, ao longo do tempo, as decisões iniciais dos agentes - adoção da regra *default* ou opção escolhida a partir da promoção de escolhas ativas - perduraram por um longo tempo,

isto é, os *nudges* parecem ter tido efeitos bastante persistentes. Com relação àquelas pessoas que seguiram a regra *default*, entre 2000 e 2016, apenas 27,4% mudaram de opinião e passaram a ser decisores ativos. Por outro lado, as pessoas que optaram, inicialmente, por decidir ativamente mantiveram-se ainda mais vinculadas à sua decisão inicial: apenas 2,9% desses agentes optaram pela aderência à opção padrão (CRONQVIST; THALER; YU, 2018, p. 6; THALER; SUNSTEIN, 2021, p. 211).

Mesmo quando aconteceram mudanças substanciais na realidade dos dois grupos de agentes analisados acima, mudanças com relação à opção inicial permaneceram sendo raras. Com relação à opção padrão, ainda que, com o tempo, o seu portfólio tenha sido alavancado e as suas opções de investimento tenham se tornado bastante arriscadas, poucas pessoas que a haviam escolhido optaram por mudar para escolher ativamente o seu portfólio (CRONQVIST; THALER; YU, 2018, p. 8-9; THALER; SUNSTEIN, 2021, p. 213). Já quanto às escolhas ativas, um dos fundos que estava disponível no sistema – coordenado pela empresa Allra - começou a atrair suspeitas de comportamento fraudulento e, diante dessas acusações, apenas 1,4% dos investidores venderam suas participações no fundo; mesmo após a Deloitte - então auditora da Allra - ter renunciado ao seu posto e denunciado a Allra às autoridades, apenas 16,5% dos investidores escolheram por direcionar seus recursos a outro fundo (CRONQVIST; THALER; YU, 2018, p. 10; THALER; SUNSTEIN, 2021, p. 214).

Nem sempre, no entanto, *nudges* produzirão efeitos persistentes. Christina Gravert e Linus Olsson Collentine (2021), por exemplo, testaram a efetividade em longo prazo - período de oito meses - de *nudges* baseados em mensagens que descreviam normas sociais para aumentar a utilização do transporte público na Suécia. A estratégia foi experimentada como parte de um experimento natural de campo com mais de 14.000 pessoas, no condado da Escânia, na Suécia. O experimento fazia parte de uma campanha de *marketing* em prol da utilização do transporte público, em que, ao registrar um novo endereço, as pessoas recebiam uma oferta para testar o transporte público de maneira gratuita por duas semanas.

Nesse contexto, Gravert e Collentine testaram duas estratégias alternativas para ampliar o número de usuários do transporte público. A primeira estratégia - o *nudge* mencionado acima - enfatizava, juntamente com a oferta de um cartão que viabilizava a utilização gratuita do transporte por duas semanas, que 72% dos vizinhos do usuário utilizavam o transporte público ocasionalmente. A segunda estratégia era baseada no aumento dos incentivos monetários para o uso do transporte público: o período de gratuidade era estendido por quatro semanas.

Para mensurar os efeitos de ambas as estratégias, Gravert e Collentine acompanharam as respostas à oferta de gratuidade, assim como a ativação do cartão e o seu uso ao longo de

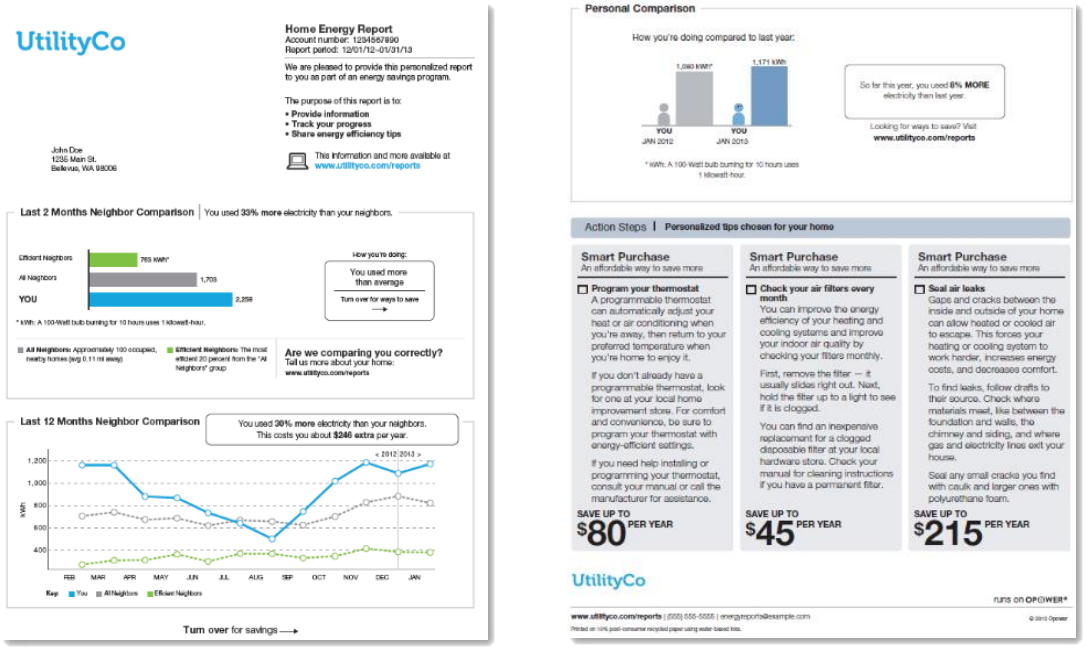
oito meses. Os *nudges* testados, no entanto, não conseguiram aumentar as respostas do público à oferta de gratuidade. O aumento de incentivos, em contrapartida, ampliou as respostas iniciais em 16%, o que parecia indicar que os indivíduos estavam interessados em utilizar o transporte público, se tivessem os incentivos adequados. Em longo prazo, ainda, o grupo de indivíduos que recebeu a gratuidade por quatro meses apresentava maior probabilidade - na ordem de 33% - de permanecer na base de cartões ativos da agência de transportes locais.

Em interpretação dos seus resultados, Gravert e Collentine (2021, p. 3) ponderam que *nudges* devem ser testados em seus respectivos domínios, antes de serem implementados. Enquanto normas sociais já se provaram vantajosas na conservação de energia e água, assim como no consumo de alimentos saudáveis, no domínio dos transportes essa estratégia parece não funcionar. Nesse domínio - discutem Gravert e Collentine - pode ser que os esforços e os custos necessários para a mudança de comportamento sejam mais relevantes para o agente do que a alteração do consumo de um alimento específico ou a redução do tempo no banho. Por essa razão, incentivos parecem funcionar melhor.

Hunt Allcott e Todd Rogers (2014), por outro lado, avaliaram a persistência dos efeitos do "relatório de energia doméstica" (*home energy report*) produzido pela empresa Opower – adquirida pela Oracle por mais de 500 milhões de dólares (BRANDON *et al*., 2022, p. 1) – a partir de uma combinação de *feedback* personalizado da utilização de energia elétrica, comparações sociais e informações de conservação de energia. As comparações sociais, em especial, enquadram-se na definição de *nudge* adotada nesta dissertação. Embora a intervenção tenha sido analisada como um todo - isto é, os efeitos das comparações sociais não foram isolados –, a pesquisa de Allcott e Rogers apresenta dados relevantes a respeito dos efeitos em longo prazo de intervenções informadas por *insights* comportamentais.

No momento em que a pesquisa foi conduzida, 6,2 milhões de casas recebiam os relatórios de energia doméstica em 85 locais diferentes nos Estados Unidos. Nesse contexto, Allcott e Rogers tentaram responder duas perguntas: a) qual seria a persistência dos efeitos dessa intervenção após ela deixar de ser implementada?; e b) se os efeitos pós-intervenção não forem persistentes, a continuidade do tratamento – isto é, a repetição da intervenção – seria necessária à obtenção de efeitos continuados? Para respondê-las, foram avaliados três programas da Opower, implementados como estudos randomizados controlados em 234.000 casas, entre 2008 e 2013.

Figura 18 - Exemplo relatório de energia doméstica (frente e verso)



Fonte: Allcott e Rogers (2014, p. 3008-3009).

A princípio, Allcott e Rogers (2014, p. 3004-3005) identificaram um efeito de "ação e retrocesso" (*action and backsliding*): após o recebimento dos relatórios, os consumidores reduziam o consumo de energia elétrica, mas esses esforços imediatos reduziam-se gradativamente, de modo que os efeitos tendiam a desaparecer em poucos meses, a não ser que a intervenção fosse reiterada. Esse padrão, no entanto, não era constante: depois dos primeiros quatro relatórios, a redução imediata do consumo de energia era cinco vezes menor do que a redução inicial. Em outros termos, com a repetição da intervenção, o padrão de ação e retrocesso é atenuado. De fato, para os grupos em que os relatórios só foram descontinuados depois de dois anos, a intervenção perdia os efeitos a uma taxa de aproximadamente 10 a 20% ao ano – valor de quatro a oito vezes inferior à taxa de decaimento de intervenções que não foram reiteradas. De qualquer forma, Allcott e Rogers (2014, p. 3005) verificaram que os consumidores ainda não estavam completamente habituados ao novo comportamento, ainda que tivessem permanecido recebendo relatórios de energia doméstica por dois anos.

Ao testar a intervenção ao longo de um período de quatro anos, os efeitos do tratamento eram muito mais robustos - 50 a 60% mais fortes - se a intervenção não fosse interrompida. Essa diferença, na interpretação de Allcott e Rogers (2014, p. 3005), parece indicar que, com a repetição da intervenção, as pessoas desenvolvem gradualmente um novo "estoque de capital"

- isto é, desenvolvem novos hábitos ou adquirem aparelhos mais eficientes em termos energéticos -, o que gera alterações persistentes de resultado. Em outras palavras, a repetição da intervenção por um período mais longo gera efeitos de transbordamento.

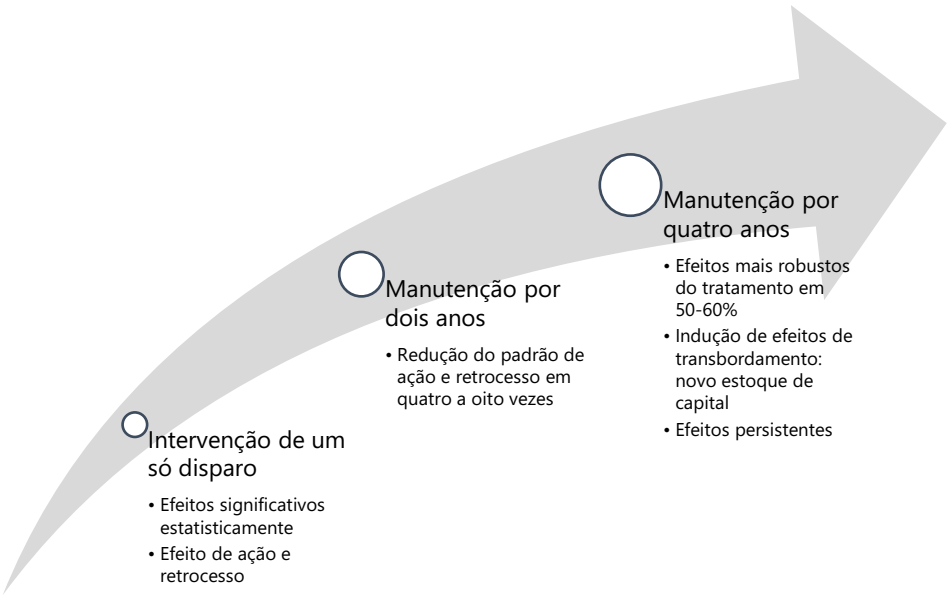
Allcott e Rogers (2014, p. 3025-3028) estimaram, ainda, a correlação entre esses resultados de longo prazo e a participação em programas de eficiência energética, que promovem grandes alterações de capital físico, a partir de incentivos à aquisição de eletrodomésticos eficientes, sistemas de aquecimento, ventilação e ar-condicionado e utilidades para aprimoramento energético das residências. De acordo com suas estimativas, esses aprimoramentos seriam responsáveis por explicar não mais do que 2% dos efeitos de longo prazo dos relatórios de energia doméstica. Esse resultado sugere que uma combinação de alteração nos hábitos de consumo e mudanças menores e não observadas no estoque de capital físico seriam os meios que conduziriam à alteração de comportamento em longo prazo.[282]

A partir desses dados, Allcott e Rogers (2014, p. 3006) entendem que os efeitos de políticas como os relatórios de energia doméstica são subestimados. Entende-se, geralmente, que a alteração de comportamento persistente, no que diz respeito à conservação de energia, depende de alterações de "estoque de capital", como a aquisição de equipamentos que consomem pouca energia. Em outras palavras, admite-se, como premissa, que técnicas como os relatórios não teriam efeitos persistentes. Por isso, essas técnicas não costumam ser aplicadas de maneira reiterada – aplica-se como intervenções de um só disparo (*one-shot intervention*). Além de demonstrar a falsidade dessa premissa, Allcott e Rogers indicaram que ela faz com que os gestores subestimem as economias de energia que seriam provocadas pela política por um fator superior a dois – enquanto as previsões no cenário sem persistência, girariam em torno de 41 a 63 dólares por residência, Allcott e Rogers observaram reduções equivalentes a 100 a 149 dólares, com a manutenção dos relatórios por um longo prazo.

[282] Nesse sentido, Erin Frey e Rogers (2014, p. 174), também analisando os efeitos persistentes de intervenções conduzidas pela Opower, identificaram quatro caminhos para que uma intervenção produza efeitos persistentes: a) a construção de hábitos psicológicos, em que a intervenção produz uma tendência automática à repetição de uma determinada resposta comportamental, desencadeada pelo contexto estável em que o comportamento é realizado; b) mudanças em como e no que as pessoas pensam, quando a intervenção muda um elemento da forma de pensar dos sujeitos, como crenças, identidades e interpretações, que está causalmente ligado ao comportamento que se busca estimular; c) alterações dos custos futuros, quando a intervenção induz o sujeito a se comportar de modo a alterar o custo do comportamento estimulado no futuro - por exemplo, a partir da aquisição de eletrodomésticos e eletrônicos eficientes em termos energéticos, o que facilita a conservação de energia; e d) reforço externo, quando a intervenção induz as pessoas a se comportar de forma que as expõe a processos externos, incluindo pressões sociais, a que elas não estariam expostas sem a intervenção - por exemplo, se o sujeito é incluído em listas de *marketing* de equipamentos com eficiência energética, essa intervenção pode fazer com que ele pense com mais frequência em adquirir esses equipamentos.

Mais do que isso, o entendimento da persistência e da medida em que intervenções induzem mudanças de hábito permite que se otimize o momento e a intensidade da utilização de determinada técnica. Nesse sentido, Allcott e Rogers ponderam que, no caso da conservação de energia, o arquiteto de escolhas poderia elaborar políticas mais eficientes se: (i) repetisse uma intervenção como os relatórios de energia doméstica; e (ii) reduzisse a sua frequência, conforme os participantes desenvolvessem um novo "estoque de capital" em termos de hábito ou de tecnologias.

Figura 19 - A persistência de efeitos do relatório de energia doméstica



Fonte: Elaborado pelo Autor com base em Allcott e Rogers (2014).

Em linha com esse experimento, Alec Brandon e outros autores (2022) conduziram uma revisão de 38 experimentos naturais de campo envolvendo o relatório de energia doméstica proposto pela Opower, para decompor a parcela dos seus efeitos de longo prazo que seria atribuível à adoção de tecnologia e a parcela atribuível à formação de hábitos. Para isso, os autores contaram com um elemento específico dessa intervenção: a sua descontinuidade quando um sujeito submetido a ela decidia se mudar. Nesse caso, a observação do comportamento de consumo daquela residência após a continuidade permitiria a identificação do tamanho de efeito atribuível à adoção de tecnologia - a partir da premissa de que melhorias energéticas permaneceriam na casa após a mudança do seu morador inicial. A parcela restante dos efeitos de longo prazo do relatório de energia doméstica seria, então, atribuída à mudança de hábitos.

A partir da análise de 140.000 pessoas que haviam se mudado em 38 programas diferentes em que o relatório de energia doméstica foi implementado, Brandon e seus coautores identificaram, inicialmente, que, em longo prazo, essas pessoas reduziam o seu consumo de energia em 2,1% após receber o referido relatório. Essa redução equivale a reduções de 0,6 a 0,9 kWh por dia, o que, na prática, seria equivalente à utilização de 1 lâmpada incandescente de 60 watts por 10 a 15 horas por dia ou à substituição de 2 a 4 lâmpadas incandescentes de 60 watts que são usadas por 5 horas ao dia por lâmpadas fluorescentes equivalentes (BRANDON *et al*., 2022, p. 21). Em caso de mudanças, 51,4% dessa redução - isto é, da efetividade de longo prazo da medida - permanecia nas casas originais, enquanto 48,6% seriam atribuíveis à mudança de hábitos.

Com base nesses dados, Brandon e outros autores (2022, p. 30-31) apresentam três contribuições principais a respeito dos efeitos de *nudges* em longo prazo.

A primeira delas é fornecer uma potencial explicação para as divergências na persistência dos efeitos de *nudges* em diferentes domínios: a disponibilidade de tecnologias produtivas que facilitam a manutenção do comportamento estimulado. Em domínios como a conservação de água e energia, essas tecnologias estão disponíveis, e essa seria a razão, segundo Brandon e outros autores, para o maior sucesso de *nudges* em tais contextos. Em outros domínios, como no estímulo à conformidade com as regras, doações para a caridade, poupança, evasão fiscal e eleições, essas tecnologias não estão facilmente disponíveis, o que pode dificultar a persistência dos efeitos de *nudges*.

Além disso, essa pesquisa indica que os arquitetos de escolhas podem tentar replicar o sucesso em longo prazo dos relatórios de energia doméstica de duas maneiras distintas. A primeira é o foco em comportamentos que podem ser influenciados por tecnologias já disponíveis. A depender da tecnologia, esse aspecto pode permitir, ainda, a personalização da intervenção, o que pode fazer com que grupos de pessoas distintos sejam direcionados pelo *nudge* – o que evitaria o segundo fator que List aponta como indutor de quedas de tensão, como discutido na Subeção 5.2.1 –, embora esse efeito não seja objeto do estudo de Brandon e seus colegas. Em outra perspectiva, o arquiteto de escolhas pode combinar *nudges* que exploram normas sociais, com oportunidades para adotar novas tecnologias. Se, por exemplo, a comparação com os pares é adotada em conjunto com regras *default* que, por exemplo, tornam o ato de votar menos custoso (*e.g.*, voto por correio) ou facilitem a poupança para o futuro, os efeitos de longo prazo podem ser mais robustos (BRANDON *et al*., 2022, p. 30).

O terceiro ponto sugerido por Brandon e seus coautores é a importância de incluir os custos da tecnologia na análise do custo-benefício de *nudges*. Se o efeito de longo prazo é

motivado, primordialmente, pela implementação de novas tecnologias, se elas forem muito custosas, o custo-benefício da política será reduzido, ainda que a implementação do *nudge* seja barata. Esses custos indiretos, segundo Brandon e seus colegas, deveriam ser considerados nas análises, para que sejam obtidas avaliações mais robustas. Afinal, como List aponta, é importante conhecer os ingredientes não negociáveis do sucesso de uma política; se o custo de os implementar for inviável em larga escala, talvez a política enfrente quedas de tensão.

Apesar de Brandon e seus coautores focarem bastante nos efeitos da mudança de tecnologia na persistência dos efeitos de *nudges*, a parcela de efeitos atribuída à mudança de hábitos – 48,6 % – não parece ser irrelevante. Nesse caso, ainda que não existam tecnologias disponíveis para reduzir o custo da aderência ao *nudge*, a promoção da mudança de hábitos em larga escala deve ser estimulada pelo arquiteto de escolhas. Afinal, se ela for bem-sucedida, a política poderá ter sucesso, ainda que não conte com tecnologias mais eficientes.[283]

Os dados expostos acima mostram que *nudges* - a partir da mudança de hábito, da adoção de novas tecnologia ou por outros motivos, como o próprio viés do *status quo*, no caso de regras *default* - podem produzir mudanças de comportamento em longo prazo. Seus efeitos podem, de fato, ser persistentes, como os exemplos da previdência na Suécia e dos relatórios de energia doméstica promovidos pela Opower demonstram. Parece mais provável, inclusive, que *nudges* que alteram de maneira mais brusca aspectos da estrutura da escolha, em vez de apenas fornecer mensagens - distinção proposta por Congiu e Moscati (2020), conforme demonstrado no Capítulo 3 - podem ter efeitos mais persistentes, como indica o caso da Suécia.

Em alguns cenários, no entanto, como no caso dos transportes públicos, *nudges* podem ter efeitos limitados em curto e em longo prazo. Esses resultados indicam que a experimentação é, ainda, essencial para esclarecer os verdadeiros efeitos de *nudges* em longo prazo. Embora os resultados expostos nesta Seção tenham indicado que esses efeitos podem ser persistentes, não se pretendeu fazer uma revisão exaustiva. Boa parte das obras mencionadas é bastante recente e compreender a efetividade dos *nudges* em longo prazo é, no momento, uma das agendas de

[283] A mudança de hábitos duradoura, no entanto, nem sempre é factível. Beshears e outros autores (2021), nesse sentido, elaboraram um experimento para verificar se incentivos a rotinas de exercícios físicos – pagamentos a cada vez que o sujeito fosse à academia e cumprisse uma rotina específica planejada previamente – seriam capazes de promover hábitos de exercício mais persistentes em comparação a incentivos flexíveis – pagamentos a cada vez que o sujeito fosse à academia, independentemente do tempo que lá permanecesse. De acordo com seus resultados, os sujeitos que recebiam incentivos a adotar uma rotina específica apresentavam menor volume de atividade física do que aqueles que receberam incentivos flexíveis. Mesmo no cenário em que os incentivos eram maiores para a rotina de exercícios do que para os incentivos flexíveis – sete dólares *versus* três dólares, respectivamente –, por mais que a frequência de visitas à academia fosse similar, os efeitos dos incentivos à rotina decaíram consideravelmente com o término da intervenção, isto é, com o fim do incentivo à rotina, as pessoas submetidas a essa intervenção apresentavam um volume muito menor de exercícios físicos. Nesse caso, a criação de um programa de incentivos que promova rotinas rígidas pode ser contraproducente à formação de hábitos.

pesquisa mais atuais e importantes, como expressamente reconhecido pela literatura (BESHEARS; KOSOWSKY, 2020, p. 14-15; SUNSTEIN, 2019a, p. 113; THALER; SUNSTEIN, 2021, p. 216).

5.2.3. *Nudges* em larga escala: efeitos significativos, porém menores do que o esperado

Expostas a) as armadilhas que devem ser evitadas para evitar quedas de tensão e b) algumas questões preliminares a respeito da persistência dos efeitos dos *nudges* e dos aspectos que a compõe, é necessário, então, avaliar se *nudges* conseguem ter efeitos relevantes em larga escala. A persistência dos efeitos dos *nudges* do Plano Sueco de Previdência Premium e dos relatórios de energia doméstica da Opower parecem indicar que sim: embora os efeitos sejam modestos, *nudges* parecem ter sucesso, ainda que implementados em larga escala. Em outras palavras, em alguma medida, ainda que para grandes populações, *nudges* parecem ser efetivos e ter bom custo-benefício, o que confirmaria os dados das meta-análises citadas na Seção 5.1.

De todo modo, algumas considerações extraídas de outras iniciativas em larga escala podem ser mais esclarecedoras a respeito dos efeitos de *nudges* em populações maiores. Entender esses efeitos é necessário não só para avaliar o custo-benefício dos *nudges*, como também para indicar os caminhos que devem ser percorridos para que os arquitetos de escolha possam maximizar os efeitos dos *nudges* em larga escala.

Abhijit Banerjee e outros autores (2021a, p. 1) afirmam que, por mais que meta-análises sejam a maneira mais comum de comparar políticas, elas podem indicar distorções por comparar populações e intervenções variadas. Nesse caso, é difícil concluir que as diferenças são puramente o resultado de uma intervenção, em vez de distinções entre a população que participou do experimento; ao mesmo tempo, se intervenções diferentes são avaliadas em diferentes contextos, identificar interações entre elas é impossível. Por essa razão, Abhijit Banerjee[284] e seus colegas sugerem que, antes de implementar um programa, ele seja testado em um experimento de larga escala, no contexto em que ele será implementado e com a comparação de várias estratégias diversas.

A crítica posta se aplica a algumas das meta-análises discutidas na Seção 5.1. Mertens e outros autores (2022), por exemplo, compararam experimentos conduzidos em laboratórios

[284] Opta-se por manter a menção ao nome e ao sobrenome de Abhijit Banerjee apenas para que não haja qualquer tipo de confusão entre os seus trabalhos e os trabalhos de Sanchayan Banerjee.

com experimentos de campo, nos Estados Unidos e em outros países, com populações muito diversas e em vários contextos. Do mesmo modo, Hummel e Maedche (2019) avaliaram experimentos *online*, em laboratório, de campo, *surveys* e outros, que reuniam ferramentas diversas de alteração da arquitetura de escolhas, em contextos de aplicação muito variados. Nesse caso, por mais robustas que sejam as descobertas, pode ser que elas não sejam generalizáveis para todo tipo de população ou contexto de aplicação.

Por outro lado, Luo e seus coautores (2021), em sua meta-análise dos efeitos de *nudges* e de *sludges*, decidiram limitar o escopo de seu trabalho a experimentos de campo - estudos randomizados controlados - com efetiva mensuração de alteração de comportamento, isto é, sem incluir resultados autorreferidos. Cadario e Chandon (2020), do mesmo modo, só incluíram experimentos de campo, em sua meta-análise de *nudges* voltados à promoção da alimentação saudável. Há, assim, maior probabilidade de indicação de relações de causalidade entre os *nudges* analisados e as mudanças de comportamento que foram estimuladas.

De todo modo, além da crítica às meta-análises, Abhijit Banerjee e seus coautores (2021a) avaliaram qual a combinação de *nudges* que produziria o melhor e o mais eficiente resultado para ampliar as taxas de vacinação de crianças na região de Haryana, na Índia. O experimento envolveu 2.360 vilarejos nessa região, que tem população superior a 8 milhões de pessoas, sendo 295.038 crianças. Embora a região tivesse uma estrutura desenvolvida para a aplicação de vacinas, dados de 2015 e 2016 indicavam que apenas 62% das crianças estavam completamente imunizadas. Nos distritos selecionados à implementação do estudo, esse índice era ainda menor e era inferior a 40%, segundo relatos dos próprios pais.

Abhijit Banerjee e seus coautores avaliariam combinações diferentes de três tipos de *nudges*, com diferentes intensidades, em prol da imunização. Como o objetivo era encontrar a melhor política ou a melhor combinação de políticas antes de implementá-la em grande escala, em um estudo randomizado e cruzado, Abhijit Banerjee e seus coautores identificaram 57 políticas em potencial, a partir de combinações das estratégias descritas abaixo.

O primeiro conjunto de estratégias envolvia alterações de incentivos, a partir do fornecimento de créditos para telefone celular. Quatro tipos de alterações foram testadas: a) altos incentivos constantes - 90 rúpias indianas por imunização, o que levava a 450 rúpias indianas pela imunização completa; b) altos incentivos progressivos - 50 rúpias indianas para as três primeiras imunizações, 100 para a quarta e 200 para a quinta; c) baixos incentivos constantes - 50 rúpias indianas por imunização; e d) baixos incentivos progressivos - 10 rúpias indianas para cada uma das três primeiras imunizações, 60 para a quarta e 160 para a quinta. Embora Abhijit Banerjee e seus coautores denominem essa intervenção um *nudge*, ela não

parece se enquadrar nessa definição. Afinal, a alteração de incentivos era substancial considerando as condições da população envolvida. De fato, 90 rúpias indianas eram equivalentes a um quilograma de lentilhas na região e, mesmo os incentivos mais baixos – na ordem de 50 rúpias indianas – eram relevantes: equivaliam, em média, a 100 minutos de conversas por telefone (BANERJEE *et al*., 2021a, p. 8).

A outra estratégia - esta já efetivamente enquadrada na definição de *nudge* adotada nesta dissertação - era a utilização de lembretes, por mensagens de texto enviadas aos pais. A mensagem, em geral, enfatizava que a época de vacinar os filhos contra determinada doença havia chegado. No conjunto de participantes que também recebia incentivos, os lembretes também enfatizavam que uma determinada quantia em créditos para celulares seria recebida como recompensa pela imunização da criança. As mensagens poderiam ser enviadas a 33% (baixa saturação) ou a 66% (alta saturação) da comunidade envolvida.

Por fim, a terceira estratégia envolvia a utilização de um "embaixador" para a imunização, com o objetivo de alavancar redes sociais para a proliferação da informação. A intenção parece ser provocar efeitos de transbordamento em rede, o que pode favorecer a escalabilidade de políticas públicas, como indicado na Subseção 5.2.1. Em linhas gerais, Abhijit Banerjee e seus coautores identificaram indivíduos influentes que pudessem difundir entre os moradores dos vilarejos a existência dos campos de imunização, e, quando relevante, os incentivos disponíveis aos pais que vacinassem seus filhos. Foram definidos embaixadores de quatro formas diferentes: a) aleatória; b) pessoas consideradas boas em transmitirem informações (núcleos de informação – *information hubs*); c) pessoas consideradas confiáveis no fornecimento de bons conselhos sobre questões de saúde e de agricultura (confiáveis – *trusted*); e d) pessoas consideradas confiáveis e boas em retransmitirem informações (núcleos de informação confiáveis – *trusted information hub*) (BANERJEE *et al*., 2021a, p. 10).

Duas combinações apresentaram resultados significativamente melhores do que o grupo de controle. A primeira foi a combinação de incentivos progressivos (de pequena ou alta intensidade) com embaixadores considerados núcleos de informação (*information hubs*) e lembretes via SMS (em ambas as saturações). Essa combinação gerou um aumento de 55% nas imunizações, em comparação ao grupo de controle. A segunda foi a combinação de embaixadores confiáveis com altos incentivos progressivos e lembretes, que provocou um

aumento nas imunizações na ordem de 44% em relação ao grupo de controle *(BANERJEE et al.*, 2021a, p. 28)[285].

Abhijit Banerjee e seus coautores (2021a, p. 28) ponderam, no entanto, que, em comparação às políticas tradicionais, as combinações de estratégias que produziram melhor resultado aumentam o número de imunizações a um custo consideravelmente maior por imunização. Enquanto as medidas tradicionais induzem 43,6 imunizações mensais por vilarejo a cada 1.000 dólares gastos, as combinações propostas acima induzem 36 imunizações mensais por vilarejo a esse mesmo custo. Essa diferença reflete o custo dos incentivos: embora eles contribuam para ampliar o número de pessoas vacinadas, os custos das intervenções aumentam consideravelmente. Ao mesmo tempo, embora mais eficiente, a estratégia do grupo de controle induziu um número consideravelmente menor de vacinações; se a questão é ampliá-lo o governo pode optar por uma política menos eficiente, mas que realmente seja efetiva, considerando, ainda, que a vacinação gera efeitos de transbordamento positivos, como discutido na Subseção 5.2.1[286].

De qualquer forma, a depender do orçamento do governo que vai aplicar uma intervenção como essa, o aumento dos custos por imunização pode apontar para uma possível queda de tensão: sem recursos suficientes para escalar a proposta para todo o país, por exemplo, o governo pode ser obrigado a reduzir ou eliminar os incentivos a fim de implementar a proposta. No limite, o aumento de custos pode, eventualmente, não ser justificado pelos resultados obtidos. Eventuais reduções dos incentivos, ainda, podem eliminar a motivação dos pais a levarem seus filhos para vacinação[287].

Em outras palavras, se o governo não tiver recursos suficientes para arcar com essa política em larga escala, pode ser necessário pensar em uma solução que não seja perfeita, mas que seja escalável, para evitar quedas de tensão, como List sugere. Nesse caso, se o arquiteto de escolhas tem restrições orçamentárias à imunização, o caminho adequado pode ser buscar a política mais eficiente – isto é, que produz o maior volume de vacinações por dólar gasto. Abhijit Banerjee e seus coautores identificaram que a estratégia mais eficiente seria combinar

---

[285] Além disso, a combinação de altos incentivos progressivos com lembretes via SMS teve um efeito positivo, mas ruidoso. Por outro lado, a combinação de núcleos de informação confiáveis com incentivos progressivos e lembretes não produziu resultados significativos. Em todas as combinações, ainda, incentivos constantes não apresentaram efeitos positivos (BANERJEE *et al.*, 2021a, p. 28).

[286] Nesse sentido, Abhijit Banerjee e seus coautores (BANERJEE *et al.*, 2021a, p. 28) mencionam que a imunização traz tantos benefícios que um aumento dos seus custos pode, ainda assim, ser mais eficiente do que outras despesas para lutar contra doenças infantis.

[287] Como será mais bem discutido na Seção 5.4 abaixo, incentivos monetários podem substituir normas sociais e valores morais que motivam determinadas ações, e serem, dessa forma, contraproducentes. Cf., nesse sentido: (BOWLES, 2016).

embaixadores considerados núcleos de informação (confiáveis ou não), com lembretes via SMS para 33% da comunidade, sem incentivos. Essa estratégia induziu um aumento de 9,1% nas vacinações a cada dólar gasto.

Outro experimento que testou os efeitos de *nudges* em larga escala foi desenvolvido por Kelli A. Bird e outros autores (2021). O objetivo foi verificar se, em larga escala, *nudges* com efeitos positivos em menor escala poderiam ter impacto em campanhas nacionais e estaduais para encorajar estudantes a se candidatarem ao processo de solicitação de auxílio estudantil. De maneira geral, a amostragem mediana de estudantes nos estudos avaliados por Bird e seus coautores era de 6.233 participantes; essas intervenções seriam promissoras pela possibilidade de implementá-las em larga escala[288]. Nesse contexto, Bird e seus coautores fizeram dois estudos controlados randomizados, com amostragem de 800 mil estudantes, em parceria com uma organização nacional e com uma organização estadual.

A primeira intervenção testada foi direcionada a todos os estudantes de baixa renda no último ano do Ensino Médio (*high school*), que estivessem registrados em um portal federal unificado que permite a realização de inscrições para a universidade. Esses estudantes receberiam mensagens que os encorajariam a completar o formulário de solicitação do auxílio estudantil, a fim de maximizar o auxílio que poderia ser recebido. Algumas variações nessas mensagens foram testadas, como mudanças de enquadramento (*framing*); diferentes canais de envio das mensagens (carta, e-mail ou mensagem de texto); acesso a consultoria individual; e um *nudge* social para encorajar os pares a também completar o formulário.

Já a segunda intervenção teve como foco todos os estudantes de um Estado de grandes dimensões, que estivessem inscritos no portal unificado estadual, destinado às inscrições às faculdades estaduais públicas, assim como a algumas instituições privadas e faculdades comunitárias. Desta vez, os estudantes recebiam mensagens informando duas importantes alterações na política de financiamento estudantil: a) a possibilidade de apresentar o formulário de pedido de auxílio em outubro do ano anterior à inscrição na faculdade; e b) a possibilidade de utilizar declarações de imposto de renda que suas famílias já tivessem apresentado, durante o processo de solicitação de auxílio estudantil. As mensagens variavam em múltiplas dimensões, como o momento em que eram implementadas e o uso de infográficos para aumentar a saliência visual.

Em ambos os casos, as mensagens não tiveram qualquer efeito na inscrição de estudantes ou nos resultados em termos de auxílio financeiro. A falta de efeitos permaneceu

[288] Convém mencionar que alguns desses estudos foram avaliados na revisão sistemática de Benartzi e outros autores (2017), que avaliou o custo-benefício de *nudges*, em comparação com outras estratégias.

consistente em diferentes amostragens, conteúdo da mensagem, momento do seu envio, apresentação visual e ofertas de ajuda personalizada. Ao especular as razões do seu resultado, enquanto a literatura majoritária, formada por experimentos em pequena escala, apresenta efeitos significativos com estratégias semelhantes, Bird e seus coautores (2021, p. 5) mencionam que os estudos pregressos utilizavam parceiros locais, com conexões mais próximas e conhecimento dos estudantes envolvidos. A mensagem, assim, seria mais facilmente assimilada, em linha, inclusive, com a utilização de embaixadores propagadores de informação no experimento de Abhijit Banerjee e outros autores. Além disso, como o experimento foi aplicado em grande escala, as mensagens eram mais genéricas e menos personalizadas do que em estudos anteriores, o que reduzia a saliência da proposta. Por fim, Bird e seus coautores especulam que os estudantes atuais podem ter mais informações sobre o preenchimento do formulário de auxílio estudantil, de modo que há menos estudantes para os quais os *nudges* gerariam uma diferença marginal.

Os resultados obtidos por Abhijit Banerjee e outros autores (2021a) e por Bird e outros autores (2021) parecem ser conflitantes. Os diferentes domínios em que as intervenções foram aplicadas - saúde e educação/finanças -, assim como a forma pela qual as decisões devem ser tomadas em ambos os casos e as fricções associadas a essas opções, certamente auxiliam a explicar essa distinção. De qualquer forma, para avaliar de maneira mais precisa os efeitos de *nudges* em larga escala, convém mencionar uma extensa revisão bibliográfica recentemente publicada por Stefano DellaVigna e Elizabeth Linos (2022).

Os autores coletaram dados de inúmeros experimentos conduzidos por duas *nudge units* dos Estados Unidos. A partir desses estudos, DellaVigna e Linos restringiram sua amostragem aos estudos randomizados controlados, em que houvesse um grupo de controle claro, sem alteração de incentivos financeiros e apenas analisaram experimentos com resultados binários como variáveis dependentes. DellaVigna e Linos também excluíram os experimentos envolvendo mudanças em regras *default*, que curiosamente correspondiam a apenas dois *nudges*, em um único experimento, embora essa seja uma intervenção bastante discutida na literatura.

O objetivo dessa última restrição foi garantir que as intervenções examinadas fossem comparáveis; ao fim e ao cabo, elas consistiam, em geral, em uma combinação de simplificação, personalização, estímulo à implementação de intenções, lembretes e comparações de normas sociais. O resultado foi uma amostragem de 126 experimentos, com 243 *nudges* diferentes e impacto em 23 de milhões de pessoas. Claramente, os dados de DellaVigna e Linos permitiam a avaliação de *nudges* implementados em larga escala.

Mais do que avaliar os impactos desses *nudges*, DellaVigna e Linos também compararam os seus efeitos com estudos públicos em periódicos acadêmicos. Para isso, basearam-se nas revisões sistemáticas conduzidas por Benartzi e outros autores (2017) e por Hummel e Maedche (2019), já mencionadas na Seção 5.1, além de fazer restrições semelhantes àquelas aplicadas à base de estudos empíricos de *nudges units*. Com isso, foram analisados 26 artigos na base de periódicos acadêmicos, com 74 *nudges* distintos.

Definidos os dois conjuntos amostrais que seriam avaliados - experimentos de *nudge units* e experimentos de periódicos acadêmicos -, DellaVigna e Linos identificaram uma enorme distinção dos tamanhos de efeito de ambas as bases. Enquanto em estudos acadêmicos, o efeito médio de um *nudge* é equivalente a 8,68 pontos percentuais, o efeito médio de estudos de *nudge units* é de 1,38 pontos percentuais. Embora esse efeito seja relevante estatisticamente, sobretudo se considerado o baixo custo dessa intervenção – custo marginal zero ou próximo disso –, ele é equivalente a apenas 1/6 do impacto de *nudges* em experimentos acadêmicos.

DellaVigna e Linos apontam três razões principais para tamanha diferença nos valores: a) o poder estatístico dos experimentos em cada uma das amostragens selecionadas[289]; b) o viés de publicação; e c) as características heterogêneas dos estudos analisados. Quando todos esses fatores foram controlados, quase toda a diferença identificada foi eliminada.

Inicialmente, quanto ao poder estatístico, DellaVigna e Linos verificaram que os estudos acadêmicos apresentavam tamanho de efeito mínimo detectável (*minimum-detectable effect size - MDE*) mediano de 6,30 pontos percentuais e o MDE médio de 8,18 pontos percentuais; na amostragem de *nudge units*, por outro lado, o MDE mediano era de 0,78 pontos percentuais e o MDE médio de 1,72 pontos percentuais. Em outras palavras, enquanto a maior parte dos experimentos acadêmicos está preparada para detectar apenas efeitos de tratamento muito amplos, os experimentos de *nudge units* são mais precisos e podem detectar tamanhos de efeitos menores. Essa precisão está relacionada a uma distinção na amostragem mediana dos grupos de tratamento de experimentos de artigos científicos e de experimentos de *nudge units*: enquanto os primeiros tinham uma amostragem mediana de 484 pessoas, a amostragem

[289] A análise do poder estatístico é realizada a partir da identificação do menor efeito que o experimento pode identificar com relação à variável observada, influenciada por uma determinada intervenção. A medida do poder estatístico, então, é o tamanho de efeito mínimo detectável (*minimum-detectable effect size – MDE*). Como explicam Eric. W. Djimeu e Deo-Gracias Houndolo (2016, p. 508), a análise do poder estatístico nas ciências sociais é essencial por três razões principais. Inicialmente, o cálculo do poder estatístico ajuda a se evitar as consequências de ter uma amostragem muito pequena para detectar a menor magnitude de interesse da variável influenciada pela política. Os pesquisadores, nesse caso, podem concluir que a intervenção não traz qualquer impacto, ainda que isso não seja verdade. Ao mesmo tempo, como a coleta de dados é custosa, os pesquisadores devem evitar a utilização de amostragens muito grandes sem necessidade. Finalmente, estudos sem poder estatístico podem ser antiéticos, já que o seu valor científico é muito baixo.

mediana dos segundos era de 10.006 pessoas. Como DellaVigna e Linos ponderam, essa diferença é crucial à escalabilidade de uma intervenção, já que amostragens mais amplas permitem estimativas de efeitos mais precisas.

A imprecisão de amostragens menores, ao mesmo tempo, favorece o viés de publicação, se o processo de publicação selecionar majoritariamente artigos com resultados significativos. Nesse sentido, se apenas resultados significativos estatisticamente forem publicados, a falta de estudos menos precisos levará a uma inferência enviesada de efeitos de tratamento maiores (DELLAVIGNA; LINOS, 2022). DellaVigna e Linos constataram que, em artigos publicados em periódicos, a probabilidade de publicação de artigos sem resultados significativos era apenas uma fração da probabilidade de publicação de artigos com pelo menos um resultado significativo – foram identificados apenas três casos de publicação de resultados não significativos, sendo que dois deles apresentaram resultados significativos como subamostragem ou em resultados diferentes[290,291]. Por outro lado, ao analisar os experimentos de *nudge units* que foram publicados em periódicos acadêmicos, DellaVigna e Linos notaram que, por mais que os incentivos ao viés de publicação permanecessem, o seu efeito não enviesava os resultados, já que os estudos eram bastante robustos estatisticamente.

Por fim, as características heterogêneas dos estudos avaliados por DellaVigna e Linos também podem explicar a distinção de tamanhos de efeitos que eles identificaram. Apenas para exemplificar, estudos publicados em periódicos acadêmicos envolvem mais intervenções pessoais, enquanto *nudge units* costumam se basear em cartas e e-mails, o que pode ter menos efetividade. Além disso, *nudge units* costumam se basear, amplamente, em técnicas de simplificação, que podem ser muito simples para gerar estudos para publicações em periódicos.

Outro elemento relevante é o raro uso de alterações de regras *default* por *nudge units*, o que também pode levar a intervenções menos efetivas. A exposição de meta-análises e revisões sistemáticas da Seção 5.1, afinal, indicou que regras *default* são, em geral, o *nudge* mais efetivo. Apesar disso, como DellaVigna e Linos ponderam, restrições institucionais podem limitar a implementação dessa técnica por *nudge units*; feitas para ser enxutas e trabalhar sem a

[290] Esse resultado é consistente com os dados encontrados por Mertens e outros autores (2022), por Luo (2021) e outros autores e por Cadario e Chandon (2020) que indicam um viés de publicação, de modo que os verdadeiros tamanhos de efeitos de *nudges* são provavelmente menores do que os dados encontrados nas meta-análises citadas na Seção 5.1.

[291] Jachimowicz e outros autores (2019), em contrapartida, em sua meta-análise de efeitos de regras *default*, identificaram um viés de publicação em sentido contrário: efeitos muito significativos não seriam reportados, talvez porque os pesquisadores não se importassem em reportar replicação do que já entendem ser efeitos fortes. De toda forma, como há alta heterogeneidade, os autores ponderam que esses resultados devem ser avaliados com cautela.

necessidade de alterações legislativas, *nudge units* tentam trabalhar com técnicas de intervenções mais simples e esse pode ser um fator relevante à sua menor efetividade.

DellaVigna e Linos apresentam, ainda, uma meta-análise que reconcilia os efeitos de *nudges* publicados em estudos acadêmicos e *nudges* implementados por *nudge units*. Comprovando que o poder estatístico, o viés de publicação e características heterogêneas dos estudos eram fatores de confusão, DellaVigna e Linos controlaram esses fatores e estimaram novamente os tamanhos de efeito de ambas as amostras selecionadas. Enquanto os efeitos de *nudges* implementados por *nudge units* aumentaram pouco, as estimativas de efeitos de *nudges* publicados em periódicos foram bastante reduzidas e se aproximaram ao acréscimo de 1,4 pontos percentuais.

Os dados apontados acima parecem indicar que, em larga escala, os efeitos de *nudges* parecem ser mais modestos do que se imagina. Há, de fato, alguns exemplos muito bem-sucedidos, como o experimento de Abhijit Banerjee e outros autores (2021a) ou os exemplos da Opower, discutidos na seção anterior. Outras propostas não foram bem-sucedidas, como evidenciam os experimentos de Kelli A. Bird e outros autores (2021) e de Gravert e Collentine (2021). Esses dados reforçam que *nudges* não são uma bala de prata e devem ser elaborados à luz do contexto em que eles serão aplicados e da população-alvo que será influenciada.

Os dados de DellaVigna e Linos, por seu turno, trazem ainda mais ceticismo à efetividade dos *nudges* em larga escala e em cenários reais. Apesar disso, a profundidade das conclusões que podem ser extraídas a partir da busca de um tamanho de efeito médio de todas as intervenções é incerta. O *hype* em torno das ciências comportamentais pode estar levando a uma percepção superestimada dos efeitos de *nudges*.[292] Nesse contexto, ao elaborar uma política pública, o arquiteto de escolhas deve buscar *nudges* que funcionam bem, reconhecendo que os resultados de experimentos nem sempre serão concretizados na realidade e em larga escala.

Experimentos em menor escala, no entanto, não são dispensáveis, mas complementares. Como Soman e Tanjim Hossain (2021, p. 87) argumentam, intervenções comportamentais frequentemente exigem uma abordagem sequencial das suas evidências, já que são altamente dependentes do contexto. Em razão de tal heterogeneidade, diferentemente de ciências naturais

[292] Além da meta-análise apresentada acima, DellaVigna e Linos (2022) fizeram um experimento a partir de questionários com acadêmicos e *policy-makers*, a respeito da efetividade de *nudges*. Seus resultados indicaram que, de fato, essa percepção é superestimada no que diz respeito às intervenções aplicadas por *nudge units*. As previsões média e mediana, de fato, eram de 4 e 5,8 pontos percentuais, respectivamente, enquanto, como demonstrado acima, os efeitos identificados eram de 1,38 pontos percentuais.

ou médicas, uma meta-análise rigorosa de evidências comportamentais não se traduz automaticamente em intervenções bem-sucedidas em larga escala.

Nesse contexto, o caminho para políticas públicas comportamentais bem-sucedidas envolve, em primeiro lugar, o teste de intervenções em ambientes de baixo-custo e de rápido resultado, como laboratórios ou painéis virtuais. As técnicas que apresentarem bons resultados podem ser testadas em estudos-piloto em escala cada vez maior, antes de serem implementadas enquanto políticas públicas concretas e em larga escala. Um alerta de Soman e Hossain (2021) é, ainda, fundamental: esses estudos devem ser feitos no contexto em que a intervenção será implementada em grande escala.

Mesmo depois de comprovar os seus efeitos em tais experimentos, é fundamental que a implementação em larga escala seja monitorada, com testes e adaptações contínuos, para que sejam evitadas quedas de tensão. Em outras palavras, pesquisadores e agentes públicos devem trabalhar em conjunto para produzir evidências; ou, como List (2022, p. 87) sugere, para obter ganhos de escala, *policy-makers* devem inverter o roteiro baseado em políticas públicas baseadas em evidências em direção à obtenção de evidências baseadas em políticas públicas[293].

5.3. Fortes preferências antecedentes, o equilíbrio da pescaria e reatância psicológica

A exposição das Seções anteriores evidencia que, por mais que *nudges* tenham efeitos estatisticamente significativos e com bom custo-benefício, eles nem sempre são efetivos, sobretudo quando implementados em larga escala e em longo prazo. O viés de publicação, inclusive, conforme demonstrado na Subseção 5.2.3, faz com que os efeitos de *nudges* sejam superestimados na literatura acadêmica. Há, ainda, mais alguns fatores que ajudam a explicar as situações em que *nudges* podem falhar e que serão mencionados nesta seção.

Ao analisar as razões que levam regras *default* a serem inefetivas[294], Sunstein menciona, inicialmente, que fortes preferências são provavelmente um motivo suficiente para garantir que as regras *default* não sejam seguidas. A inércia e a procrastinação, a aversão à perda e o sinal informacional explorados pela regra *default* - a facilidade, a aprovação e o efeito dotação

[293] O custo dessa alteração, evidentemente, não é desprezível. Soman e Hossain (2021) sugerem, como uma de suas propostas, a busca pela redução do custo da experimentação. Abhijit Banerjee e outros autores (2021a, p. 29) sugerem, por sua vez, que, diante do número cada vez maior de combinações de estratégias que um *policy-maker* pode utilizar, uma abordagem baseada em dados, em que técnicas de aprendizado de máquina podem ser utilizadas para identificar o conjunto esparso de políticas que afetam o resultado de interesse de maneira significativa, deve ser considerada.

[294] Como as regras *default* costumam ser os *nudges* mais efetivos, é provável que as situações em que *defaults* não funcionam sejam ainda mais relevantes a outras espécies de *nudge*. O raciocínio, então, pode ser replicado a esses outros casos.

demonstrados por Dinner e outros autores (2011) - não seriam suficientes para fazer com que o sujeito contrariasse uma preferência que lhe é muito importante.

Elizabeth F. Emens (2007) apresenta um exemplo interessante nesse sentido, ao analisar a regra *default* que disciplina a escolha do sobrenome por mulheres que se casam com homens nos Estados Unidos. Emens demonstrou, em primeiro lugar, que, atualmente, a regra *default* a esse respeito é a manutenção, pela mulher, do seu próprio sobrenome. Essa seria a opção que envolveria menos escolhas ativas e menos custos; inércia e procrastinação, assim como sugestionamento informativo e o efeito dotação, favoreceriam essa escolha. Apesar disso, a maioria das mulheres acaba escolhendo o sobrenome do marido. Diante dessa preferência forte e bem-definida – moldada, como Emens sugere, por convenções sociais -, um *nudge* como uma regra *default* é pouco efetivo: as mulheres optam ativamente pelo sobrenome dos maridos.

Do mesmo modo, Erin Todd Bronchetti e outros autores (2013) exemplificam experimentalmente a inefetividade de regras *default* diante de fortes preferências antecedentes. Os autores elaboraram um experimento de campo para avaliar em que medida a alteração de uma regra *default* poderia impactar as decisões de famílias de baixa renda, a respeito da poupança de suas restituições tributárias.

A ideia era avaliar em que medida a introdução de uma regra *default* que destinava parte dessas quantias a títulos do tesouro americano (*U.S. Saving Bonds*) aumentaria os índices de poupança entre essa população. Os resultados evidenciaram que a regra *default* não teve efeitos detectáveis. Segundo um questionário feito por Bronchetti e seus colegas, os participantes do experimento tinham consciência de que provavelmente receberiam as restituições e tinham planos de antemão para gastar esses valores. Mais uma vez, diante de fortes preferências antecedentes, a efetividade de um *nudge* se esvai.

Em outra perspectiva, Sunstein (2019a, p. 103-104) também menciona o risco de que agentes privados venham a se opor aos *nudges*, por meio de diversas estratégias, inclusive técnicas comportamentais, para que as pessoas abandonem a escolha *default*; pode ser que esses agentes usem contra-*nudges* (*counternudges*), o que pode levar a política a ser inefetiva. Se, por exemplo, um *nudge* - ou um *sludge* - é adotado para reduzir o consumo de álcool, pode ser que os grandes *players* desse mercado tentem "empurrar" as pessoas novamente a esse consumo, por meio de diversas estratégias comerciais[295].

[295] Nesse sentido, Mark Petticrew e outros autores (2020) identificaram, inicialmente, a existência de *dark nudges* - isto é, *nudges* que encorajam o consumo de produtos danosos - em comunicações da indústria do álcool a respeito do consumo responsável de álcool. Em tais comunicações, foram identificados tanto elementos de normas sociais - como avisos de que a maioria das pessoas está bebendo - quanto sinais verbais e pictoriais referentes à bebida,

Como discutido no Capítulo 4, em equilíbrio, o mercado é um campo aberto à exploração de falhas comportamentais das pessoas por empresas, isto é, à pescaria de tolos, descrita por Akerlof e Shiller (2015). Em outras palavras, se as pessoas forem estimuladas de alguma forma por fatores supostamente irrelevantes - e essas fraquezas puderem ser exploradas com o objetivo de auferir mais lucros, no equilíbrio da pescaria (*phishing equilibrium*), algum agente do mercado vai tirar proveito dessa situação. Esse elemento deve ser sempre levado em consideração pelo arquiteto de escolhas. Afinal, como List (2022) demonstrou, o equilíbrio de mercado é um dos fatores tendentes a gerar efeitos de transbordamento negativos, quando uma medida é implementada em larga escala.

As maiores possibilidades de um *nudge* - como uma regra *default* - falhar em razão de contra-*nudges* ou de outras estratégias promovidas por agentes do mercado são identificadas na aplicação de *nudges* impuros. No Capítulo 4, demonstrou-se que *nudges* impuros são implementados por agentes privados a partir de uma determinação prévia do Estado; são situações em que há uma relação jurídica privada entre um arquiteto de escolhas e um decisor a ser influenciado, e o Estado interfere indiretamente - por meio de medida de comando e controle - na arquitetura de escolhas, para definir como algum de seus aspectos deve ser organizado. Assim, embora haja um *nudge* para o decisor, o arquiteto de escolhas privado será submetido a uma medida coercitiva, caso deixe de organizar a arquitetura de escolhas da forma determinada pelo Estado.

A forma como *nudges* impuros podem ser inefetivos em razão de pressões dos próprios arquitetos de escolhas foi demonstrada extensamente por Lauren E. Willis (2013). Em primeiro lugar, Willis avaliou as condições que fazem uma regra *default* ter sucesso - ser "pegajosa" (*sticky*)[296]. Ao avaliar regras *default* implementadas por agentes privados e as ferramentas utilizadas para torná-las "pegajosas", Willis argumenta que as empresas rodeiam suas regras *default* com uma teia de regras de *opt-out* - também chamadas por Willis (2013, p. 1171-1173)

---

ao mesmo tempo em que se alertava a respeito dos riscos advindos do consumo excessivo de álcool. Além disso, Petticrew e seus colegas também identificaram *sludges*, como fontes, cores e *design*, que parecem explorar vieses cognitivos para dificultar o acesso às informações referentes aos danos à saúde provocados pelo álcool.

[296] Embora tenha apontado fatores semelhantes aos identificados por Dinner e outros autores (2011) para esse efeito, Willis (2013) aponta outros elementos como chave do sucesso de regras *default*. Se o ambiente de decisões for confuso e houver incerteza quanto às preferências, por exemplo, *defaults* tendem a ser "pegajosos". Também contribui para esse efeito barreiras de transação – o que depois veio a ser denominado de *sludge*, como demonstrado no Capítulo 4 – que dificultem o *opt-out*, como custos, confusão para fazer essa escolha ou até mesmo a sua invisibilidade, quando o sujeito não consegue definir como deixar de seguir a regra *default*. Além da aversão à perda, do efeito dotação e da procrastinação, Willis aponta outros vieses cognitivos como mecanismos tendentes a favorecer a regra *default*, como descontos intertemporais e o viés de omissão, em linha, inclusive, com Camerer e outros autores (2003) – vide nota 182 *supra*. Por fim, com relação à formação de preferências, além do aconselhamento implícito, Willis pondera que regras *default* podem ser positivas a partir da própria experiência do sujeito; pode ser, por exemplo, que, com o *default*, ele adquira novos hábitos.

de regras de alteração - e técnicas de enquadramento (*framing*) que dificultam a escolha de outra opção. Alguns exemplos são: a) processos burocráticos e demorados – repletos de *sludges* - para cancelamento de assinaturas automáticas; b) a reversão reiterada do consumidor ao *default*, o que exige repetidas escolhas pelo *opt-out*; c) aumento do número de opções necessárias para se afastar do *default*[297]; d) exploração do viés de omissão a partir de avisos de que o usuário toma uma decisão afirmativa ao mudar o *default* e de que a empresa não será responsável pelos resultados correspondentes; e) a tentativa de moldar as preferências dos usuários, a partir de recomendações e conselhos para utilização do *default* etc.

A partir de todas essas constatações, Willis chega a uma primeira conclusão importante: ao utilizar todas essas técnicas, "as empresas implicitamente reconhecem que defaults, por si só, não são necessariamente tão poderosos. É a combinação de um default, processos de *opt-out* e o enquadramento (framing) que é coletivamente poderosa" (WILLIS, 2013, p. 1173)[298]. Nesse caso, Willis aponta que, para que uma regra *default* seja efetiva, o arquiteto de escolhas deve identificar ou criar condições que explorem mecanismos que tornam os *defaults* pegajosos. Em relação a políticas públicas, para criar essas condições, pode ser necessário dificultar o *opt-out* a partir da utilização de *sludges* e explorar intensamente os diversos vieses comportamentais ligados à prevalência de regras *default*, nas comunicações a respeito da medida.

Por outro lado, a maior contribuição do trabalho de Willis (2013, p. 1200-1204) é apontar o papel primordial do alinhamento de preferências entre arquitetos de escolha privados e os governos que implementam regras *default*, para que essas regras sejam bem-sucedidas. De fato, Willis argumenta que a principal diferença entre algumas regras *default* bem-sucedidas e outras que falham é a presença de uma parte que se oponha suficientemente à política.

No exemplo clássico da poupança para aposentadoria - em planos como o *Save More Tomorrow* -, por exemplo, o arquiteto de escolhas - empregador - não se opõe à regra *default* e tem incentivos para participar; os prestadores de serviços de planos de previdência, por seu turno, beneficiam-se dos resultados da política. Em equilíbrio, nesse contexto, não haverá incentivos à adoção de *contra-nudges* ou de outras medidas que inviabilizem a regra *default*. Na presença de partes que se oponham à regra *default*, no entanto, os resultados são diferentes.

O principal exemplo dado por Willis (2013, p. 1174-1200) é a alteração da regra *default* envolvendo a utilização de limites de crédito de cheque especial (*overdraft coverage*) nos

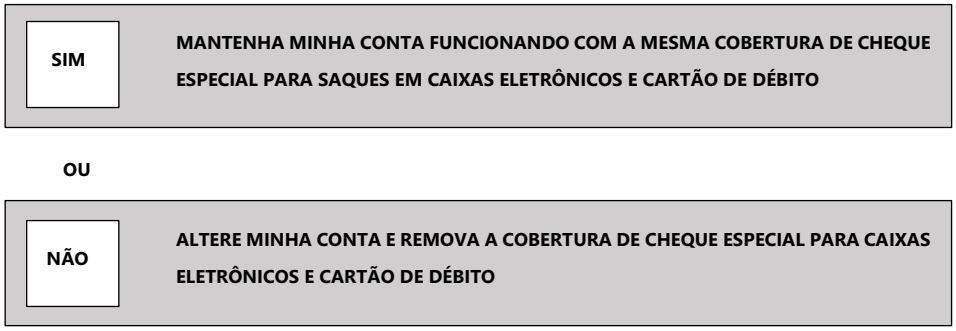
[297] Segundo Willis (2013, p. 1172), para rejeitar as regras *default* de privacidade do Facebook, por exemplo, os usuários precisam percorrer 50 botões e escolher entre 170 opções diferentes.
[298] Tradução livre. Texto original: "[...] *firms implicitly acknowledge that defaults, per se, are not necessarily all that powerful. It is the combination of a default, opt-out process, and framing that is collectively powerful*."

Estados Unidos. Em termos simplificados, com o objetivo de reduzir o endividamento a partir dessa linha de crédito, uma reforma promovida nos Estados Unidos exigiu que: a) a contratação desse limite de crédito fosse feita por escolhas ativas dos usuários, de modo que a regra *default* seria não contar esse limite e, consequentemente, não ser capaz de concluir transações via cartão de débito ou saques sem que houvesse fundos em suas respectivas contas; b) bancos fornecessem os mesmos termos e condições aos usuários que permanecessem com o *default*, em comparação àqueles que contratassem o limite de cheque especial; e c) técnicas de *framing* tentavam definir a forma pela qual o *opt-out* deveria ser apresentado aos consumidores. Em outras palavras, trata-se de uma combinação de *nudges* impuros que os bancos deveriam seguir.

Apesar disso, diferentemente do exemplo da previdência, bancos perdem muita receita com a redução da demanda por linhas de crédito de cheque especial. No equilíbrio da pescaria, nesse caso, os bancos tenderiam a explorar fatores supostamente irrelevantes para fazer com que o *default* fosse inefetivo e fazer com que as pessoas optassem por contratar o crédito.

Embora os reguladores bancários dos Estados Unidos esperassem que a regra *default* fosse prevalente, a política foi um fracasso: os bancos que buscaram obter receita com essa linha de crédito apresentaram altos índices de *opt-out*, especialmente entre os sujeitos com um histórico de utilização frequente do cheque especial e com novos correntistas. Para isso, uma multiplicidade de técnicas comportamentais, como a remoção dos custos de *opt-out*, o aumento dos custos da manutenção do *default*, mudanças de *framing*, entre outras, foram utilizadas para que consumidores aderissem ao cheque especial. Ao mesmo tempo, em razão do novo equilíbrio de mercado - provocado pela menor demanda posterior à regra *default* -, as taxas de juros do cheque especial aumentaram e usuários com maior perfil de endividamento acabaram arcando com essas taxas de forma desproporcional. Em outros termos, a política gerou efeitos de transbordamento negativos - e, como List discute, quedas de tensão em escala.

Figura 20 - Excertos de formulário utilizado pelos bancos para o *opt-out*



Fonte: Adaptação do exemplo apresentado por Willis (2013, p. 1192).

A mensagem que os exemplos de Willis parecem fornecer é extremamente relevante à elaboração de *nudges* que pretendam ser efetivos. Ao propor *nudges* impuros, o regulador deve estar atento ao equilíbrio de mercado e à estrutura de incentivos do arquiteto de escolhas. Se os seus incentivos forem substancialmente contrários a uma regra *default* ou a outros tipos de *nudges*, o arquiteto de escolhas pode agir ativamente para garantir a inefetividade da política proposta. No limite, pode haver uma batalha de *nudges*[299] - como no exemplo da previdência na Suécia[300] -, o que poderia levar à completa inefetividade da política.

Ao propor um *nudge* impuro, o regulador reduz a liberdade do arquiteto de escolhas para definir um aspecto específico da arquitetura de escolhas. Apesar disso, no âmbito de uma relação privada, se essa redução de liberdade estiver contrária aos próprios interesses econômicos do arquiteto de escolhas privado - e, portanto, ele tiver incentivos para agir de maneira contrária à política -, ele terá capacidade de alterar todos os demais aspectos da arquitetura de escolhas não afetados pelo *nudge* impuro, a fim de torná-lo inefetivo[301]. Controlar

[299] Agradeço à Professora Dra. Rute Saraiva por chamar a minha atenção a este ponto.

[300] O exemplo sueco evidencia, ainda, que se os incentivos para se oporem aos *nudges* acabarem, os agentes do mercado passarão a reagir de outro modo. De fato, no momento em que houve a reforma da previdência - e um grande volume de suecos economicamente ativos adeririam aos planos de contribuição fixa -, havia um grande incentivo para que gestores de planos de previdência investissem em propaganda e em maneiras de afastar as pessoas das regras *default*. Com o tempo, novos aderentes à previdência sueca não crescem com a mesma velocidade; o retorno a esses investimentos não seria tão grande. Afinal, as propagandas e demais estratégias para afastar as pessoas da regra *default* teriam que ser mantidas continuamente, mas não haveria o mesmo fluxo de possíveis novos clientes proporcionado pela reforma, que afetou toda a população economicamente ativa.

[301] Esse mesmo aspecto vale a questões ligadas à privacidade. Por mais que se exija o consentimento ativo dos titulares de dados pessoais para determinadas formas de tratamento de dados, o arquiteto de escolhas - normalmente, grandes plataformas – tem meios para fazer com que esse consentimento seja dado de maneiras praticamente automáticas. Os avisos de *cookies*, por exemplo, em que os usuários devem dar consentimento para acessar páginas específicas em um site, acabam sendo aceitos automaticamente pela maior parte dos usuários. A consequência talvez seja a detenção de um volume de dados por empresas digitais superior ao que os próprios usuários gostariam, o que eventualmente pode lhes causar prejuízos. Há, nesse cenário, um paradoxo da

coercitivamente todos esses aspectos seria uma medida muito drástica, que, provavelmente, neutralizaria o custo-benefício da política, além de impor custos desproporcionais ao arquiteto de escolhas privado. Essa discussão será retomada no Capítulo 6, na avaliação da medida em que *nudges* são capazes de proporcionar maior bem-estar social e individual.

Por fim, um terceiro elemento que pode gerar *nudges* inefetivos é a reação negativa de algumas pessoas a eles, simplesmente por serem tentativas do Estado de direcionar o seu comportamento (SUNSTEIN, 2019a, p. 113). Esse fenômeno é conhecido - e já descrito na literatura da psicologia há décadas - como reatância psicológica. Essa teoria envolve o conjunto de consequências motivacionais que pode ser esperado sempre que a liberdade for ameaçada ou perdida, e como o estado motivacional subsequente irá se manifestar (BREHM; BREHM, 1981, p. 4).

De acordo com Sharon S. Brehm e Jack W. Brehm (1981, p. 4), em obra clássica a esse respeito, essa teoria sustenta, em geral, que uma ameaça ou uma perda de liberdade motiva o sujeito a restaurá-la. Essa sensação, então, pode ser seguida por uma tentativa de exercer a liberdade. No exemplo de Brehm e Brehm, se o sujeito descobre que quotas de importação foram estabelecidas em veículos da Honda, ele poderá apresentar uma maior motivação para adquirir um veículo como esse, ainda que precise adquiri-lo de um vendedor muito distante de sua casa - muito mais distante do que ele se deslocaria normalmente.

O risco, então, é que um *nudge* - especialmente mais interventivo, como uma regra *default* - seja percebido como uma limitação da liberdade. A reação poderia ser simplesmente decidir em sentido contrário, o que eliminaria o efeito do *nudge*. Como mencionado anteriormente[302], um dos mais clássicos trabalhos das ciências comportamentais evidenciou como a presunção de consentimento - regra *default* - levava a níveis mais elevados de consentimento para doações de órgãos em diversos países (JOHNSON; GOLDSTEIN, 2003). Apesar disso, a criação de uma regra como essa pode desencadear efeitos semelhantes à reatância psicológica.

Um exemplo brasileiro pode ilustrar esse ponto. Stancioli e outros autores (2011, p. 134-136) relatam que a Lei n.º 9.434 estabeleceu, inicialmente, a figura da doação presumida de transplantes - isto é, a presunção de consentimento dos doadores que não se manifestassem em sentido contrário. Os dados do Ministério da Saúde no momento em que a regra foi elaborada

privacidade: por mais que a valorizem, as pessoas renunciam à privacidade por recompensas muito baixas, em diversos contextos. Nesse sentido, Cf. (ACQUISTI; BRANDIMARTE; LOEWENSTEIN, 2020). Esse elemento parece estar indissociavelmente ligado à pouca efetividade de leis de proteção de dados pessoais, identificada, por exemplo, por Filippo Lancieri (2021).

[302] Cf. nota 194 *supra*.

indicavam que 70% dos brasileiros aceitariam doar seus órgãos, mas acabavam não o fazendo. O que se seguiu à alteração da regra, no entanto, foi uma enorme repercussão social, com inúmeras pessoas posicionando-se ativamente para reverter a regra *default*. Por mais que possa ser explicada por outros fatores[303], essa reação pode estar relacionada à reatância psicológica. A senadora Benedita Silva, por exemplo, chegou a dizer que a população se manifestava no sentido de que havia uma interferência e uma invasão em seus direitos (STANCIOLI *et al.*, 2011, p. 135). Diante de tal reação, Stancioli e seus colegas acrescentam: alterações legislativas substituíram a doação presumida pela decisão familiar a respeito da doação de órgãos – regra que permanece vigente até hoje[304].

Há, na literatura, inúmeros estudos que já identificaram efeitos consistentes com a reatância psicológica, em resposta a *nudges*[305]. Apenas para exemplificar, Arad e Rubistein (2018) identificaram experimentalmente que, em uma situação hipotética, mais pessoas permaneceriam vinculadas a um plano de poupança se optassem por se cadastrar mediante uma escolha ativa, do que pessoas permaneceriam nesse plano com o estabelecimento de uma regra *default*. No experimento de Arad e Rubinstein, um grupo considerável de indivíduos optava por se desvincular do plano estabelecido como regra *default* (*opt-out*), quando tinha ciência do *nudge*.

Por outro lado, ainda que demonstrem substancial apoio das pessoas aos *nudges*, inúmeros experimentos também constataram que um grupo significativo de indivíduos rejeita essas estratégias de políticas públicas, especialmente no caso de regras *default*, o que pode desencadear a reatância psicológica e fazer com que os respectivos *nudges* falhem[306]. Questões políticas – especialmente em ambientes polarizados - podem, ainda, estimular a reatância psicológica, já que pode haver uma maior reprovação a propostas apresentadas por políticas de espectro contrário ao do indivíduo. Dora L. Costa e Matthew E. Kahn (2013), por exemplo, demonstraram, no caso de um *nudge* elaborado para reduzir o consumo de energia, que

---

[303] Stancioli e seus coautores (2011, p. 134-136) mencionam questões como a insegurança ligada às deficiências do sistema de recepção e distribuição de órgãos, as condições de pesquisa científica e a realidade do comércio ilegal de órgãos.

[304] Em exemplo semelhante ao caso brasileiro, Saraiva (2020, p. 46) recorda-se do exemplo do repúdio nos Países Baixos à alteração da lei de doação de órgãos, nos seguintes termos: "Face às dificuldades encontradas na doação de órgãos, a Câmara Baixa do Parlamento holandês, em 2016, decidiu alterar a solução por defeito, seguindo o exemplo bem aceite e sucedido em muitos outros Estados, fixando uma regra de opt-out. Ora, o que parecia uma decisão acertada, até pela bondade inerente da causa, foi repudiada pelos cidadãos com os números de recusas a baterem recordes, obrigando consequentemente a ajustes. Ou seja, um nudge robusto mais do que testado e aprovado revelou-se afinal problemático, impostor e eticamente reprovável."

[305] Para um mapeamento de tais estudos, Cf.: (OSMAN, 2020).

[306] Nesse sentido, Cf.: (ARAD; RUBINSTEIN, 2018; BRANSON *et al.*, 2012; FELSEN; CASTELO; REINER, 2013; HAGMAN *et al.*, 2015; JUNG; MELLERS, 2016; REISCH; SUNSTEIN, 2016; SUNSTEIN, 2016b; TANNENBAUM; FOX; ROGERS, 2017).

conservadores apresentavam uma tendência de se posicionar contrariamente à política, se comparados com liberais.

A relação entre *nudges* e a reatância psicológica - assim como outros efeitos de transbordamento ligados a posições políticas ou outros aspectos interpessoais - é mais uma agenda de pesquisas que deve ser estimulada na literatura a respeito da efetividade de *nudges*. A mensagem que fica, desta breve exposição, é a necessidade de o arquiteto de escolhas não transparecer uma tentativa de limitar a liberdade ao propor a utilização de um *nudge*. Elaborar políticas públicas de maneira transparente - além de um dever ético, como será discutido no Capítulo 6 - é uma medida útil para garantir, inclusive, a efetividade da proposta.

5.4. O papel dos *nudges* no conjunto de técnicas regulatórias: necessidade de combinação de diferentes técnicas de intervenção

A medida em que *nudges* podem realmente ser uma técnica efetiva para resolver problemas e alterar o comportamento é uma questão inerentemente empírica. Os dados apresentados ao longo das últimas seções evidenciam que *nudges* são efetivos e têm bom custo-benefício, mas seus efeitos são menores do que o esperado. Os resultados experimentais nem sempre se traduzem da mesma maneira em grande escala, e esses resultados podem não ser persistentes. Há, ainda, dificuldades de implementação de *nudges* se as pessoas tiverem fortes preferências antecedentes, agentes do mercado tiverem incentivos para boicotar e reduzir a dificuldade do *nudge* e se os decisores sentirem a sua liberdade reduzida, o que pode gerar a reatância psicológica.

De qualquer forma, no argumento que critica a efetividade dos *nudges*, exemplificado acima pelas abordagens de Burgess (2012, p. 13) e de Selinger e Whyte (2012), há duas premissas equivocadas que devem ser corrigidas: a) a ideia de que há problemas exclusivamente comportamentais; e b) pensar *nudges* como ferramenta isolada e alternativa às medidas tradicionais e às demais técnicas de intervenção informadas por *insights* comportamentais. Corrigir essas duas premissas é uma tarefa importante para finalizar este Capítulo, já que, embora os efeitos de *nudges* isolados sejam modestos, *nudges* podem ampliar a efetividade de políticas tradicionais e podem ser implementados em conjunto a elas.

Como discutido no Capítulo 2, para inúmeros problemas, as razões para se intervir podem ser apresentadas pela economia tradicional, pela economia comportamental ou por uma análise híbrida entre esses dois paradigmas (LOEWENSTEIN; CHATER, 2017, p. 29). Thaler (2017), nesse sentido, afirma que a boa economia comportamental deve incluir tanto fatores

estruturais e econômicos, como fenômenos psicológicos. Afinal, esse é o traço que diferencia a economia comportamental da psicologia: o reconhecimento de que as forças do mercado atuam de maneira direta e central. Dizer que o viés do presente contribui à obesidade, por exemplo, não significa sustentar que esse viés é a principal ou a única causa para esse fenômeno; outros elementos estruturais e econômicos não são ignorados pela economia comportamental. Assim, na maior parte dos casos, não há problemas puramente comportamentais.

Por outro lado, conforme demonstrado no Capítulo 4, *nudges* são apenas uma técnica de intervenção. Os fundamentos para utilizá-los podem, muito bem, ser estruturais, econômicos ou comportamentais. A decisão da Comissão Europeia de exigir que o Google oferecesse espaço no topo de sua plataforma para outras ferramentas de busca além do Google Shopping, por exemplo, fundamenta-se por uma falha de mercado clássica: o poder de mercado. Elementos comportamentais como o viés do *status quo* e a saliência podem servir como fundamentos complementares que explicam a decisão. A implementação de um programa como o *Save More Tomorrow*, por outro lado, é baseada em larga medida em elementos comportamentais como a inércia, o viés do *status quo* e a procrastinação, mas também pode ser explicada por fatores estruturais como a analfabetização financeira e as externalidades geradas pelos gastos do Estado com a seguridade social.

Assim como há motivos de natureza diversa que podem justificar a implementação de um *nudge*, Loewenstein e Chater (2017, p. 28-29) também ponderam que técnicas diversas podem ser combinadas. Dessa forma, diante de um problema específico, diversas estratégias podem ser utilizadas em conjunto, incluindo técnicas tradicionais, como impostos, subsídios, imposições ou proibições, e técnicas comportamentais, como *nudges* e outras intervenções alternativas, como *boosts*, *nudges plus*, *thinks* e *sludges*. Compreender a efetividade de todas essas estratégias, em conjunto com *nudges* ou de forma separada, é mais uma agenda de pesquisas incipiente, mas importante nas ciências comportamentais[307,308].

[307] Beshears e Kosowsky (2020, p. 15) apontam que, dos 174 artigos da amostragem de sua revisão sistemática, apenas 48 combinavam *nudges* com pelo menos um teste com alguma outra intervenção, seja ela um *nudge*, seja ela uma intervenção de outra natureza. Os autores ressaltam, ainda, que mais pesquisas combinando essas estratégias são necessárias para guiar políticas públicas a respeito da combinação ótima de intervenções, a fim de atingir objetivos específicos em termos de políticas públicas.

[308] Já é possível identificar na literatura algumas experiências isoladas a respeito da efetividade de técnicas de intervenção comportamental diferentes de *nudges*. Luo e outros autores (2021), na meta-análise discutida na Seção 5.1, também mensuraram os efeitos de *sludges* e identificaram uma média meta-analítica de 0,27 no índice *d* de Cohen, o que indica um tamanho de efeitos pequeno. Jacob Bradt (2022), por outro lado, comparou os efeitos de *nudges* e *boosts* no aumento da demanda por seguros contra enchentes. *Nudges* que apelavam ao afeto e à informação foram as técnicas mais efetivas, enquanto o *boost* acabou reduzindo a demanda por seguros, embora não de forma robusta. Henrico van Roekel, Joanne Reinhard e Stephan Grimmelikhuijsen (2022), por sua vez, compararam *nudges* e *boosts* elaborados para aumentar a conformidade com a higiene das mãos de enfermeiros

No caso da regulação do uso de cigarros, por exemplo, é possível pensar em técnicas tradicionais, como impostos extrafiscais ou a proibição de propagandas, e em técnicas comportamentais, como anúncios gráficos que evocam os malefícios trazidos à saúde pelo uso de cigarros ou a alteração da disposição de cigarros em lojas, para que eles fiquem em posições menos visíveis. Como destacam Loewenstein e Chater (2017, p. 35), "[i]soladas, é provável que nenhuma dessas intervenções teria tido muito impacto nos índices de fumantes, mas a combinação [das intervenções] parece ter tido um grande impacto"[309].

No mesmo sentido, Benartzi e seus coautores (2017, p. 11), assim como Beshears e Kosowsky (2020, p. 15-16), ressaltam que *nudges* nem sempre são a melhor ferramenta em busca de determinados objetivos em termos de políticas públicas. Como discutido no Capítulo 2 e como Thaler e Sunstein ressaltam em diversas oportunidades, há casos que exigem proibições e imposições. O combate ao crime é um exemplo. Em outros casos, *nudges* funcionam como parte de outras políticas já estabelecidas. Em programas de poupança para a aposentadoria, por exemplo, *nudges* apenas facilitam a adesão dos participantes a planos já existentes. Estímulos ao planejamento da vacinação, por outro lado, são elementos acessórios de políticas de vacinação já estabelecidas. Inúmeros outros exemplos podem ser pensados no mesmo sentido.

De maneira semelhante, Mertens e outros autores (2022, p. 7) afirmam que a aplicabilidade da arquitetura de escolhas não é restrita a intervenções "autônomas"; técnicas de intervenção híbrida, que combinam a arquitetura de escolhas com intervenções mais tradicionais também devem ser utilizadas. Inclusive, intervenções tradicionalmente econômicas, baseadas na alteração da estrutura de incentivos, podem ser aprimoradas a partir da arquitetura de escolhas[310]. Esse papel complementar é importante, sobretudo, quando se

---

hospitalares. Embora ambas as estratégias tenham se provado efetivas, *nudges* apresentaram efeitos imediatos mais robustos, enquanto os efeitos de *boosts* foram mais persistentes e permaneceram estáveis por uma semana após a remoção da intervenção. Por fim, Sanchayan Banerjee e outros autores (2021b) avaliaram a efetividade de *nudges*, *thinks*, *boosts* e *nudges plus* para reduzir o consumo de alimentos intensivos em carbono, em um *survey online* conduzido no Reino Unido. Seus resultados indicaram que *nudges plus* foram a técnica mais efetiva, com redução de emissões de gases de efeito estufa em 0,32 unidades padronizadas. *Thinks*, por sua vez, reduziram essas emissões em 0,21 unidades padronizadas, enquanto *nudges* e *boosts* geraram uma redução de 0,2 unidades padronizadas.

[309] Tradução livre. Texto original: "*In isolation, it is likely that none of these interventions would have had much impact on smoking rates, but the combination does appear to have had a major impact*".

[310] David J. Hardisty, Johnson e Elke U. Webber (2010), por exemplo, demonstraram que a alteração do enquadramento de uma escolha - a partir da rotulagem de cobranças de valores para custear o meio-ambiente como um imposto ou como uma compensação - pode ser capaz de alterar as preferências das pessoas a depender de suas posições políticas. No experimento de Hardisty, Johnson e Webber, Republicanos e Independentes escolhiam produtos mais caros, com uma sobretaxa destinada ao dióxido do carbono emitido, quando essa sobretaxa era enquadrada como uma compensação. As preferências de Democratas mantinham-se constantes em ambos os enquadramentos. Tatiana A. Homonoff (2018), por sua vez, avaliou se enquadrar um incentivo como uma

considera que nem sempre técnicas tradicionais funcionam da maneira esperada pelos modelos econômicos neoclássicos.

Bowles (2016), por exemplo, argumenta que tentar explorar o autointeresse sob a forma de incentivos monetários pode enfraquecer impulsos morais que são necessários para que determinadas políticas funcionem. Sem esses impulsos morais, os resultados de tais políticas são subótimos - piores do que antes da instituição do incentivo financeiro[311]. *Nudges*, nesse contexto, podem contribuir para apresentar a mensagem correta ao inserir incentivos em uma determinada política; fazer com que os incentivos apelem para o autointeresse, sem substituir valores morais intrínsecos. Esse papel complementar não pode ser negligenciado.

Ao propor a utilização de *nudges*, Thaler e Sunstein não pretendiam que *nudges* fossem a única ferramenta disponível ao regulador. Thaler (2017) argumenta, inclusive, que não conhece nenhum economista comportamental, *policy-maker* ou jornalista que tenha afirmado que *nudges* são uma panaceia ou são a ferramenta apropriada para enfrentar todo problema de política pública. Sunstein (2018b, p. 65), no mesmo sentido, afirma que o programa de abordagens informadas por *insights* comportamentais vai além dos *nudges* e pode incluir imposições, proibições e incentivos, que podem, inclusive, ser recomendados como técnicas mais efetivas em determinadas situações, o que também foi apontado no Capítulo 2.

Sunstein (2018b, p. 65) aponta, ainda, que *nudges* podem, de fato, não ser suficientes para resolver alguns problemas, como a eliminação da pobreza, o desemprego e a corrupção. De todo modo, pensar em estratégias isoladas para qualquer um desses problemas não parece ser uma estratégia razoável. Como pondera Thaler (2017), diante de problemas específicos, o caminho ideal pode ser conduzir uma "auditoria da arquitetura de escolhas" (*choice architecture audit*), para identificar as decisões mais críticas que os vários atores envolvidos devem tomar, assim como as alavancas – comportamentais e econômicas - que os reguladores podem utilizar para melhorar os resultados.

Portanto, para problemas complexos, como a mudança climática ou os cuidados com a saúde, abordagens informadas por *insights* comportamentais devem tentar incorporar um rol

---

penalidade ou uma recompensa teria influência na efetividade da política. Foram avaliadas duas políticas que almejavam reduzir a utilização de sacolas plásticas descartáveis. Em linha com aversão à perda, o enquadramento como penalidade reduzia a utilização das sacolas plásticas em 40%, enquanto a recompensa não alterava o comportamento.

[311] Em um experimento clássico, Uri Gneezy e Aldo Rustichini (2000) elaboraram um experimento de campo em que foram introduzidos incentivos monetários para os pais que atrasavam para buscar seus filhos em uma creche e forçavam os professores a ficarem até mais tarde. Esses pais, de acordo com o experimento, foram obrigados a pagar uma multa monetária pelo atraso. O resultado da política é significativo a respeito da possibilidade de incentivos afastarem o papel (*crowd-out*) de valores morais: após a introdução dos incentivos, o número de pais atrasados aumentou consideravelmente. Mesmo depois da exclusão da multa, o atraso permaneceu acontecendo.

bastante diverso de intervenções (THALER, 2017). Nesse cenário, interpretar resultados modestos em termos estatísticos como um argumento contrário à utilização de *nudges* não parece ser uma estratégia adequada. Como Sunstein (2018b, p. 65) pondera, uma iniciativa pequena direcionada a uma grande população pode ter efeitos massivos. Do mesmo modo, Thaler (2017) argumenta que, como não há balas de prata para resolver questões complexas, pequenas mudanças podem ser importantes, especialmente se tiverem um custo-benefício razoável (*cost-effective*).

5.5. Principais conclusões do Capítulo 5

Este Capítulo pretendeu realizar uma revisão bibliográfica da efetividade dos *nudges*. Inicialmente, foi apresentada uma revisão de algumas importantes meta-análises e revisões sistemáticas de experimentos que implementam *nudges*. Em seguida, foram discutidos os desafios à implementação de *nudges* em larga escala. Feita essa exposição, foram avaliadas algumas questões adicionais que podem levar à inefetividade dos *nudges* – notadamente, fortes preferências antecedentes, incentivos aos agentes do mercado para prejudicar a efetividade do *nudge* e a reatância psicológica. Por fim, argumentou-se que *nudges* devem ser vistos como uma ferramenta complementar no arcabouço regulatório, que, além de ter efeitos autônomos, pode complementar o efeito de técnicas tradicionais e de técnicas comportamentais. As principais conclusões deste Capítulo foram as seguintes:

1. *Nudges* têm efeitos estatisticamente significativos e, de um modo geral, alteram o comportamento influenciado com tamanhos de efeito pequenos para médios. Esses efeitos são mais intensos em *nudges* que exploram fatores supostamente irrelevantes e se baseiam em processos automáticos, como a alteração de regras *default*. Ao mesmo tempo, o baixo custo de implementação faz com que *nudges* tenham bom custo-benefício, por vezes superior ao de técnicas tradicionais. Mais pesquisas são necessárias para maior compreensão desse ponto. Em alguns domínios, como alimentação, saúde, consumo e meio-ambiente, *nudges* conseguem gerar efeitos mais significativos, enquanto em outros, como nas finanças, seus efeitos são mais modestos;
2. Os efeitos estatisticamente significativos da literatura não são imediatamente traduzíveis em políticas de grande escala. Para elaborar *nudges* que afetem populações maiores, o arquiteto de escolhas deve evitar quedas de tensão, que inviabilizem um empreendimento em larga escala. Cinco fatores colaboram para que esse efeito ocorra:

(i) falsos positivos; (ii) o desconhecimento do público-alvo; (iii) o desconhecimento dos elementos não negociáveis que sustentam o sucesso da política; (iv) efeitos de transbordamento negativos; e (v) custos muito altos em grande escala. Além disso, a literatura parece indicar, em primeiro lugar, que *nudges* podem produzir mudanças de comportamento em longo prazo, a partir da mudança de hábito, da adoção de novas tecnologias ou de outros fatores, como o viés do *status quo*. Indica, ainda, que *nudges* implementados em políticas concretas por *nudge units* têm efeitos cerca de seis vezes menores do que os efeitos de resultados publicados em periódicos acadêmicos. Nesse contexto, para elaborar políticas bem-sucedidas - e identificar *nudges* realmente efetivos -, governos devem trabalhar em conjunto com acadêmicos, em prol da produção de evidências baseadas em políticas públicas;

3. Outros fatores auxiliam a compreensão das circunstâncias em que *nudges* falham. Em primeiro lugar, quando o público-alvo tiver fortes preferências a respeito de um determinado tipo de decisão, *nudges* não serão suficientes para provocar verdadeira mudança de comportamento. Ao mesmo tempo, no caso de *nudges* impuros, se houver um desalinhamento dos incentivos do governo e do arquiteto de escolhas privado, no equilíbrio da pescaria, esse último agente tomará medidas para tornar o *nudge* inefetivo, como contra-*nudges*. Esse efeito pode gerar uma batalha de nudges públicos e privados, que poderia levar à inefetividade da política ou eliminar o seu custo-benefício. Finalmente, se *nudges* forem entendidos pela população-alvo como uma restrição de sua liberdade, a população poderá tomar medidas para exercer essa liberdade e, assim, decidir de forma contrária à indicada pelo *nudge*. Trata-se do fenômeno da reatância psicológica;

4. *Nudges* devem ser vistos como mais uma ferramenta disponível ao Estado para regular o comportamento privado. A implementação conjunta de *nudges*, técnicas de regulação tradicional - como impostos, subsídios, proibições e imposições - e técnicas de regulação comportamental, como *sludges*, *boosts*, *thinks* e *nudges plus*, deve ser estimulada diante de um problema cujas soluções são promovidas por políticas públicas. De todo modo, ainda que os efeitos de *nudges* sejam modestos, os efeitos de uma iniciativa pequena aplicada em larga escala e com bom custo-benefício podem ser massivos.

Apresentada a revisão bibliográfica da efetividade dos *nudges*, é necessário, então, avaliar a última questão que será enfrentada nesta dissertação. No Capítulo 6, a legitimidade

dos *nudges* será posta em debate. Se o Capítulo 5 avaliou a medida em que *nudges* podem ser efetivos, o Capítulo 6 discutirá três pontos principais: a) *nudges* são capazes de gerar ganhos de bem-estar individual e social?; b) *nudges* são compatíveis com a autonomia?; e c) como evitar que *nudges* sejam utilizados para finalidades abusivas, a partir das restrições institucionais existentes no Estado Democrático de Direito.

6. *NUDGES* QUE PROMOVEM O BEM-ESTAR SEM PREJUÍZOS À AUTONOMIA: ARGUMENTO EM FAVOR DA UTILIZAÇÃO DE *NUDGES* EM POLÍTICAS PÚBLICAS[312]

> "Pois bem: imagine um quadrado, um bonito quadrado vivo. E ele precisa falar sobre si mesmo, sobre sua própria vida. Compreenda que a última coisa que passaria pela cabeça do quadrado seria falar que possui quatro lados iguais: isso é algo que ele simplesmente já não vê, para ele isso é muito habitual, cotidiano" (ZAMIÁTIN, 2017, p. 41).

> "Espera-se submeter ao jugo benéfico da razão os seres desconhecidos, habitantes de outros planetas, que possivelmente ainda se encontrem em estado selvagem de liberdade. Se não compreenderem que levamos a eles a felicidade matematicamente infalível, o nosso dever é obrigá-los a serem felizes" (ZAMIÁTIN, 2017, p. 16).

Como demonstrado ao longo dos últimos cinco Capítulos, os *nudges* englobam um conjunto de intervenções que altera a arquitetura de escolhas para, a partir da exploração de fatores supostamente irrelevantes, reduzir as fricções hedônicas, sociais ou cognitivas associadas a uma determinada opção de escolha e estimular as pessoas a escolhê-la. Em outras palavras, há uma redução sutil de custos não monetários, para que, sem coerção ou alteração de incentivos, as pessoas sejam estimuladas a optar por uma determinada opção. Embora os seus efeitos sejam modestos, *nudges* têm bom custo-benefício e podem provocar mudança de comportamento, com potencial para implementação em larga escala e com potenciais efeitos persistentes.

Apesar disso, como já mencionado nesta dissertação, embora seja bastante influente, a arquitetura de escolhas passa por vezes despercebida. Assim como o quadrado descrito pelo engenheiro D-503, protagonista de "Nós", a célebre distopia do russo Ievguêni Zamiátin, não vê que tem quatro lados iguais e os peixes descritos por Foster Wallace desconhecem a água[313], as pessoas não costumam notar que são influenciadas a todo momento pela arquitetura de escolhas.

Como qualquer intervenção na vida privada, especialmente quando adotados a partir de fundamentos paternalistas, *nudges* geram preocupações ligadas aos limites da regulação do

---

[312] Algumas passagens deste Capítulo foram desenvolvidas no artigo "Ética, Epistemologia e Prática dos Nudges", escrito em coautoria com Professor Dr. Renato César Cardoso, como resultado do meu trabalho de conclusão de curso, e no artigo "Para além da coerção, da manipulação e da doutrinação: como nudges podem ser legítimos em termos de autonomia", escrito como trabalho de conclusão da disciplina Temas de Direito Civil: Bases Teóricas e Novos Contornos da Autonomia Privada, oferecida pelo Professor Dr. Fábio Queiroz Pereira e pela Professora Dra. Mariana Alves Lara, no primeiro semestre de 2020.

[313] Cf. nota 189 *supra*.

comportamento privado e à correspondente geração de bem-estar. Há, ainda, preocupações relacionadas à autonomia, à expansão da intervenção estatal na vida privada e a um possível aumento do controle sobre as pessoas. Afinal, por trás de benefícios de bem-estar, podem existir razões escusas. O Estado Único descrito na obra distópica de Zamiátin citada na epígrafe, por exemplo, tinha o objetivo levar a felicidade matematicamente infalível para seres desconhecidos e habitantes de outros planetas, ainda que em detrimento da sua liberdade.

Nesse contexto, desde a elaboração do conceito de *nudge*, uma enorme controvérsia se instalou na literatura a respeito da sua legitimidade, com críticas de diversas matrizes.[314] O teor distópico e alarmista dos exemplos da obra de Zamiátin reflete o tom de algumas das críticas aos *nudges* e ao paternalismo libertário. Segundo o economista francês Giles Saint-Paul (2011, p. 72), por exemplo, na sua crítica às ciências comportamentais e ao paternalismo libertário, "[q]uando as preferências reveladas são jogadas pela janela, o Grande Irmão[315] entra pela porta dos fundos. Qualquer ação individual pode ser examinada e penalizada sob o fundamento de que ela exerce alguma externalidade indesejável nos sentimentos de outra pessoa"[316]. Em outra passagem, Saint-Paul afirma que, quando o Estado se move em uma direção em que se busca maximizar os fluxos agregados de felicidade das pessoas, há um amplo espaço para "manipular as preferências individuais para que as pessoas se sintam bem. Isto é, na verdade, o que a sociedade do admirável mundo novo de Huxley faz"[317].

O viés desta dissertação, contudo, não é propriamente pessimista e, muito menos, distópico. Como já mencionado na Introdução, adotou-se a hipótese de que há justificativas para a implementação de *nudges* ainda que com fundamento paternalista, considerando a influência dos fatores supostamente irrelevantes no processo decisório e a inviabilidade de se valer de escolhas ativas em alguns contextos. Para começar a testá-la, foi necessário, em primeiro lugar, superar as armadilhas teóricas que criam impasses ao debate normativo – como o Grande Debate da Racionalidade, a definição de *nudge* e as demais questões conceituais discutidas nos Capítulos 3 e 4. Em seguida, no Capítulo 5, apresentou-se um panorama geral a respeito da efetividade dos *nudges* e os contextos em que eles mais bem se aplicam, a fim de evidenciar o alcance e as limitações dos *nudges*, além de pontos que devem ser levados em consideração para maximizar a efetividade de tais técnicas de intervenção.

---

[314] Para uma revisão de literatura sobre o tema, Cf.: (SCHMIDT; ENGELEN, 2020).

[315] Evidentemente, Saint-Paul faz referência à entidade que lidera a fictícia Oceânia da distopia 1984, do inglês George Orwell (2019).

[316] Tradução livre. Texto original: "*When revealed preferences are tossed out the window, Big Brother enters through the back door*."

[317] Tradução livre. Texto original: "[...] *to manipulate individual preferences so as to make people feel good. This is, in fact, what Huxley's brave new world society does*."

Nesse contexto - e para terminar de testar a hipótese adotada nesta dissertação –, é necessário mergulhar no debate normativo. De maneira geral, as críticas direcionadas aos *nudges* - e aos seus fundamentos paternalistas (libertários) - permitem pensar o próprio padrão de intervenção estatal e auxiliam a compreender como buscar consensos para tornar as intervenções mais legítimas e eticamente aceitáveis. Respondê-las diretamente, portanto, é uma tarefa imprescindível em qualquer argumento em favor da utilização de *nudges*. Este será o objetivo deste Capítulo, que encerra esta dissertação.

Toda a controvérsia em torno da legitimidade dos *nudges* gerou um volume extenso de publicações a esse respeito. Abordar de maneira sistemática e exaustiva todas essas críticas é uma tarefa virtualmente impossível, considerando os objetivos desta dissertação.[318] Nesse caso, foram selecionados livros e artigos específicos – de forma inevitavelmente enviesada, como mencionado na Introdução –, que servem para ilustrar os principais e mais relevantes argumentos contrários à utilização de *nudges* que foram identificados na literatura. Há, ainda, diversas formas de classificar esses argumentos. Sunstein (2015b, 2016a, 2019a), por exemplo, costuma reunir as principais críticas aos *nudges* no guarda-chuva mais amplo da ética, dentro do qual são apresentados argumentos ligados aos seguintes pontos: a) bem-estar; b) autonomia; c) dignidade; d) manipulação; e e) intenções e habilidades dos arquitetos de escolha. Mark D. White (2016), por sua vez, divide as críticas à aplicação de *nudges* em críticas epistemológicas, éticas e práticas[319].

Neste Capítulo, o conjunto de críticas à utilização de *nudges* será dividido em dois conjuntos mais amplos: bem-estar e autonomia.[320] Parte desses argumentos se sobrepõe, já que, por vezes, uma crítica ligada à incapacidade de um *nudge* gerar bem-estar pode estar indissociavelmente ligada a uma eventual heteronomia das alterações da arquitetura de escolhas. Pode ser que os *nudges* sejam criticados, por exemplo, por gerar escolhas que não refletem as verdadeiras preferências do sujeito. Essa mesma constatação, por outro lado, pode originar críticas epistemológicas ligadas ao bem-estar: sem refletir aquelas preferências, o *nudge* é incapaz de gerar bem-estar. Como nem todas as críticas se sobrepõem, elas serão abordadas separadamente neste Capítulo.

---

[318] Em uma pesquisa feita em abril de 2021, no Scopus, maior banco de dados de literatura revisada por pares, incluindo artigos publicados a partir de 2008, data de publicação da edição original de *Nudge*, foram identificados 628 trabalhos disponíveis na base do Scopus e 90 documentos secundários, que não estavam disponíveis em tal banco de dados. A pesquisa se limitou a artigos que discutissem a ética, a aplicabilidade, a legitimidade ou a definição dos *nudges* e do paternalismo libertário, em inglês, português, francês, italiano e espanhol.

[319] As críticas práticas à utilização de *nudges*, baseadas na sua suposta ausência de efetividade, já foram discutidas no Capítulo 5. A divisão adotada nesta dissertação, de certa forma, assemelha-se à proposta de White.

[320] Embora não seja exatamente idêntica, a divisão é inspirada na afirmação de Engelen (2019) de que a maior parte das críticas remontam aos objetivos (*goals*) e aos meios (*means*) explorados pelos *nudges*.

Inicialmente, com relação ao bem-estar, serão discutidas duas questões ligadas aos fundamentos que embasam a implementação de um *nudge*: a) há mesmo evidências de que os *nudges* são necessários, em um contexto de racionalidade ecológica e de problemas estruturais?; e b) considerando o problema do conhecimento e o abandono das preferências reveladas, os *nudges* podem, de fato, gerar incremento de bem-estar?

Em seguida, serão enfrentadas críticas ligadas à autonomia, nas suas diversas acepções. As preocupações, nesse caso, são com os meios aparentemente obscuros, que são utilizados para influenciar o comportamento. Em essência, ao explorar fatores supostamente irrelevantes, os *nudges* prejudicam a autonomia e são caminhos para a coerção, a manipulação e a doutrinação? Em linha com a definição de *nudge* adotada no Capítulo 2, pretende-se avaliar se as decisões empurradas são atos substancialmente autônomos.

Ao enfrentar todas essas questões, serão apontados os cuidados necessários para evitar os *nudges* ilegítimos, o que perpassa, inclusive, às restrições e amarras institucionais inerentes ao Estado de Democrático de Direito. De maneira geral, a transparência efetiva, o escrutínio democrático e o controle jurisdicional, incluindo o controle de constitucionalidade, são salvaguardas contra as más intenções e as eventuais falhas dos arquitetos de escolha.

6.1. Correspondência entre procedimentos de tomada de decisão, estrutura psicológica e ambiente de escolhas: *nudges* como ferramentas promotoras da racionalidade ecológica

A utilização de *nudges*, como já mencionado, tem como matriz teórica o Programa de Heurísticas e Vieses e as demais teorias que vieram a ser construídas ao longo do desenvolvimento da economia comportamental. De maneira geral, essas teorias descrevem fatores supostamente irrelevantes, que podem levar as pessoas a tomarem decisões que são prejudiciais a si próprias - notadamente quando, como discutido no Capítulo 1, os mecanismos heurísticos empregados para tomar determinada decisão são ecologicamente inadequados.

Apesar disso, essa visão não é unânime. Ao longo do Capítulo 1, a apresentação dos argumentos do Programa de Heurísticas Rápidas e Frugais evidenciou essa controvérsia. Essa linha de pesquisas destaca a excelência do raciocínio heurístico, como estratégia ecologicamente racional, de modo que a sua utilização não resultaria em prejuízos. No mesmo sentido, Grüne-Yanoff e Hertwig (2016), por exemplo, enfatizam que o objetivo desse programa de pesquisas é, simplesmente, explorar os mecanismos cognitivos que um decisor que opera em condições de capacidade computacional e informação limitadas e de incerteza emprega para tomar decisões minimamente satisfatórias.

A partir dessa distinção, uma primeira linha de críticas à utilização de *nudges* - em especial, aos seus fundamentos paternalistas libertários - rechaça a ideia de falhas de mercado comportamentais. Gigerenzer (2015, p. 365), por exemplo, afirma que o paternalismo libertário culpa as pessoas por não viverem de acordo com o arquétipo do *Homo economicus*. Essa não é uma crítica aos *nudges* em si, mas às justificativas de que eles são necessários em razão de falhas de racionalidade (GIGERENZER, 2015, p. 363). Do mesmo modo, Rebonato (2014, p. 378) argumenta que, com a rejeição da ideia de preferências reveladas pelos paternalistas libertários - tópico discutido no Capítulo 2 e que será retomado na Seção 6.2 -, esse movimento tenta direcionar as pessoas às escolhas que um *Homo economicus* faria. Nessa visão, o principal argumento em favor da aplicação de *nudges* - as falhas de mercado comportamentais – seria circular, já que dependeria do próprio arquétipo que foi destronado. Naturalmente, a ênfase em como as heurísticas são bem adaptadas ao ambiente em que as escolhas são feitas é o eixo central da crítica aqui exemplificada nas obras de Gigerenzer e de Rebonato.

Essa crítica, no entanto, não sustenta que o comportamento humano, baseado largamente em heurísticas, é imune a erros. Esses erros existem, mas, para Gigerenzer (2015, p. 367), eles não seriam o resultado de uma irracionalidade sistemática. Na verdade, o psicólogo alemão afirma que há múltiplas razões alternativas que explicam o comportamento nocivo a si próprio dos indivíduos, como campanhas publicitárias multibilionárias que seduzem as pessoas a um estilo de vida não-saudável; influências de agentes privados na tomada de decisões políticas; e falhas dos sistemas educacionais no ensinamento estatístico.

A crítica exposta acima apresenta pelo menos dois questionamentos que devem ser enfrentados, para aferir a legitimidade dos *nudges* como técnica capaz de gerar bem-estar. Em um primeiro aspecto, Gigerenzer aponta questões estruturais - *e.g.*, publicidade, a captura de arquitetos de escolha e problemas educacionais - como causa dos problemas comportamentais que justificariam a utilização de *nudges*. Em outra perspectiva, Gigerenzer questiona a medida em que esses problemas são realmente comportamentais, já que considera inadequada a leitura que paternalistas libertários fazem do comportamento humano, por ser baseada em uma suposta falta de entendimento da natureza ecológica da racionalidade.

Inicialmente, a ênfase em questões estruturais como causas de problemas sociais não é exclusiva de Gigerenzer. Como mencionado anteriormente[321], Loewenstein e Chater (2017, p. 33) criticam o paternalismo assimétrico – em uma crítica que, de certa forma, também se aplica ao paternalismo libertário - por atribuir a necessidade de paternalismo a deficiências

[321] Vide nota 148 *supra*.

individuais[322]. Na sua visão, muitos problemas associados a falhas no comportamento são, na verdade, o resultado de questões estruturais. A obesidade, por exemplo, ao invés de ser o resultado de um crescimento de descontos intertemporais ou do viés do presente, teria emergido a partir de questões como mudanças de preços absolutos e relativos de alimentos frescos e empacotados, diminuição dos tempos para preparo de alimentos, maior número de lanches e aumento do tamanho das porções.

Essa crítica, no entanto, não parece justa se analisada a obra de Thaler e Sunstein com maior cuidado. Thaler e Sunstein reconhecem, em inúmeras oportunidades, que intervenções coercitivas são necessárias e recomendadas diante de problemas mais graves, o que já foi apontado no Capítulo 2. O paternalismo libertário é, afinal, uma tentativa de acordo incompletamente teorizado entre paternalistas e libertários. Não houve, porém, qualquer intenção de que toda e qualquer política pública fosse exclusivamente baseada em *nudges* e, tampouco, defendeu-se que todos os problemas do mundo estariam reduzidos a fatores supostamente irrelevantes.

Como discutido no Capítulo 5, os *nudges* devem ser vistos como uma ferramenta complementar no arcabouço do regulador. Thaler (2017), inclusive, sugere que a boa economia comportamental inclua fatores estruturais e econômicos, para além dos fatores psicológicos. Embora o viés do presente, por exemplo, possa não ser o principal elemento que explique maiores índices de obesidade, parece claro que, na presença desse viés, os fatores estruturais que Loewenstein e Chater (2017, p. 33) mencionam terão maiores efeitos. Dentro do seu papel complementar, os *nudges* podem contribuir para reduzir o impacto desse viés e os efeitos dos fatores estruturais que promovem a obesidade.

Por outro lado, as técnicas sugeridas por Gigerenzer para aprimorar a educação não são substitutas perfeitas dos *nudges*. Evidentemente, o aprimoramento da educação pode melhorar a capacidade decisória dos agentes. Apesar disso, os *nudges* e as técnicas educativas, como os *boosts*, são complementos e não alternativas excludentes (DHAMI; AL-NOWAIHI; SUNSTEIN, 2019). Nenhum defensor dos *nudges* ou paternalista libertário sugeriria que a

[322] Autores de matriz *foucaultiana* também apresentam argumentos voltados aos aspectos sociais e estruturais. Rik Peeters (2019), por exemplo, sustenta que a engenharia das escolhas, a partir de *nudges* que exploram mecanismos psicológicos subjacentes às escolhas, é uma das técnicas de poder comportamental (*behavioural power*) utilizada pelos governos. Na sua visão, essas técnicas são empregadas como meios de responsabilização, concepção que ilustra o processo de transferência de responsabilidade do Estado aos indivíduos sociais. Will Leggett (2014), por sua vez, argumenta que, ao propor *nudges*, o Estado replica o comportamento de agentes do setor privado e se baseia em uma visão sociológica rasa e despolitizada. Leggett propõe, então, a utilização de elementos dos *thinks*, expostos no Capítulo 4, juntamente com elementos de *nudges* para que o Estado promova alterações de comportamento de forma social-democrática. Além de implementar *nudges*, essa abordagem permitiria a proteção dos agentes contra mudanças de comportamento que lhes fossem prejudiciais e que alterassem a arquitetura das escolhas para atuar pelos dos indivíduos, afetando a constituição das suas subjetividades.

educação estatística fosse abandonada, em prol da utilização de *nudges*. O fato é que, como será mais bem discutido na Subseção 6.2.1, no contexto de tempo, atenção e capacidades cognitivas escassas, as pessoas nem sempre conseguem tomar decisões ativas ou deliberar sobre todo e qualquer aspecto das suas vidas. Os *nudges*, nesse caso, podem aprimorar eventuais decisões tomadas de forma automática ou não reflexiva.

Nesse contexto, argumentar que fatores estruturais fazem com que indivíduos se comportem de maneira lesiva a si próprio não é, de fato, um argumento em desfavor da utilização de *nudges* ou, mesmo, de fundamentos paternalistas libertários. Afinal, esse argumento parte da premissa equivocada de que a utilização de *nudges* é excludente de qualquer outra técnica de intervenção. A visão mais promissora, na verdade, como já argumentado no Capítulo 5, é ver os *nudges* como apenas mais uma ferramenta, que pode ter efeitos autônomos e que pode, ainda, potencializar os efeitos de outras técnicas.

Em outra perspectiva, a defesa da utilização de *nudges* e do paternalismo libertário é absolutamente compatível com a ideia de racionalidade ecológica; o seu argumento não é circular nem depende do arquétipo do *Homo economicus*. Para começar a demonstrar esse ponto, é pertinente retomar a distinção entre a racionalidade de conteúdo e a racionalidade procedimental, exposta no Capítulo 1; isto é, entre as normas que utilizamos para avaliar a deliberação de um agente das normas que guiam essa deliberação (MORTON, 2011, p. 569).

A racionalidade ecológica, compreendida em termos procedimentais, avalia em que medida os processos de tomada de decisão levam a resultados precisos, considerando o agente decisor e o ambiente em que ele está inserido. Nessa perspectiva, as normas de racionalidade de conteúdo não são definidas *a priori*, mas avaliadas à luz do contexto ecológico em que as decisões são tomadas. Nesse sentido, Schmidt (2019, p. 528) sustenta que uma decisão é procedimentalmente racional, em termos ecológicos, na medida em que houver uma boa correspondência entre os seus procedimentos de tomada de decisão, a sua estrutura psicológica e o seu ambiente de escolhas.

Essa correspondência, no entanto, pode ser desarranjada por fatores supostamente irrelevantes, que fazem com que os procedimentos utilizados para decidir sejam incompatíveis com o ambiente de escolhas. Não por outra razão, Schmidt (2019, p. 526) sugere que os decisores avaliem continuamente os seus próprios decisórios; afinal, podem surgir eventuais descompassos entre os procedimentos empregados para decidir e o ambiente em que as escolhas são feitas. Diante de tais descompassos, pode ser necessário ajustar os procedimentos empregados pelo agente, a fim de atingir a correspondência entre eles que caracteriza a racionalidade ecológica. Como mencionado no Capítulo 1, se um sujeito percebe que toma

decisões de investimento a partir de heurísticas e, com isso, tem algum prejuízo, ele pode optar ativamente por tentar seguir as normas que conduziriam à otimização da sua utilidade ou delegar essa escolha e contratar um assessor de investimentos.

Nem sempre, no entanto, o decisor tem viabilidade de alterar os seus procedimentos decisórios. Schmidt (2019, p. 528) pondera que essa alteração pode ser impossível e excessivamente custosa ou, ainda, que, embora os procedimentos heurísticos e automáticos funcionem bem em determinados ambientes, o problema pode estar naquele contexto decisório específico. De fato, a arquitetura de escolhas pode ter sido elaborada sem que as características psicológicas dos decisores fossem levadas em consideração; ou, ainda, diante do equilíbrio da pescaria, arquitetos de escolha privados podem moldar o contexto decisório de forma a evitar a correspondência entre os procedimentos decisórios dos agentes e o ambiente de escolhas e impedi-los de atingir os seus objetivos[323].

Nesse caso, diante da falta de conformidade entre processos decisórios e o ambiente em que escolhas são feitas, isto é, diante de decisões ecologicamente irracionais, ajustes podem ser necessários para permitir que o sujeito seja capaz de atingir os seus próprios objetivos, em um determinado ambiente. A definição de quais seriam esses objetivos será o objeto da Seção 6.2. De qualquer forma, em termos ecológicos, em linha com o que Schmidt (2019, p. 528) propõe, os *nudges* podem ajustar o ambiente de escolhas, para promover a racionalidade ecológica.

Alguns exemplos de Schmidt (2019, p. 530) são ilustrativos. Schmidt pondera que é comum haver um descompasso entre o que indivíduos e a sociedade querem fazer e o que eles efetivamente fazem. Nesse caso, pode ser que a arquitetura de escolhas não esteja adaptada aos procedimentos decisórios que os agentes utilizam para tomar uma determinada decisão. Schmidt menciona o caso da utilização de energia renovável na cidade de Schönau, na Alemanha. Depois de uma reunião da prefeitura municipal, um referendo alterou a regra *default* para que todas as residências aderissem automaticamente à energia renovável. Com essa alteração, 99% das residências passaram a utilizar energia renovável.

A aprovação popular da alteração da regra indica que o ambiente anterior de escolhas não era compatível com os procedimentos decisórios que as pessoas utilizam para decidir a respeito do consumo de energia elétrica. Por mais que elas desejassem utilizar energia verde,

[323] Schmidt (2019, p. 530-531) sustenta que as correspondências nocivas (*ill matches*) podem ocorrer por acaso, quando o descompasso entre os procedimentos decisórios e o ambiente surge naturalmente, ou, ainda, de forma deliberativa. No primeiro caso, como a arquitetura de escolhas não pode ser neutra, pode ser que ela direcione os agentes para uma determinada opção aleatoriamente. Por outro lado, como mencionado, pode ser que arquitetos de escolhas privados tenham incentivos para criar correspondências nocivas, a fim de auferir lucro. Em ambos os casos, a partir da utilização de *nudges*, puros e impuros, o governo pode contribuir para originar ou restaurar essa correspondência.

como a aprovação popular da medida indica, o ambiente de escolhas prévio ao *nudge*, em que se exigia escolhas ativas - em meio, eventualmente, a entraves burocráticos -, era incompatível com a forma automática e pouco deliberada pela qual as pessoas decidem nesse domínio. Na interpretação de Schmidt, esse *nudge* fortalece a racionalidade ecológica, ao tornar racional o raciocínio social, valorizado pelos habitantes de Schönau.

Schmidt (2019, p. 531-532) pondera, ainda, que os ambientes de alimentação são frequentemente incompatíveis aos processos decisórios que as pessoas empregam. Decisões alimentares são baseadas, muitas vezes, em sinais do ambiente, como cores que podem indicar que um produto saudável, o tamanho de uma embalagem, a ordem dos produtos, dentre outros fatores. O descompasso ecológico, nesse caso, faz com que as pessoas, muitas vezes, comam demais e de maneira pouco saudável.

Se o arquétipo do *Homo economicus* fosse tido como modelo de decisor, nesse caso, prossegue Schmidt, técnicas como o fornecimento de informações nutricionais relevantes facilitariam a tomada de decisões informada. Por outro lado, na perspectiva da racionalidade ecológica, em que há influência de fatores supostamente irrelevantes, como defendido nesta dissertação, seria necessário, inicialmente, compreender de maneira sistemática e com evidências como os agentes reais respondem aos sinais externos e à arquitetura de escolhas. A partir desse entendimento, *nudges* podem ser elaborados para alterar o ambiente de escolhas e promover a racionalidade ecológica.

Portanto, o argumento em favor dos *nudges* não é circular nem depende do *Homo economicus*; ele não é, ainda, incompatível com a noção de racionalidade ecológica. Na verdade, se forem efetivos, os *nudges* podem desempenhar um papel fundamental para ajudar as pessoas a ajustar os ambientes de escolha aos seus procedimentos de tomada de decisão, ao prevenir, corrigir ou contra-atacar arquiteturas de escolha que gerem correspondências nocivas (SCHMIDT, 2019, p. 532). Em suma, os *nudges* são capazes de promover a racionalidade ecológica.

6.2. Os *Nudges* e o problema do conhecimento: os julgamentos diretos e os indiretos e a análise de custos e benefícios

O objeto do Capítulo 5 foi discutir a medida em que *nudges* são efetivos como técnica de transformação do comportamento. Em outras palavras, em que medida os decisores influenciados por um *nudge* se comportam em conformidade com a opção estimulada? Uma questão distinta, no entanto, envolve a medida em que essa opção gera bem-estar aos decisores

e à sociedade. Em outras palavras, se o *nudge* for efetivo em transformar o comportamento, esse é um resultado benéfico aos decisores que são empurrados e à sociedade como um todo?

Uma das linhas mais robustas de críticas aos *nudges* lança dúvidas sobre a medida em que eles podem gerar bem-estar. O alvo central dessa crítica é o paternalismo libertário e o seu principal objetivo: melhorar a vida dos decisores como julgados por si próprios (*as judged by themselves*). De todo modo, a partir dessa questão, é possível pensar se são necessárias intervenções diante de falhas de mercado comportamentais ou se, em outra perspectiva, as pessoas conseguem decidir de forma adequada por conta própria. O objetivo desta seção é mapear essa linha de críticas e oferecer respostas, a fim de demonstrar que os *nudges* são promissores em termos de geração de bem-estar e apontar, ainda, os cenários em que esse bem-estar talvez não seja gerado.

6.2.1. A limitação das escolhas ativas: concorrência imperfeita, aprendizado e limitada largura de banda

Como mencionado nos Capítulo 2 e 3, um dos argumentos mais utilizados por Thaler e Sunstein (2003a, p. 175) em defesa do paternalismo libertário é a inevitabilidade da arquitetura de escolha*s*. A influência permanente do contexto em que escolhas são feitas nos resultados dessas mesmas decisões fez com que Thaler e Sunstein (2003a, 2003b) sugerissem que qualquer estrutura escolhida seria interventiva. O paternalismo seria, então, inevitável.

Gregory Mitchell (2005), no entanto, afirma que essa posição está equivocada. Na sua visão, a consequência lógica da constatação de que fatores normativamente irrelevantes da estrutura de escolhas influenciam as preferências não é imediatamente a inevitabilidade da manipulação das escolhas pelos planejadores centrais. Essa conclusão, para Mitchell, depende de uma premissa indispensável: a impossibilidade de os indivíduos superarem as influências irracionais que recaem sobre as suas escolhas. Técnicas de desenviesamento, já discutidas no Capítulo 4, além de estratégias que viriam a ser consideradas *boosts* com a evolução da literatura científico-comportamental, evidenciariam para Mitchell o equívoco da premissa de Sunstein e Thaler.

Nesse contexto, antes de enfrentar propriamente a crítica posta aos parâmetros de bem-estar que fundamentam os *nudges*, quando se tem em vista o problema do conhecimento, é importante levar em conta a medida em que as pessoas conseguiriam tomar decisões ativas que lhes trouxessem mais bem-estar. Em essência, a questão se resume em tentar compreender se, por mais que existam fatores supostamente irrelevantes, as pessoas conseguiriam evitá-los ou

aprimorar o seu processo decisório, para tomar melhores decisões. Afinal, se esse for o caso, os *nudges* com fundamento paternalista podem ser desnecessários.

Um primeiro ponto que pode ser alegado em favor das escolhas livres - de índole notadamente liberal - é um eventual efeito da concorrência. Sunstein (2014a, p. 93) menciona que se abordagens paternalistas enrijecem o processo competitivo o resultado pode ser problemático. No clássico argumento de Friedrich Hayek (1945), considerando a natureza dispersa do conhecimento humano, os mercados levam vantagem por agregar tanto as informações, quanto os gostos e as preferências de inúmeras pessoas. A partir do sistema de preços, Hayek argumenta que todo esse conhecimento é incorporado e, de certa forma, tornado público, o que faz com que o sistema de preços seja um mecanismo conciso e preciso de coordenação e sinalização.

Nesse caso, o mercado poderia, em tese, contribuir para minimizar os erros derivados de fatores supostamente irrelevantes. Como Sunstein (2014a, p. 92) exemplifica, "[s]e as pessoas são otimistas, e se companhias de cartão de crédito exploram esse otimismo irrealista, as pessoas vão descobrir e buscar outros fornecedores"[324]. Nessa lógica, por mais que alguns consumidores sejam enganados, em longo prazo, o equilíbrio de mercado tenderia a fazer com que os problemas fossem minimizados.

Essa não é, contudo, a realidade. Ao longo dessa dissertação, em especial nos Capítulos 2, 4 e 5, já se demonstrou em inúmeros exemplos que, em equilíbrio, diante de fatores supostamente irrelevantes, os agentes do mercado têm incentivos para tirar proveito de tais fatores para auferir lucro. O elemento essencial, aqui, é entender que as pessoas nem sempre percebem que estão sendo submetidas aos *nudges* ou aos *sludges* privados. A influência comportamental, então, não gera necessariamente uma perda de demanda por parte da empresa.

Nesse caso, em vez de contribuir para desenviesar os decisores e levá-los às opções que mais lhes aprouvessem, as empresas têm incentivos para estimular um consumo desenfreado das opções mais lucrativas. Esse fator é ainda mais relevante no caso de empresas que detêm poder de mercado e sofrem menos pressão competitiva de seus concorrentes. Neste último caso, o consumidor pode não ter sequer a opção de buscar outras opções. Portanto, por mais que os arquitetos de escolha devam se policiar para não distorcer a concorrência, ela não é suficiente para gerar decisões informadas ou que gerem maior bem-estar aos consumidores[325].

---

[324] Tradução livre. Texto original: "[*i]f people are optimistic, and if credit card companies exploit their unrealistic optimism, people will find out and seek for other providers*".

[325] No caso Eastman Kodak Co. v. Image Technical Services, Inc. 504 U.S. 451 (1992), por exemplo, - importante caso da jurisprudência dos Estados Unidos em defesa da concorrência -, a ignorância de atributos ocultos - uma

Uma outra questão envolve o aprendizado: diante de *nudges*, as pessoas perdem oportunidades de aprender com os erros que cometeriam ao agir com base em fatores supostamente irrelevantes? Sunstein (2014a, p. 94-95, 2015b, p. 436) menciona, em algumas oportunidades, que essa é uma preocupação válida, já que a tomada de decisões pode ser como um músculo, que pode atrofiar se não for utilizado de maneira adequada.

Essa preocupação a respeito da aprendizagem pode se ligar ao argumento exposto de Gigerenzer (2015) evidenciado na Seção 6.1: em vez de *nudges*, as pessoas deveriam ser educadas de forma adequada. Neste ponto, convém retomar o conceito de *boosts*[326], apresentado no Capítulo 4. Afinal, essa ferramenta tem por objetivo o aprimoramento das capacidades decisórias dos agentes, a fim de que eles possam tomar melhores decisões ativas – eventualmente, até pela exploração de processos heurísticos rápidas, frugais e inteligentes (GRÜNE-YANOFF; HERTWIG, 2016; HERTWIG; GRÜNE-YANOFF, 2017).

Inicialmente, quanto aos potenciais erros pedagógicos que os *nudges* evitariam, convém reiterar que a revisão de literatura apresentada no Capítulo 5 indica que os efeitos de *nudges* são modestos, embora sejam significativos. Se as pessoas erram e aprendem com os erros, esse processo continuará a acontecer na presença de um *nudge* e os seus efeitos apenas serão mitigados, o que pode gerar ganhos de bem-estar àqueles que deixam de errar. Como Conly (2013) argumenta, ainda, a forma como as pessoas reagem a erros causados por vieses cognitivos - ou fatores supostamente irrelevantes - não é racional. O excesso de otimismo, ao mesmo tempo, cria uma dificuldade adicional ao aprendizado com os erros. Nesse caso, simplesmente deixar que as pessoas errem e esperar que elas aprendam com os erros não parece uma estratégia normativamente atraente.

---

das falhas de mercado comportamentais indicadas no Capítulo 2 – foi um dos elementos essenciais para explicar os potenciais prejuízos que a Kodak poderia causar ao mercado a partir da sua conduta – exclusividade no mercado de cartuchos. O cerne da questão envolvia a ignorância, pelos consumidores, dos custos futuros que teriam com suprimentos para o produto principal (*atributos ocultos*); isto é, ao pensar em comprar uma impressora, apenas o preço da própria impressora era levado em consideração pelos consumidores e o valor dos cartuchos era ignorado. Por isso, os consumidores eram menos responsivos à alteração nos preços dos cartuchos (*atributos ocultos*). Diante disso e do custo significativo para adquirir uma impressora de outra marca (efeito *lock-in*), a *Kodak* acabaria conseguindo impor preços mais altos no mercado de cartuchos, ainda que não tivesse poder de mercado no mercado de impressoras. Cf., nesse sentido: (TOR, 2014).

[326] Uma ressalva, neste ponto, é necessária. Hertwig e Grüne-Yanoff (2017, p. 982) ressaltam expressamente que os *boosts* não equivalem à educação escolar, embora alguns deles possam ser incluídos nos currículos das escolas. Algumas diferenças entre as estratégias são marcantes. Em primeiro lugar, o objetivo dos *boosts* é aprimorar as competências decisórias em áreas que normalmente não são parte do currículo escolar, como decisões financeiras, avaliação de risco, decisões alimentares saudáveis, decisões médicas informadas e a autorregulação efetiva. *Boosts*, por outro lado, assim como *nudges*, são informados por evidências científicas, o que nem sempre é exigido no currículo escolar. Além disso, os *boosts* buscam aprimorar ou desenvolver novas competências em condições de tempo e recursos limitados, normalmente para pessoas que não podem mais ser submetidas a um período adicional nas escolas. Diferentemente da educação, ainda, os *boosts* não têm um enfoque na informação por si só, mas trabalham em cima de competências motivacionais e decisórias, como procedimentos decisórios, heurísticas e habilidades para implementação de objetivos.

Em outra perspectiva, por mais que já se tenha argumentado ao longo desta dissertação que *nudges* e *boosts* não são excludentes, mas complementares, esse argumento também deve ser analisado no contexto da geração de bem-estar. Em tese, se os *nudges* forem plenamente substituíveis por *boosts*, não haveria razões para se cogitar a elaboração de *nudges*, uma vez que são mais intrusivos e podem esbarrar em considerações epistemológicas e éticas, principalmente. A realidade, porém, parece indicar que nem sempre essa substituição será possível.

Evidentemente, políticas que buscam aprimorar a capacidade decisória dos indivíduos devem ser incentivadas. Assim, técnicas que priorizem a "alfabetização" estatística, a elaboração de árvores de decisão rápidas e frugais e outros processos heurísticos ou, ainda, o ensino de conhecimentos elementares importantes devem, certamente, ser incentivados. Apesar disso, por ser uma estratégia ainda recente, a medida em que os *boosts* são efetivos é ainda incerta. No Capítulo 5, embora não se tenha feito uma revisão completa da literatura nesse sentido, os exemplos que testaram a efetividade de *boosts* encontraram efeitos ambíguos[327].

Nem mesmo os proponentes dos *boosts*, no entanto, sugeriam que as abordagens fossem intrinsecamente excludentes. Nesse sentido, Hertwig e Grüne-Yanoff (2017, p. 983) afirmam expressamente que os *boosts* e os *nudges* não são substitutos perfeitos e que, em alguns domínios, tanto os *nudges*, quanto os *boosts* podem ser utilizados. Notadamente, os autores exemplificam, no âmbito das escolhas alimentícias, das decisões financeiras e de problemas ligados ao autocontrole. Mais do que isso, Hertwig e Grüne-Yanoff ressaltam que o objetivo deles não era defender os *boosts* em detrimento dos *nudges*, mas ressaltar a necessidade de uma análise dos objetivos e circunstâncias que fundamentam a política, a fim de definir qual das duas estratégias é a mais adequada - o que pode incluir, evidentemente, uma combinação das duas.[328]

A definição de qual intervenção deve ser utilizada, ao fim e ao cabo, é essencialmente empírica. Ainda que essas evidências ainda sejam restritas, alguns princípios podem auxiliar a identificar esses objetivos e circunstâncias. Nesse sentido, Hertwig (2017) desenvolveu seis regras para facilitar a análise de custos e benefícios entre *nudges* e *boosts* e facilitar a escolha entre essas intervenções.

---

[327] Cf. nota 308 *supra*.

[328] No mesmo sentido, Hertwig (2017, p. 158) sustenta que em muitos domínios, *boosts* e *nudges* podem se complementar e amplificar os efeitos comportamentais desejados. No caso do clássico exemplo da reestruturação dos alimentos de uma cafeteria, o fornecimento de regras simples e inteligentes para escolha de alimentos (*boost*) pode corroborar e ampliar os efeitos buscados.

Inicialmente, se as pessoas não tiverem a habilidade cognitiva ou a motivação necessárias para adquirir novas habilidades ou competências, os *nudges* serão provavelmente a intervenção mais eficiente. Por outro lado, se os *policy-makers* estão incertos a respeito dos objetivos das pessoas, se houver heterogeneidade entre esses objetivos ou se um indivíduo tiver objetivos conflitantes, os *boosts* serão a intervenção menos propensa aos erros. Na sua terceira regra, Hertwig argumenta que, se um *nudge* não puder ser transparente ou tiver que ser invisível ao decisor para funcionar, ele não será facilmente reversível e deverá ser considerado paternalista. A quarta regra sugere que, se os governos deixarem de agir de forma benevolente ou permitirem que o setor privado crie arquiteturas de escolha tóxicas, os *boosts* serão meios de fornecer maior proteção aos indivíduos. Além disso, Hertwig também sugere que, se os arquitetos de escolha quiserem promover comportamentos generalizáveis e de longa duração, os *boosts* parecem ser a estratégia mais adequada a ser utilizada.[329] Por fim, a última regra descrita por Hertwig sugere que, se houver um perigo substancial de consequências indesejadas, não antecipadas e imprevisíveis a partir de um *nudge* ou de um *boost*, a estratégia contrária deve ser considerada.

Tendo em vista todas essas regras – que funcionam como um *boost* para o arquiteto de escolha -, é preciso reconhecer que, em algumas circunstâncias, a escolha ativa realmente será a melhor opção aos agentes. Os *boosts*, nesse contexto, serão extremamente importantes. Como descrito por Sunstein (2015a), há inúmeras situações em que a melhor opção é a promoção de escolhas ativas. Evitam-se, assim, problemas relacionados à inércia, aos arquitetos de escolha ruins ou mal-intencionados, às preferências que podem se modificar ao longo do tempo – e deixar de ser compatíveis com um *default*, por exemplo -, à heterogeneidade da população atingida, à autonomia e ao aprendizado. Nesses casos, se a escolha ativa é precedida de *boosts*, é provável que melhores resultados sejam atingidos.

Por outro lado, Sunstein (2015a, p. 114-115) afirma que não se pode exigir que as pessoas optem por escolher ativamente em todas as circunstâncias. Uma imposição como essa conduziria a uma espécie de "paternalismo que exige escolha" ("*choice-requiring paternalism*"), o que gera questionamentos em termos de autonomia e de bem-estar. Algumas pessoas, afinal, podem optar por não escolher. Nesses contextos, exigir uma escolha pode ser uma intervenção coercitiva e prejudicial em termos de bem-estar, especialmente se a escolha ativa não trouxer qualquer dos benefícios mencionados no último parágrafo. Os *nudges*, então,

[329] A discussão empreendida no Capítulo 5, no entanto, mostra que os *nudges* podem gerar efeitos persistentes, especialmente se gerarem efeitos de transbordamento positivos, como o caso dos relatórios de energia doméstica da Opower.

podem ser exigidos como opção voluntária de indivíduos que não querem decidir ativamente e que têm, nesta escolha, a opção que mais lhes traz utilidade. Seria esse o caso, por exemplo, de um indivíduo que escolhe aderir a uma determinada regra *default*, que presume o seu consentimento em decisões subsequentes.

Também é preciso considerar que existem hipóteses em que as pessoas simplesmente não se importam em escolher ativamente. Ainda que bons resultados lhes interessem, essas são situações complexas, com variadas opções e *feedbacks* separados temporalmente do momento em que a escolha é feita. Nesse caso, ainda que obrigadas a escolher, é improvável que as pessoas atinjam bons resultados. Um exemplo claro são as múltiplas e variadas opções que um indivíduo pode escolher como plano de previdência para o futuro. O exemplo da previdência da Suécia, discutido no Capítulo 5, evidencia essa afirmação[330].

O problema se agrava na medida em que as pessoas têm limitada largura de banda (*bandwith*) - conceito que mensura a capacidade cognitiva[331] e o controle executivo[332] das pessoas (MULLAINATHAN; SHAFIR, 2016, p. 64). Nesse contexto, exigir que elas prestem atenção em questões que não são do seu interesse, desvia recursos cognitivos escassos de outros empreendimentos da vida (SUNSTEIN, 2015a, p. 146). Por exemplo, se o sujeito é obrigado a escolher o plano de previdência que deseja, dentre uma miríade de opções do mercado, sem ter uma opção *default*, é muito provável que opte por um plano ruim; o tempo gasto com essa escolha inadequada poderia ter sido utilizado para decisões em que a escolha ativa lhe fosse mais aprazível.

Em face de escolhas nitidamente ruins ou, mesmo, da ausência de escolhas, deixar de utilizar um *nudge* pode gerar nítido prejuízo em termos de bem-estar. A autonomia do indivíduo também fica prejudicada, já que não será possível uma dedicação maior às decisões que são mais decisivas para seus projetos pessoais, em que a sua autoridade lhe é mais importante. Nesse sentido, também é possível que exigir escolhas ativa em todas as circunstâncias traga prejuízos em termos de dignidade, já que o indivíduo não consegue se empenhar em maximizar os fins que julga mais nobres.

---

[330] Cf., nesse sentido, mais uma vez: (CRONQVIST; THALER; YU, 2018; THALER; SUNSTEIN, 2021).

[331] A capacidade cognitiva está ligada aos "mecanismos psicológicos subjacentes à nossa capacidade de resolver problemas, guardar informações, raciocinar de maneira lógica e assim por diante. Talvez o mecanismo mais proeminente dessa categoria seja a inteligência fluida, a capacidade de pensar e raciocinar de maneira abstrata e resolver problemas independentemente de qualquer aprendizado ou experiência específica" (MULLAINATHAN; SHAFIR, 2016, p. 72-73).

[332] O controle executivo "é subjacente à nossa capacidade de administrar nossas capacidades cognitivas, incluindo planejamento, atenção, iniciar e inibir ações e controlar impulsos. Em grande parte, como um processador central, o controle executivo é essencial para nossa capacidade de funcionar bem" (MULLAINATHAN; SHAFIR, 2016, p. 73).

Outro aspecto relevante, no debate em torno das escolhas ativas, é o fato de que decidir é mais custoso cognitivamente para pessoas que se encontram em situação de escassez. Na metáfora utilizada por Anandi Mani, Sendhil Mullainathan, Eldar Shafir e Jiaying Zhao (2013, p. 980), é possível que uma maior "tributação cognitiva" (*cognitive taxes*) incida sobre as pessoas mais pobres, no sentido de que a situação de pobreza, por si só, poderia gerar padrões comportamentais perniciosos, capazes de reforçar a própria condição de pobreza[333].

Mullainathan e Shafir (2016) argumentam que a situação de escassez de qualquer tipo - incluindo tempo e recursos financeiros – gera um dividendo de foco, ao capturar a atenção das pessoas inconscientemente. Por mais que essa captura possa gerar efeitos benéficos, ao gerar resultados positivos em situações de escassez, ela também faz com que as pessoas entrem em túneis, isto é, concentrem a sua atenção na administração da escassez, em detrimento de outras coisas que podem ser mais importantes. Daí a ideia de que a escassez tributa a limitada largura de banda, de modo que uma menor capacidade cognitiva e um menor controle executivo disponíveis para outras tarefas importantes[334].

Nesse contexto, direcionar as pessoas para escolhas que talvez elas não fariam e que lhes trazem bem-estar permite que elas se empenhem nas decisões que mais lhes importam e nas quais o seu esforço será mais bem recompensado. Reduzir o peso do fardo de escolher em todas as circunstâncias pode, ainda, ter efeitos redistributivos, na medida em que deixa de agravar as consequências cognitivas da escassez. Dessa forma, os *boosts* e as escolhas ativas não podem nem devem substituir os *nudges* que, ao menos em alguma medida, devem ser aplicados. Afinal, essa pode ser uma medida de redução da tributação cognitiva, extensamente demonstrada por Mullainathan, Shafir e outros autores dessa importante linha de pesquisas[335].

---

[333] Nesse sentido, Anuj K. Shah, Mullainathan e Shafir (2012) documentaram a prevalência, em situações ligadas à pobreza, de comportamentos econômicos prejudiciais em longo prazo, como baixa poupança, excesso de empréstimos, muitas vezes com juros elevados, e falhas para se inscrever em programas assistenciais. Assim, exigir que escolhas sejam feitas em todas as circunstâncias pode, além de tudo, gerar consequências regressivas, já que isso tende a impor maiores custos cognitivos a pessoas mais pobres, cuja capacidade cognitiva se encontra ocupada em resolver sua própria situação de escassez.

[334] Evidentemente, as evidências de Mullainathan e Shafir não sugerem que pessoas em situação de escassez são piores, despreparadas ou descuidadas. A redução de largura de banda não se refere à "capacidade inerente de uma pessoa, mas o quanto dessa capacidade está disponível para uso no momento" (MULLAINATHAN; SHAFIR, 2016, p. 72). Nas palavras de Mullainathan e Shafir (2016, p. 98): "O problema não é a pessoa, e sim o contexto da escassez".

[335] Para uma revisão bibliográfica, Cf.: (SILVA *et al.*, 2018).

6.2.2. O problema do conhecimento e a mensuração de bem-estar: crítica ao critério como julgado por si próprios

Ao apresentar o paternalismo libertário, no Capítulo 2, demonstrou-se que Thaler e Sunstein (2009, p. 5, 2021, p. 7), com esse movimento, pretendiam defender políticas não coercitivas que almejassem gerar ganhos de bem-estar às pessoas segundo os seus próprios critérios (*as judged by themselves*). Afinal, o ponto de partida do paternalismo libertário é a afirmação de que é falsa a premissa antipaternalista de que as pessoas fazem escolhas que favorecem o seu melhor interesse ou que, pelo menos, são melhores, segundo as suas próprias lentes (THALER; SUNSTEIN, 2003b, 2003a). Os fatores supostamente irrelevantes, de fato, lançam dúvidas sobre essa premissa, como amplamente discutido ao longo desta dissertação.

Apesar disso, na perspectiva da economia comportamental e dos fatores supostamente irrelevantes, como também mencionado no Capítulo 2, a preferência revelada pela escolha do indivíduo nem sempre evidenciará a opção que maximiza o seu bem-estar. Esse é, afinal, o ponto de partida da ideia de que existem falhas de mercado comportamentais que podem exigir a regulação, ainda que sutil, do comportamento privado, a fim de gerar bem-estar aos próprios indivíduos que tomam decisões.

Nesse caso, para operacionalizar um paternalismo leve – como o paternalismo libertário – é necessário especificar o que deve ser considerado um erro ou percepção equivocada e, mais ainda, como os resultados de políticas paternalistas devem ser julgados, considerando que esses erros e percepções equivocadas existem (DHAMI, 2016, p. 1581). Diante de um descompasso entre as preferências reveladas e as chamadas preferências normativas, isto é, aquelas que realmente trazem bem-estar ao decisor, o paternalista libertário - assim como o proponente de outras espécies de paternalismo suave - deve especificar qual é o seu critério de bem-estar. Thaler e Sunstein, no entanto, não fazem isso com clareza ao propor o paternalismo libertário.

Daí decorre um primeiro conjunto de críticas aos *nudges* paternalistas libertários, que retoma o clássico problema do conhecimento: se as escolhas deixam de ser um guia às opções que beneficiam um agente, como o arquiteto de escolhas pode saber quais são essas opções? Em outras palavras, como é possível definir a opção que deve ser estimulada por um *nudge* se

as preferências das pessoas não são reveladas pelas escolhas e podem, inclusive, ser construídas[336] pelas próprias tarefas de julgamento[337]?

Esse é o cerne do que Sunstein (2014a) denomina de "O Argumento Epistêmico contra à utilização de *nudges*". A rigor, o argumento remonta a debates filosóficos inesgotáveis e muito anteriores ao surgimento do paternalismo libertário e dos *nudges*[338].

O clássico argumento de Friedrich Hayek (2014, p. 385), em favor do liberalismo, por exemplo, sustenta que: "[...] a consciência de nossa ignorância irremediável da maior parte do que é conhecido por alguém é a base principal do argumento pela liberdade".[339] Em um sentido semelhante, Mill, como indicado na introdução do Capítulo 2, justifica o princípio do dano com base na ideia de que o indivíduo é a pessoa mais interessada no seu próprio bem-estar e é quem tem o conhecimento mais profundo sobre os seus próprios sentimentos e as suas próprias circunstâncias. Mergulhar nesse profundo, antiquíssimo e interminável debate filosófico ultrapassa os objetivos dessa dissertação. O que convém ressaltar, aqui, é que inúmeros autores vêm criticando o paternalismo libertário e a implementação de *nudges* com base em argumentos que remontam às preocupações de Mill e de Hayek.

Douglas Glen Whitman e Mario J. Rizzo (2011, 2015), por exemplo, apresentam uma crítica elaborada aos paternalismos comportamentais, sob o argumento de que os seus fundamentos de bem-estar são falhos. Whitman e Rizzo apontam, a princípio, que critério normativo tido como referência por autores como Thaler e Sunstein é o modelo racional neoclássico - que, contudo, é paradoxalmente desprezado pela análise positiva posta pela economia comportamental, tema que começou a ser discutido na Seção 6.1 acima.

---

[336] Cf., nesse sentido: (SLOVIC, 1995).

[337] Em 21 de outubro de 2020, em um evento promovido pelo Grupo de Estudos em Economia Comportamental da Universidade de São Paulo, tive a oportunidade de perguntar para Sunstein (FEAST..., 2021) como os arquitetos de escolha poderiam conhecer as preferências das pessoas para direcioná-las para a escolha correta. A resposta foi simples: pergunte para elas. Mais uma vez, Sunstein utilizou um GPS como exemplo para o seu argumento. Ele reconhece que, no caso de governos, fazer essa pergunta é mais difícil, mas, geralmente, afirma ser possível aprender o que a maioria das pessoas deseja fazer. Para muitos casos, essa resposta seria muito clara para o autor. Por exemplo, os motoristas preferem não morrer. Cf.: https://youtu.be/Q_6HSdcUlkw?t=2600. Acesso em: 13 ago. 2021.

[338] O debate remonta, inclusive, a uma discussão do utilitarismo de preferências, a respeito de como fazer comparações interpessoais, já que as pessoas podem decidir e valorizar opções distintas de maneira diversa. Como Walter Sinnott-Armstrong (2021) explica no verbete a respeito do consequencialismo na Enciclopédia de Filosofia de Stanford: "Se queremos saber o que uma pessoa prefere, podemos perguntar o que essa pessoa escolheria em conflitos. Não podemos, contudo, utilizar o mesmo método para determinar se a preferência de uma pessoa é mais forte ou mais fraca do que a preferência de outra pessoa, já que essas diferentes pessoas podem escolher de forma distinta em conflitos decisivos. [...] Para determinar se é certo fazer ou não fazer A, precisamos ser capazes de comparar as forças das preferências de Jones e de Smith (ou as quantias de prazer que cada um receberia em seus resultados preferidos) a fim de determinar a medida em que fazer ou não fazer A seria melhor no geral." (tradução livre).

[339] Tradução livre. Texto original: "[...] *the awareness of our irremediable ignorance of most of what is known to somebody is the chief basis of the argument for liberty*".

Além disso, Whitman e Rizzo (2015) sustentam que os paternalismos comportamentais são marcados por uma falácia *non sequitur*[340] no seu cerne: se, como a economia comportamental demonstra, existem evidências de que um agente pode ter preferências tanto pelo conjunto X, quanto pelo conjunto Y, não há base analítica para designar X ou Y como o verdadeiro conjunto de preferências do agente.

Além de epistemológico - já que a inconsistência das escolhas criam o problema de conhecimento para o arquiteto de escolhas, que não terá meios para conhecer as preferências das pessoas -, para Whitman e Rizzo (2015), há, ainda, um desafio metafísico. Na realidade, um agente pode não ter verdadeiras preferências, isto é, preferências que satisfaçam os axiomas neoclássicos da escolha racional e não sejam apenas evocadas ao fazer escolhas. Se essas preferências não existem, acrescentam Whitman e Rizzo, nem mesmo um planejador onisciente poderia conhecê-las e, assim, não haveria bases para que um arquiteto de escolhas apurasse a medida em que uma determinada intervenção paternalista suave gera bem-estar.

Alguns exemplos citados por Whitman e Rizzo (2015) ajudam a ilustrar o seu argumento. No caso de descontos hiperbólicos ou quase hiperbólicos, por exemplo, a prescrição, em termos de política pública do paternalismo libertário, é a de favorecer as taxas de desconto dos agentes mais pacientes de forma consistente; não se favorece, por exemplo, decisões mais voltadas ao presente – menos pacientes – ou taxas de desconto intermediárias. No caso das lacunas de empatia quente-frio, por sua vez, as prescrições do paternalismo libertário favorecem que o sujeito se comporte como se estivesse em estados emocionais frios, em vez de quentes. Apesar disso, os fundamentos de bem-estar do paternalismo libertário não justificariam essas escolhas de maneira adequada, na visão de Whitman e Rizzo.

Inúmeros outros exemplos podem ser extraídos da literatura para expressar argumentos semelhantes.

Gregory Mitchell (2005, p. 1267), por exemplo, afirma que, quando uma definição ordinal de utilidade ou de bem-estar individual é utilizada, terceiros não conseguiriam, por definição, fazer julgamentos a respeito da utilidade de outra pessoa. Como a ordem de preferências é puramente subjetiva, sem que haja uma meta objetiva ou implícita, essa tarefa seria simplesmente impossível. Na visão de Mitchell (2005, p. 1269), essas dificuldades levariam os paternalistas libertários à imposição de algumas concepções de bem-estar às pessoas irracionais, o que geraria objeções às políticas propostas.

[340] Em uma falácia *non sequitur*, falta conexão entre as conclusões apresentadas e os argumentos que as sustentam.

White (2010, p. 207), por sua vez, afirma que o paternalista libertário não teria meios para apurar as preferências informadas ou racionais de um indivíduo, já que não pode se valer de suas escolhas; as preferências utilizadas como medidas de bem-estar pelos paternalistas libertários, nesse caso, seriam artificiais e construídas arbitrariamente por um terceiro. Na visão kantiana de White (2010, p. 209-210), as preferências ou o bem-estar não explicam completamente as escolhas dos indivíduos, que podem se mover por aspectos que podem lhes prejudicar, como determinados princípios. Assim, ninguém poderia saber se uma escolha é realmente ruim a não ser o próprio decisor, razão pela qual White (2010, p. 211) sustenta que: "[o] verdadeiro paternalismo no 'paternalismo libertário' consiste em substituir os fins dos agentes que são 'empurrados' pelos fins dos próprios elaboradores das políticas públicas"[341].

Em um argumento anti-anti-antipaternalista, isto é, em desfavor da postura crítica ao antipaternalismo de autores como Thaler e Sunstein, Claire Hill (2007) questiona como os paternalistas poderiam saber o que realmente as pessoas querem. Não haveria, na sua visão, nenhuma razão apriorística para privilegiar o futuro em detrimento do presente, por exemplo. Mais do que isso, se as preferências são construídas, as coisas que as pessoas realmente querem frequentemente não existiriam *a priori*, de maneira não mediada por alguma escolha. A postura paternalista, então, não seria consistente com o libertarianismo, embora Hill repute corretas as direções em termos de política pública propostas pelos proponentes dos paternalismos comportamentais.

Grüne-Yanoff e Hertwig (2016) também sustentam que, como preferências construídas são medidas de forma inerentemente reativa e sujeita a vieses cognitivos, os seus resultados são, em certa medida, arbitrários. Em outras palavras, pode-se chegar a diferentes preferências a depender do contexto e a da estrutura em que a avaliação é feita. Apesar disso, para que os *nudges* tenham legitimidade, o arquiteto de escolhas precisa saber se as pessoas envolvidas na respectiva política almejam, de fato, o objetivo ao qual estão sendo direcionadas. Como essa avaliação é arbitrária, Grüne-Yanoff e Hertwig (2016) chegam à conclusão de que os *nudges* não seriam coerentes com a teoria que lhes é subjacente.

Rebonato (2014, p. 378), por sua vez, afirma que o paternalismo libertário adota a visão de utilidade das preferências-informadas. Nessa perspectiva, não há incremento de utilidade pela satisfação de quaisquer preferências, mas apenas daquelas que são consideradas racionais ou informadas. Como observar as escolhas que são feitas não é uma opção para definir as preferências de um indivíduo, Rebonato (2014, p. 378) indica dois problemas que deveriam ser

[341] Tradução livre. Texto original: "*[t]he true paternalism in 'libertarian paternalism' consists of substituting the policy makers' own ends for those of the agents being 'nudged'*".

resolvidos para que o paternalismo libertário possa cumprir seus objetivos: (a) determinar qual é a utilidade oculta que será maximizada; e (b) como ter acesso a essa utilidade. Para Rebonato, o paternalismo libertário resolveria essas questões recorrendo ao padrão de escolhas racionais. Em outras palavras, os *nudges* tentariam direcionar as pessoas para as escolhas que seriam feitas se o padrão de racionalidade econômica fosse seguido.

No mesmo sentido, Joshua D. Wright e Douglas H. Ginsburg (2015, p. 29–30) argumentam que, quando os "*behaviorists*" descartam as preferências reveladas como expressões das verdadeiras preferências, perdem um princípio coerente para identificar escolhas que maximizam bem-estar. Sendo assim, para prescrever determinado rumo de ação, podem basear-se apenas no que esperam que um indivíduo racional faria ou deveria fazer. Não se cumpriria, dessa forma, o objetivo de maximizar o bem-estar econômico dos indivíduos segundo seu próprio julgamento, mas haveria a imposição de escolhas consideradas corretas por um terceiro.

Sem a pretensão de exaurir as críticas aos *nudges* ligadas ao bem-estar, o conjunto de críticas citado acima parte, em geral, de uma mesma preocupação: ao destronar o critério de preferências reveladas pelas escolhas, o paternalismo libertário e os demais paternalismos comportamentais careceriam de fundamentos normativos? Os nudges são incoerentes com a teoria que lhes é subjacente? Essas questões serão enfrentadas nas próximas subseções.

6.2.3. Para onde empurrar: como o arquiteto de escolhas pode buscar consensos?

Se as preferências reveladas não podem ser usadas para orientar os arquitetos de escolha a respeito das opções que trazem maior bem-estar aos decisores, como aferir se um determinado *nudge* influencia os decisores a adotarem melhores decisões? Thaler e Sunstein (2021, p. 332) afirmam, como já mencionado, que a sua estrela guia é o bem-estar humano; se, no entanto, o objetivo é gerar bem-estar individual, como o arquiteto de escolhas deve se guiar?

6.2.3.1. Na prática, um paternalismo suave não criaria problemas reais: argumento fraco em favor dos ganhos de bem-estar dos nudges

Ao discutir essas questões, um primeiro ponto importante é ter atenção ao alerta constante de Sunstein (2015b, p. 424, 2016a, p. 26, 2019a, p. 117), para que, nas discussões de questões éticas ligadas às políticas públicas, os princípios sejam analisados à luz de práticas

concretas. A armadilha da abstração, nas palavras de Sunstein, pode criar bastante confusão quando políticas regulatórias estão sendo avaliadas. Na sua visão, a avaliação ética dos *nudges*, incluindo em termos de bem-estar, deve ser feita à luz do caso concreto e não do seu status como *nudge*.

Nesse caso, evitar a armadilha da abstração significa, em primeiro lugar, avaliar os impactos reais do problema do conhecimento na prática de políticas públicas, com o foco especial na situação do Brasil. A análise do cenário normativo brasileiro evidencia uma miríade de normas com fundamento nitidamente paternalista[342]. Alguns exemplos ilustram essa afirmação. Empregados têm restrições para movimentar as quantias que empregadores depositam na sua conta do Fundo de Garantia do Tempo de Serviço (FGTS)[343]. O Código de Trânsito exige que motociclistas utilizem capacetes de segurança, com viseira ou óculos protetores, além de vestuário de proteção[344]. O Código Civil impede, via de regra, que as pessoas renunciem ao exercício dos direitos da personalidade[345].

Se medidas de cunho paternalista, ainda que coercitivas, são aplicadas em vários contextos, sem maiores questionamentos - e não se pretende, nesta análise, questionar o mérito de tais medidas -, parece improvável que medidas paternalistas suaves seriam questionáveis. Na prática, normas paternalistas surgem e permanecem válidas ainda que o regulador não tenha meios precisos de apurar se aquelas regras favorecem o bem-estar individual das pessoas.

Além disso, na prática, os *nudges* têm em si um mecanismo de autocontenção, que limita a sua possibilidade de gerar prejuízos individuais em termos de bem-estar.

No caso de uma medida coercitiva, o problema do conhecimento gera um desafio maior já que as pessoas cujo bem-estar individual fosse ampliado na direção oposta à da regra arcariam com prejuízos. Considere, por exemplo, uma possível proibição do consumo de bebidas alcóolicas. Em qualquer hipótese, o sujeito cujo bem-estar é ampliado com o consumo de bebidas teria grande probabilidade de arcar com perdas de bem-estar. Se ele cumprisse a

---

[342] Em análise semelhante do cenário de Portugal, Saraiva (2020, p. 37) pondera que: "[e]m Portugal, não se adivinham tantas reticências ou dificuldades em justificar medidas paternalistas, em especial moles, visto que social, cultural, política ou juridicamente o padrão tolerado de actuação do Estado assume-se menos libertário".

[343] O art. 20 da Lei n.º 8.036, de 11 de maio de 1990, prevê as situações em que a conta vinculada do trabalhador no FGTS poderá ser movimentada (BRASIL, 1990).

[344] Essa regra está prevista no art. 54, da Lei n.º 9.503, de 23 de setembro de 1997 - o Código de Trânsito.

[345] O art. 11 do Código Civil prevê que: "[c]om exceção dos casos previstos em lei, os direitos da personalidade são intransmissíveis e irrenunciáveis, não podendo o seu exercício sofrer limitação voluntária" (BRASIL, 2022). Stancioli (2017, p. 133) critica essa redação e afirma que: "[o] Código Civil Brasileiro agrilhoou as pessoas naturais a seus direitos". Nesse contexto, como relata Stancioli (2017, p. 134), o Professor João Baptista Villela, em um esforço hermenêutico, propôs que esse dispositivo fosse lido como se o exercício dos direitos da personalidade pudesse sofrer limitações voluntárias, desde que não de forma permanente ou geral. Na visão de Stancioli (2017, p. 142), a renúncia ao exercício dos direitos da personalidade "é a mais perfeita expressão do direito ao livre desenvolvimento da personalidade".

regra e deixasse de consumir bebidas alcóolicas, haveria perdas de bem-estar na medida em que ele deixaria de escolher da forma que lhe seria mais aprazível. Em caso de descumprimento, por outro lado, uma eventual sanção que lhe fosse aplicada faria com que ele arcasse com prejuízos de toda forma.

O paternalismo suave, por outro lado, não é coercitivo. Um *nudge* que estimula o consumo de alimentos e produtos saudáveis ou um *sludge* que desestimula o consumo de álcool, por exemplo, não impedem, em princípio, que bebidas alcóolicas sejam consumidas. Como discutido no Capítulo 5, por diversas razões, os *nudges* frequentemente falham. De acordo com Sunstein (2019a), uma dessas razões é a existência de fortes preferências antecedentes. Em outras palavras, embora o arquiteto de escolhas não saiba com precisão quais são as preferências dos decisores, o fato de um *nudge* falhar pode indicar que os decisores não gostam da direção à qual eles foram empurrados[346]. Nesse caso, "o fato de que as pessoas podem rejeitá-lo [o nudge] é uma valiosa salvaguarda" (SUNSTEIN, 2019a, p. 105)[347]. Afinal, essa salvaguarda evitaria os prejuízos de bem-estar, que seriam inevitavelmente trazidos por uma medida coercitiva.

Ao mesmo tempo, ao longo dos últimos dez anos, uma série de experimentos vem demonstrando que *nudges* parecem contar com aprovação popular, embora ela não seja indiscriminada e alguns tipos de *nudges* específicos gerem preocupações por parte da população. Em geral, são realizados questionários com pessoas de diferentes países e que parecem confirmar a ideia de que não há tanta rejeição à aplicação de *nudges* e mesmo técnicas tradicionais de comando e controle parecem contar com aprovação popular em diferentes países. Alguns exemplos auxiliam a compreender esses dados[348].

William Hagman e outros autores (2015), a princípio, compararam a aprovação de *nudges* nos Estados Unidos e na Suécia. De maneira geral, os autores identificaram um alto nível de aprovação para *nudges* em ambos os países - variando entre 54% para o *nudge* com menos apoio popular e 85,5% para os mais aprovados. De todo modo, foi identificada bastante variabilidade e os *nudges* de caráter pró-social contaram com menor aprovação do que aqueles

[346] A partir dessa constatação, Tor (2020) pondera que a efetividade de um *nudge* deve ser avaliada à luz do seu potencial de gerar bem-estar aos decisores. Na visão de Tor, embora Sunstein (2019a) tenha indicado as diferentes razões pelas quais *nudges* falham, ele não explicou como determinar a causa de uma falha específica. Sem tal determinação, diante de um *nudge* que falha, o arquiteto de escolhas não tem meios se a falha deve ser apoiada ou corrigida. Também não tem meios de saber se o *nudge* que falhou deve ser substituído por um *nudge* melhor ou por uma técnica de intervenção mais assertiva.

[347] Tradução livre. Texto original: "[...] *the fact that people can reject it is a valuable safeguard*".

[348] Para mais exemplos, com o controle de diferentes variáveis, Cf.: (JUNG; MELLERS, 2016; PETRESCU *et al.*, 2016; REISCH; SUNSTEIN; GWOZDZ, 2017; REYNOLDS *et al.*, 2019; TANNENBAUM; FOX; ROGERS, 2017; WACHNER; ADRIAANSE; DE RIDDER, 2020).

focados no bem-estar privado. Hagman e os seus coautores identificaram, ainda, que uma grande proporção de participantes via os *nudges* como intrusivos no que diz respeito à liberdade.

Reisch e Sunstein (2016), por sua vez, avaliaram a aprovação popular de *nudges* em seis países - Dinamarca, França, Alemanha, Hungria, Itália e Reino Unido. A maior parte das intervenções avaliadas - 12 de 15 - contou com um apoio amplo dos participantes do experimento. Três *nudges* específicos, no entanto, foram amplamente reprovados: uma regra *default* direcionada às doações à caridade; uma regra *default* direcionada às compensações de carbono; e uma mensagem subliminar direcionada à prevenção do consumo de cigarros e da obesidade. Além de demonstrar um apoio amplo aos *nudges* de maneira geral, Reisch e Sunstein extraíram dois princípios que devem ser levados em consideração pelos arquitetos de escolha: a) *nudges* inequivocamente manipulativos geram rejeição - tema que será retomado na Seção 6.3; e b) parece adequado contar com o consentimento das pessoas antes de implementar uma política que transfira os seus recursos pessoais, ainda que seja para uma boa causa.

De maneira semelhante, Sunstein (2016b) avaliou a aprovação de *nudges* nos Estados Unidos e identificou uma série de princípios que parecem emergir da opinião dos participantes do experimento. Inicialmente, os americanos rejeitaram os *nudges* que pareçam promover fins ilícitos – como religiosos ou políticos –, assim como os *nudges* que pareçam inconsistentes com os valores ou com os interesses da maior parte das pessoas. Havia apoio, por outro lado, aos *nudges* com finalidades legítimas e consistentes com esses interesses e com esses valores.

O terceiro princípio, em linha com o experimento que Sunstein conduziu com Reisch, é a tendência à rejeição pelas pessoas de *nudges* que provoquem perdas financeiras: antes desses *nudges*, é importante que as pessoas possam expressar os seus interesses. Em outros termos, "os arquitetos de escolha não devem produzir perdas a partir do uso da inércia ou da desatenção das pessoas contra si próprias" (SUNSTEIN, 2016b, p. 203)[349]. Ao mesmo tempo, os dados obtidos por Sunstein indicaram que a maioria das pessoas rejeita a inscrição automática em programas de caridade, sobretudo se promovida por instituições públicas. Os *nudges* considerados inaceitavelmente manipulativos, como mensagens subliminares, foram rejeitados pelos participantes.

Por fim, Sunstein (2016b) também apresentou uma distinção relevante a depender do partido político que os participantes defendiam. A princípio, a aprovação de Republicanos, Democratas ou Independentes a um *nudge* específico dependia do seu conteúdo, isto é, da

[349] Tradução livre. Texto original: "[...] *choice architects should not produce losses by using people's inertia or inattention against them*".

medida em que os valores promovidos pela política se alinhavam aos valores políticos defendidos por esses partidos. Além disso, Sunstein identificou que os Republicanos se mostraram mais céticos com relação aos *nudges* governamentais em si: mesmo que os fins promovidos fossem relevantes e consistentes com os interesses dos decisores, Republicanos, por vezes, rejeitaram tal medida. Em linha com essa descoberta, Sunstein concluiu que alguns Republicanos e, até mesmo, alguns Democratas e Independentes parecem apoiar uma presunção refutável contra os *nudges*, se o governo puder evitá-los.

Todos esses estudos mostram, no mínimo, que boa parte dos *nudges* contam com aprovação popular. Apesar disso, é relevante destacar que a definição de *nudge* adotada por Sunstein é mais ampla, como já demonstrado no Capítulo 3. Alguns dos pontos que geram rejeição, nos termos apresentados acima, estão ligados aos *nudges* que se encaixam na definição estrita adotada nesta dissertação, como a exploração de fatores supostamente irrelevantes, a manipulação das capacidades decisórias e a utilização da presunção de consentimento.[350] Replicar esses estudos a partir de definições estritas de *nudge* pode informar melhor a maneira como as pessoas realmente se comportam com relação a essas medidas.

De todo modo, os argumentos expostos acima indicam que, na prática, os *nudges* e outras medidas de paternalismo suave talvez não gerem preocupações concretas. As pessoas estão acostumadas a medidas paternalistas e podem aprová-las, especialmente se estiverem de acordo com as finalidades promovidas por elas. Fins ilícitos até podem ser promovidos, mas essa não é uma questão referente ao problema do conhecimento; trata-se, na verdade, de uma questão institucional, relacionada à maneira de definir os objetivos coletivos a serem promovidos por normas e políticas públicas.

Embora amenize parte das preocupações ligadas ao problema do conhecimento, o argumento exposto neste item em favor da promoção de bem-estar por *nudges* é insuficiente por si só. Na prática, ele não tenta reorientar a coerência entre a teoria subjacente aos *nudges* e a aplicação prática dessa técnica de intervenção. Por mais que a abstração seja uma armadilha ser evitada na discussão filosófica de políticas públicas, oferecer os fundamentos normativos subjacentes aos *nudges* paternalistas é essencial para que o arquiteto de escolhas saiba quais perguntas devem ser feitas e respondidas antes de justificar um *nudge* específico. Afinal, por

[350] Gidon Felsen, Noah Castelo e Peter B. Reiner (2013), nesse sentido, demonstraram que as influências ostensivas (*overt*) são mais aceitáveis que influências ocultas (*covert*). Em diferentes cenários, como alimentação saudável, consumo prudente, promoção de exercícios físicos, prudência em investimentos e produtividade no trabalho, os participantes aprovaram, em maior grau, as medidas que tentavam melhorar decisões conscientes, em comparação àquelas que exploravam procedimentos automáticos de tomada de decisão.

trás de fundamentos paternalistas, podem existir razões escusas[351] e deixar claros os seus fundamentos em termos de bem-estar é um meio de facilitar o seu controle, questão que será debatida na Seção 6.4.

6.2.3.2. Superando o problema do conhecimento: preferências em condições epistemologicamente favoráveis

Paternalismos suaves, como o paternalismo libertário, sofrem mais com o problema do conhecimento, em razão das suas estratégias tentarem influenciar o meio escolhido pelos decisores para atingir seus próprios fins. Trata-se de um paternalismo fraco e orientado apenas à alteração dos meios adotados pelos decisores, como discutido no Capítulo 2. Paternalismos tradicionais, por sua vez, tentam ampliar o bem-estar das pessoas, de acordo com o julgamento do próprio paternalista (TOR, 2020).

Conforme o exemplo de Tor (2020), uma política paternalista tradicional pode desencorajar as pessoas a dirigirem sem cintos de segurança pura e simplesmente porque os *policy-makers* acreditam que essa decisão traria prejuízos aos decisores. Não há, aqui, uma preocupação em ampliar o bem-estar, segundo os próprios critérios do decisor. Pouco importa, então, se o decisor tem predileção por aventuras e pelo risco e não quer utilizar cintos de segurança.

A situação é diversa no caso de paternalismos suaves e fracos: a ideia, nessa hipótese, é permitir que aqueles que não desejam seguir a direção recomendada pela política tenham a liberdade de agir de outro modo. Como, no entanto, definir qual é a opção que melhora o bem-estar dos agentes, se as suas escolhas deixam de ser um guia para essa análise?

O primeiro passo para superar o problema do conhecimento é pensar em outros critérios de bem-estar que poderiam substituir o critério das preferências reveladas. Mergulhar profundamente nessa tarefa exigiria uma leitura mais ampla de toda a economia do bem-estar e a vertente comportamental dessa área. Evidentemente, esse empreendimento seria inviável nesta dissertação. De todo modo, apresentar as principais alternativas às preferências reveladas, ainda que sem entrar em todos os seus detalhes, é uma etapa necessária à discussão empreendida nesta seção.

[351] Nesse sentido, em linha com o que foi exposto no Capítulo 4, Christine Clavien (2018) pondera que os *nudges* podem ser elaborados com três objetivos distintos. Em primeiro lugar, existem objetivos egoístas, a partir do qual os arquitetos de escolha tentam favorecer o seu próprio interesse. Há, ainda, objetivos sociais, quando são promovidos *nudges* welfaristas. Por fim, há objetivos ligados ao decisor (*nudgee-driven*), que se confundem com os *nudges* paternalistas libertários.

Inicialmente, as teorias hedônicas de bem-estar poderiam ser adotadas. Kahneman, um dos adeptos dessas teorias, diferencia a utilidade da decisão (*decision utility*) – termo utilizado para representar a utilidade em sentido econômico – da utilidade experimentada (*experienced utility*, isto é, a experiência hedônica que é realmente proporcionada por uma determinada decisão. Por mais que os sujeitos acreditem saber prever qual seria a utilidade da sua decisão, erros de previsão frequentemente ocorrem e a utilidade experimentada não corresponde à imaginada previamente (KAHNEMAN, 2012; KAHNEMAN; WAKKER; SARIN, 1997)[352].

Apesar disso, como Sunstein (2014a, p. 109-110, 2020c) pondera, as pessoas se preocupam com outros elementos desvinculados da sua experiência hedônica para tomar decisões. Viver uma vida com significado e propósito, promover o bem-estar de sua família, ter controle sobre a própria vida, buscar a excelência e vários outros valores podem guiar as decisões das pessoas, ainda que, eventualmente, não lhes deixem mais felizes.

Outro ponto problemático dessa abordagem é a utilização de medidas de bem-estar autorreferido, extraídas de questionários como critério para definição de políticas públicas.

B. Douglas Bernheim (2016), por exemplo, aponta problemas nesse método ligados a, de maneira simplificada: a) como definir o que realmente é o bem-estar; b) a confiabilidade das respostas, considerando que nem sempre as pessoas vão reportar o seu sentimento verdadeiro; e c) a falta de escala adequada para mensurar o bem-estar. Loewenstein e Haisley (2008), por seu turno, criticam essa abordagem sob a ótica da adaptação hedônica: as pessoas se adaptam a circunstâncias felizes e infelizes e retornam ao ponto hedônico original depois de um certo tempo. Por mais que pacientes de hemodiálise apresentem níveis de felicidade semelhantes aos de pessoas saudáveis, não parece adequado considerar que graves doenças crônicas sejam um resultado desejável. A mensuração da utilidade experimentada, no entanto, poderia levar a essa conclusão e afastar a conclusão de que a doença é um resultado negativo.

As teorias hedônicas, assim como aquelas baseadas nas preferências, levam em consideração o bem-estar subjetivo das pessoas. Outra vertente teórica, no entanto, leva em consideração critérios objetivos de bem-estar, baseados em bens que seriam valorizados por

[352] Kahneman (2012; KAHNEMAN; WAKKER; SARIN, 1997) demonstraram, ainda, a existência de erros sistemáticos na avaliação retrospectiva a utilidade experimentada. Em outras palavras, a experiência da decisão é diferente da recordação da decisão, o que leva Kahneman a diferenciar o eu experiencial (*experiencing self*) do eu recordativo (*remembering self*) (KAHNEMAN, 2012, p. 476). Nos experimentos em que essa distinção foi demonstrada, duas regras principais foram descobertas: (a) a regra do pico-fim (*peak-end rule*), segundo a qual a média do nível hedônico experenciado no ápice da experiência hedônica e no seu fim previu de maneira adequada o nível de bem-estar relatado pelos participantes dos experimentos; e (b) a negligência com a duração (*duration neglect*), já que a duração das experiências não alterou as avaliações hedônicas dos participantes (KAHNEMAN, 2012, p. 474-475).

qualquer pessoa (SUNSTEIN, 2020c, p. 15)[353]. Essas visões dão origem às posições perfeccionistas, que recomendam, enquanto visão política, que os governos tentem guiar a vida das pessoas em direções consideradas objetivamente boas pelo próprio governo (CONLY, 2013, p. 103). Em termos de bem-estar, o perfeccionismo parte da ideia de que: a) há uma noção objetiva do que significa o bem-estar humano; e b) deixar de seguir essa noção, representa um prejuízo às pessoas, ainda que elas estejam satisfeitas com as escolhas que fazem (CONLY, 2013, p. 106). Vidas objetivamente ruins, nesse caso, deveriam ser rejeitadas e, nessas hipóteses, com o Sunstein pondera, teorias que levam bens objetivos em consideração são relevantes (SUNSTEIN, 2020c, p. 16).

O grande problema dessas abordagens, como pondera Conly (2013), é definir em que realmente consiste o bem-estar objetivo. Fazer escolhas livres, amar, ser capaz de produzir arte, ter acesso a uma vida digna, dentre outros valores, podem ser considerados aspectos constituintes de uma boa vida. A dificuldade, acrescenta Conly, é impor esses valores àqueles que não os veem como fatores essenciais à sua satisfação pessoal. Como um acordo acerca de tais elementos é improvável, na prática, o perfeccionismo pode ligado a um aspecto moral, ligado à noção de que as pessoas devem seguir determinadas posições, independentemente do bem-estar que elas lhes proporcionam. Integrá-lo à legislação ou a políticas públicas concretas pode resultar em moralismo jurídico, que busca, apenas, impor determinados padrões de comportamento.

A apresentação de todas essas teorias, de maneira agrupada, é uma simplificação. Não se pretende, nesta dissertação, aprofundar-se nos profundos debates filosóficos e econômicos que envolvem o bem-estar. Nesse ponto, vale a máxima de Rawls, escrita, segundo Sunstein (2020c, p. 16), em um manuscrito não publicado: "Colocamos uma sinalização. Sem pensamento profundos aqui; as coisas já vão suficientemente mal"[354].

Feita essa ressalva, os problemas em todos os critérios apresentados acima parecem sinalizar que, no âmbito de decisões privadas, definir o que gera maior bem-estar às pessoas não é uma tarefa trivial. Nesse caso, Sunstein (2020c) está correto ao defender, como tentativa de um acordo incompletamente teorizado entre as posições levantadas acima, uma presunção refutável de manutenção da liberdade de escolhas, no caso de escolhas pessoais que não geram dano a terceiro. Em termos de bem-estar, essa é a razão pela qual o paternalismo coercitivo comportamental de Conly (2013) é rejeitado nesta dissertação. Por mais que ela sustente que o paternalismo coercitivo só é justificável se uma análise de custos e benefícios comprovar

[353] Cf., por exemplo: (NUSSBAUM, 2011; RAZ, 1986).
[354] Tradução livre. Texto original: "*We post a signpost. No deep thinking here; things are bad enough already*".

empiricamente a sua superioridade, na prática, o problema do conhecimento parece indicar que dificilmente medidas coercitivas que limitassem decisões privadas, que só causam prejuízos ao próprio agente, seriam efetivas e gerariam ganhos de bem-estar.

De todo modo, fatores supostamente irrelevantes geram prejuízos de bem-estar aos decisores. Mais do que isso, esses fatores podem ser explorados por agentes privados, que auferem lucro com tal exploração. Pessoas reais, como já discutido nesta dissertação, nem sempre são capazes de maximizar ou, mesmo, satisfazer o seu próprio bem-estar. Por essa razão, os *nudges* podem ser úteis, ainda que em termos paternalistas.

Como, então, superar o problema do conhecimento? Bernheim, sobretudo nos seus trabalhos com Antonio Rangel, desenvolveu modelos baseados nas escolhas, que tentam compatibilizar a soberania das decisões individuais e, ao mesmo tempo, reconhecer as evidências da economia comportamental.[355] Embora expor toda a complexidade do seu modelo ultrapasse os objetivos desta dissertação, as suas linhas gerais podem representar uma saída adequada para enfrentar o problema do conhecimento.

Bernheim (2016) diferencia, inicialmente, os julgamentos diretos e os julgamentos indiretos que um sujeito faz. Enquanto os julgamentos diretos se referem aos resultados almejados pelo sujeito para o seu próprio bem, isto é, os seus objetivos últimos[356], os julgamentos indiretos se referem às alternativas que levam a esses resultados.

O exemplo de Bernheim (2016, p. 17) envolve a escolha de uma entre duas caixas que contêm frutas, por Norma. A caixa vermelha contém maçãs, enquanto a caixa amarela, peras. O seu objetivo era escolher a fruta que mais gostava e ela não sabia qual das frutas estava em cada uma das caixas. Por um lapso da sua memória, Norma achou que a caixa amarela continha bananas e a escolheu, pois eram a sua fruta preferida. No entanto, se ela soubesse o conteúdo das caixas, escolheria a caixa vermelha, por preferir maçãs a peras. Nesse exemplo, a escolha entre as caixas é um julgamento indireto, enquanto a escolha entre os diferentes tipos de frutas é um julgamento direto.

Na visão de Bernheim, que é adotada nesta dissertação, os fatores supostamente irrelevantes - e a variabilidade das preferências conforme a situação de escolha – lançam dúvidas sobre as habilidades de julgamentos indiretos das pessoas. Apesar disso, em grande parte, a sua habilidade de realizar julgamentos diretos – e apreciar certos valores ou objetivos - é mantida. É verdade, no entanto, que os julgamentos diretos são frequentemente expressos a

[355] Cf., por exemplo: (BERNHEIM, 2016; BERNHEIM; ANTONIO, 2011).
[356] A ideia é próxima da distinção entre preferências de primeira e de segunda ordem já exposta no Capítulo 2. Vide nota 159 *supra*.

partir de julgamentos indiretos. Apesar disso, nem todos os julgamentos indiretos são feitos da mesma forma e, na visão de Bernheim, em algumas ocasiões, eles podem revelar de maneira mais confiável quais são os julgamentos diretos das pessoas.

Nesse sentido, Sunstein (2021b) sugere, em uma abordagem neo-hayekiana, baseada nas ideias de Bernheim e de outros autores[357], que seria possível se questionar o que as pessoas escolheriam em condições epistêmicas favoráveis. Na prática, Sunstein sugere cinco questões subsidiárias para operacionalizar esse teste:

a) Decisores consistentes que não fossem afetados por fatores supostamente irrelevantes decidiriam de que forma?
b) Decisores informados escolheriam qual opção?
c) O que decisores ativos escolheriam?
d) Em situações que não desencadeiam o viés do presente e o otimismo irrealista, qual seria a escolha das pessoas?
e) Sem atenção limitada e com uma visão mais ampla da situação, o que as pessoas escolheriam?

O que seriam essas condições epistêmicas? A resposta a essa pergunta depende da compreensão dos fatores supostamente irrelevantes que afetam uma determinada decisão. Entender técnicas de desenviesamento, já discutidas no Capítulo 4, certamente contribui para a identificação de tais condições. Também é necessário atentar-se que, nem sempre que as pessoas erram ou decidem de maneira contrária aos interesses do arquiteto de escolhas, um viés cognitivo será a causa de tal comportamento. Como alertam Kahneman, Olivier Sibony e Sunstein (2021, p. 164): "[...] culpar os vieses por todos os resultados indesejáveis é uma explicação inútil. Recomendamos que a palavra viés seja reservada para erros específicos e identificáveis e para os mecanismos que os produzem"[358].

De todo modo, a distinção apresentada por Clavien (2018) entre preferências motivacionais imediatas e preferências teóricas parece oferecer um caminho adequado à identificação de tais condições epistêmicas. Preferências motivacionais imediatas são avaliadas quando uma escolha específica é feita. Nessa ocasião, cada atributo da escolha gera – consciente

[357] Cf., por exemplo: (GOLDIN, 2015)
[358] Tradução livre. Texto original: "[...] *blaming every undesirable outcome on biases is a worthless explanation. We recommend reserving the word* bias *for specific and identifiable errors and the mechanisms that produce them*".

ou inconscientemente – uma reação positiva ou negativa no decisor. Um *nudge*, assim como a própria arquitetura de escolhas em si, introduz novos atributos a essa escolha para direcioná-la. As reações a todos esses atributos geram uma preferência motivacional geral, que impele o sujeito a escolher determinada forma, isto é, realizar um julgamento indireto, nos termos de Bernheim. Já as preferências teóricas são evocadas quando o sujeito pensa sobre uma situação de escolha sem ter que decidir efetivamente, hipótese em que pode reconhecer os atributos de cada opção sem a influência da arquitetura de escolhas e valorizá-las de maneira diversa.

Na visão de Clavien (2018, p. 11), as preferências teóricas devem ser levadas em consideração pelo arquiteto de escolhas, já que participam da avaliação pessoal explícita e consciente sobre o que o decisor deve ou não escolher e sobre a medida em que ele aceita um determinado *nudge*. As situações em que as preferências teóricas são evocadas - isto é, longe da situação de escolha - parecem condições epistemologicamente favoráveis aos questionamentos sugeridos por Sunstein. Em tais condições, pode-se evitar ou minimizar o problema do conhecimento.

6.2.4. Saldo de bem-estar e o argumento cético em favor dos *nudges*

Como demonstrado na Subseção 6.2.3, o problema do conhecimento pode ser superado. De todo modo, a dificuldade para mensurar preferências existe e deve ser levada em conta no momento de se elaborar, implementar e monitorar políticas públicas. Em outras palavras, o arquiteto de escolhas sempre deve considerar a possibilidade de que esteja empurrando as pessoas em uma direção inadequada. A legitimidade dos *nudges*, nesse caso, dependerá essencialmente de sua análise de custos e benefícios, o que é sugerido inclusive por Tor (2020), na sua postura crítica aos *nudges* que exploram fatores supostamente irrelevantes.

Antes de fazer algumas considerações sobre essa análise, é pertinente deixar claro que o problema do conhecimento afeta apenas *nudges* de cunho paternalista. Os *nudges* welfaristas, que têm por objeto a promoção de bem-estar social, como demonstrado no Capítulo 4, não são afetados por essa crítica. Afinal, o seu objetivo expresso não é direcionar as pessoas para satisfazer o seu bem-estar segundo os seus próprios critérios; na verdade, os *nudges* welfaristas buscam aumentar o bem-estar social e são adotados segundo os critérios estabelecidos normativamente pela sociedade. No caso do Brasil, os *nudges* welfaristas podem ser adotados, com o objetivo de, por exemplo, efetivar os princípios da ordem econômica estabelecidos no art. 170 da Constituição da República, como a livre concorrência, a defesa do consumidor, a redução das desigualdades regionais e sociais, a defesa do meio ambiente, dentre outros fatores.

De qualquer forma, com relação aos *nudges* paternalistas, há um argumento adicional em favor da sua implementação, que enfraquece parte das preocupações atraídas pelo problema do conhecimento. Os paternalismos suaves, em geral, são *paternalismos mistos*, que, como exposto no Capítulo 2, não são a única razão que legitima uma determinada intervenção do Estado. Para além das razões paternalistas, isto é, da crença de que as pessoas nem sempre promoverão o seu próprio bem-estar, cumulada com medidas para alterar escolhas em proveito dessas mesmas pessoas (SUNSTEIN, 2013a, p. 1854, 2014a, p. 54), a maior parte dos *nudges* defendidos por paternalistas suaves também incorpora elementos de bem-estar social.

Francesco Guala e Luigi Mittone (2015), nesse sentido, argumentam que há justificativas políticas – independentes do paternalismo[359] - à utilização de *nudges*. Na sua visão, danos que um indivíduo causa a si mesmo – por exemplo, por não poupar o suficiente para a velhice ou por fumar demais – causam efeitos não negligenciáveis a terceiros. Nesse sentido, no caso de uma política como o *Save More Tomorrow*, ainda que um sujeito específico prefira ter mais recursos à sua disposição no presente, de modo que o seu bem-estar não seria promovido pela intervenção, a política ainda poderia ser legítima: o governo pode desejar a ampliação da poupança para a aposentadoria, para limitar as consequências sociais da falta de recursos na velhice (TOR, 2016, p. 8).

Em um sentido semelhante, Russell Korobkin (2009) argumenta que as políticas do paternalismo libertário podem ser justificadas de maneira mais convincente por um welfarismo libertário (*libertarian welfarism*). Em sua visão, um welfarista libertário prescreve a utilização de *nudges* para encorajar os agentes a agirem de maneira a ampliar o bem-estar social, ainda que essa ampliação gere efeitos negativos ou incertos nas suas utilidades individuais (KOROBKIN, 2009, p. 27). Haveria sobreposição entre as políticas do welfarismo libertário e do paternalismo libertário, conforme acrescenta Korobkin (2009, p. 28), sempre que um dado comportamento crie externalidades positivas, nenhuma externalidade ou poucas externalidades negativas, em comparação com o benefício proporcionado aos agentes individuais.

Essas abordagens fornecem justificativas para a utilização de *nudges*, ainda que os efeitos sobre os decisores sejam incertos ou indeterminados, caso a valência das externalidades que seriam criadas seja clara (KOROBKIN, 2009, p. 29). O argumento, no entanto, não é tão substancial quanto Guala e Mittone ou Korobkin sugerem, já que as externalidades associadas às decisões influenciadas por um *nudge* paternalista são secundárias (GOLDIN, 2015, p. 233).

[359] Ressalta-se, ainda, que Guala e Mittone (2015) não consideram a visão de Thaler e Sunstein propriamente paternalista; em sua visão, ao ter como fundamento a promoção do bem-estar dos indivíduos, a proposta de Thaler e Sunstein seria essencialmente welfarista.

Os efeitos principais de tais decisões recaem sobre os decisores e é essa a razão pela qual essas decisões normalmente são tomadas sem qualquer influência de intervenções estatais.

De qualquer forma, diante de um cenário em que a) a arquitetura de escolhas sempre influenciará as decisões dos agentes; b) existe a possibilidade - a ser esclarecida em condições epistemologicamente favoráveis, como argumentado na Subseção 6.2.3 - de que as pessoas estejam criando prejuízos para si mesmas em razão de fatores supostamente irrelevantes; e c) existe uma decisão que pode, além de reverter esses potenciais prejuízos individuais, criar benefícios sociais indiretos, não há razões em termos de bem-estar para não utilizar *nudges*.

Evidentemente, como já mencionado, essa utilização deve ser amparada por análises de custos e benefícios. O modelo do paternalismo assimétrico que avalia se uma intervenção é válida nessa proposta parece indicar elementos importantes que devem integrar essa análise. Como já exposto no Capítulo 2, esse modelo avalia: a) os benefícios (*B*) criados para indivíduos afetados por fatores supostamente irrelevantes (*p*); b) os custos (C) aos agentes que não precisam do *nudge*, isto é, cujas preferências lhe sejam contrárias; c) os custos de implementação que serão arcados pelo governo (I); e d) os resultados que a intervenção gera às empresas (ΔΠ). Considerados todos esses elementos, se o saldo líquido for positivo, é provável que o *nudge* gere bem-estar suficiente para justificar a política.

Algumas preocupações, de todo modo, devem ser levadas em consideração pelo arquiteto de escolhas.

Em primeiro lugar, é necessário entender as características da população-alvo, que será submetida à política e entender a sua heterogeneidade. Um mesmo *nudge* pode não servir para todo e qualquer grupo social, isto é, não há um tamanho que sirva para todos (*one size does not fit all*). Desconhecer essa realidade e não testar a intervenção na população local antes de implementá-la pode fazer com que a política seja inefetiva, como demonstrado no Capítulo 5, e impedir que ela gere ganhos de bem-estar. Conhecer essas populações permite apurar com maior precisão os limites do conjunto de pessoas que precisa do *nudge* (*p*) e, consequentemente, uma melhor apreciação dos benefícios que a intervenção pode gerar, assim como dos prejuízos correspondentes para os demais grupos que não a aprovam.

Além disso, no caso da utilização de *nudges* impuros, é essencial que o arquiteto de escolhas esteja atento à medida em que os interesses dos agentes privados que implementarão a política estão alinhados à direção sugerida pelo *nudge*. Se o custo para esses agentes for substancial (ΔΠ), como demonstrado por Willis (2013) e como já discutido no Capítulo 5, para além do potencial aumento de preços, as empresas tomarão medidas para garantir que o *nudge* seja ineficaz, reduzindo o benefício (B) que a população alvo (p) auferiria com a medida ou,

mesmo, agindo para reduzir o número de pessoas que seria submetida ao *nudge*. O arquiteto de escolhas, afinal, deve sempre estar atento ao equilíbrio da pescaria.

Nesse caso, por mais que o governo possa agir para se contrapor à atuação dos agentes privados, aumentando o custo de tais *contra-nudges*, o custo de implementação da política (I) pode se tornar excessivo e injustificável diante dos efeitos potenciais que ela pode gerar. Sem maior intervenção estatal para realizar tal contraposição, por outro lado, o *nudge* pode simplesmente ser inefetivo e os benefícios (B) aos agentes que precisam ser empurrados (p) podem acabar sendo inferiores aos prejuízos gerados aos agentes que rejeitam o *nudge* (C).

A medida em que *nudges* impuros podem realmente gerar bem-estar parece estar essencialmente ligada ao alinhamento dos interesses do arquiteto de escolhas privado e do governo que exige a implementação do *nudge*. Por mais que essa seja uma questão empírica, isto é, esse não é um argumento normativo que abala a legitimidade de um *nudge* específico, parece razoável manter um certo ceticismo quanto aos ganhos de bem-estar que um *nudge* impuro pode gerar nessas situações. Afinal, como o arquiteto de escolhas privado controla todos os aspectos da arquitetura de escolhas, a menos que o governo interfira em todos eles – o que equivaleria a determinar, inclusive, como a empresa deve se comunicar com o consumidor em todas as instâncias –, o arquiteto de escolhas privado conseguirá tentar direcionar os decisores a tomarem decisões contrárias à direção indicada pelo *nudge*.

Em outra perspectiva, o arquiteto de escolhas também deve estar atento à simetria *nudge/sludge*, já exposta no Capítulo 4. A implementação de um *nudge* é associada a um aumento relativo das fricções associadas à escolha das opções contrárias. Mais do que isso, como Willis (2013) demonstrou, o sucesso de um determinado nudge, especialmente de uma regra *default*, depende da combinação de *nudges*, de enquadramentos diversos (*framing*) e dos processos necessários para agir de maneira contrária à direção sugerida pelo *nudge*. Nesse caso, para maximizar a efetividade da política, o arquiteto de escolhas pode ter incentivos para ampliar as fricções relativas associadas às outras opções. Apesar disso, para manter a natureza do *nudge* e impedir custos excessivos aos agentes que não querem ser empurrados, parece razoável que o arquiteto de escolhas evite um aumento relativo substancial de tais fricções. Em outras palavras, não deve haver fricções demais para que o sujeito decida de maneira ao *nudge*. Se esse for o caso, parece mais legítimo adotar uma medida coercitiva tradicional do que implementar uma medida coercitiva às avessas, pelo aumento das referidas fricções.

6.3. Autonomia em xeque: *nudges* transparentes e facilmente resistíveis como ferramenta promotora de atos autônomos

Na definição estrita adotada nesta dissertação, os *nudges* exploram explicitamente fatores supostamente irrelevantes, para reduzir as fricções (hedônicas, sociais ou cognitivas) associadas a uma determinada opção de escolhas. Ao fazê-lo, os *nudges* podem facilmente passar despercebidos pela sua população-alvo; a alteração dos tamanhos de copos e pratos, por exemplo, pode ser imperceptível ainda que para um sujeito detalhista, caso sua atenção esteja direcionada a outras prioridades; a presunção do consentimento para que um sujeito passe a fazer parte de um determinado programa de benefícios do Estado, por sua vez, pode provocar essa adesão sem que esse sujeito perceba, conscientemente, as consequências de sua omissão.

No Capítulo 2, após apresentar as falhas de mercado comportamentais, questionou-se se a sua identificação, ao promover mais oportunidades para intervenção estatal, sobretudo de cunho paternalista, não enfraqueceria as bases da autonomia. Afinal, as regras jurídicas têm conteúdo expressivo[360] e, ao incorporarem medidas de regulação do comportamento privado, especialmente de caráter paternalista, essas regras podem limitar a possibilidade de fazer escolhas, expressar a ideia de que as pessoas não são capazes de decidir por conta própria ou, ainda, modificar as preferências e os interesses dos decisores. Não é por acaso, nesse contexto, que um conjunto robusto e bem elaborado de críticas à utilização de *nudges* é centrado na proteção à autonomia. Esse conjunto de críticas será enfrentado nesta seção.

6.3.1. Os riscos apontados pelas críticas: coerção, manipulação, doutrinação e dominação

Como mencionado no Capítulo 2, existe um debate profundo a respeito da definição de autonomia. Ainda que esse debate extrapole os objetivos desta dissertação, um primeiro passo necessário à discussão da relação entre *nudges* e a autonomia é compreender que esse debate envolve diferentes definições de autonomia. Nesse sentido, Anastasia Vugts e outros autores (2020), em uma revisão sistemática da literatura a esse respeito, identificaram que as críticas costumam adotar três abordagens distintas de autonomia: a) liberdade de escolhas (*freedom of choose*); b) agência, no sentido de capacidade de escolher e decidir com base em razões e

[360] Sunstein (2019a, p. 41), por exemplo, afirma que, assim como ações, o direito tem uma função expressiva, que consiste em fazer declarações, em vez de controlar o comportamento diretamente. Se, por exemplo, leis exigirem a reciclagem, essa imposição expressa que a reciclagem é algo a ser valorizado, o que pode afetar normas sociais a respeito do meio ambiente de um modo desejável.

motivos e c) autodeterminação, isto é, a capacidade de agir com base em motivos e preferências que sejam autênticos.

Essas definições se relacionam aos três critérios utilizados nessa dissertação para avaliar se um ato é substancialmente autônomo, em linha com Faden e Beauchamp (1986). De fato, quando se reduz a liberdade de escolhas, haverá, naturalmente, maior espaço para influências controladoras; no limite, a influência pode ser coercitiva e o ato, completamente controlado. Por outro lado, se a agência é impactada, o sujeito pode ter menos capacidade de compreender as suas ações; se o ato não for substancialmente compreendido, o espaço da autonomia pode ter sido eliminado. Finalmente, se o sujeito decide movido por motivos de terceiros inadvertidamente, por mais que compreenda a ação e esteja livre de influências controladoras, o grau de intencionalidade do ato poderá ser incompatível com a autonomia.

A partir dessas definições, as críticas aos *nudges* ligadas à autonomia serão reunidas em três grupos diferentes: a) limitação da liberdade de escolhas; b) desrespeito à racionalidade do agente e infantilização; e c) alienação do agente e substituição de preferências[361]. Em cada um deles, foram selecionados trabalhos específicos para exemplificar a crítica exposta, de modo que não se almeja, nessa oportunidade, esgotar a revisão bibliográfica do tema.

6.3.1.1. Limitação da liberdade de escolhas: nudges são coercitivos?

A primeira crítica que será destacada sustenta que *nudges* reduziriam a liberdade de escolha (*freedom to choose*) das pessoas. Como já mencionado, os *nudges* são, em essência, medidas não coercitivas. O sujeito que é empurrado sempre tem a opção de optar por um rumo de ação distinto, sem que lhe seja atribuída qualquer sanção. Apesar disso, o próprio Sunstein (2016a, p. 64) reconhece que alguns tipos de *nudge*, como regras *default*, representam, no mínimo, um fardo para as pessoas. Se a liberdade em relação às múltiplas opções de escolhas do indivíduo for reduzida significativamente, poderia haver o risco de coerção (VUGTS *et al*., 2020, p. 8).

[361] Como mencionado acima, Sunstein (2015b, 2016a) costuma dividir as críticas éticas aos *nudges* em cinco frentes distintas: a) bem-estar; b) autonomia; c) dignidade; d) manipulação; e e) intenções e vieses dos arquitetos de escolha. Os problemas de bem-estar já foram expostos na Seção 6.2. Os problemas ligados à dignidade e à manipulação acabam guardam uma relação muito próxima com acepções específicas de autonomia, razão pela qual serão tratados em conjunto. Já as preocupações ligadas às intenções e vieses dos arquitetos de escolha não estão diretamente relacionadas aos *nudges*, mas a qualquer intervenção estatal. De todo modo, essas questões serão abordadas na Seção 6.4.

Grüne-Yanoff (2012, p. 636), por exemplo, afirma que as políticas do paternalismo libertário reduzem efetivamente a liberdade das pessoas. Em primeiro lugar, Grüne-Yanoff (2012, p. 637-638) sustenta que os efeitos e as condições de legitimidade das políticas do paternalismo libertário não são totalmente transparentes[362]. O autor afirma, ainda, que a sua aplicação exige alguma medida de obscuridade em relação aos elementos que envolvem a política proposta. Por consequência, os reguladores teriam maior poder arbitrário sobre as pessoas, o que, por si só, reduziria a liberdade. Por outro lado, Grüne-Yanoff (2012, p. 638-639) afirma que o paternalismo libertário amplia deliberativamente o número de intervenções nos processos de escolha, o que poderia representar uma espécie de coerção.

Em outro sentido, Rebonato (2014, p. 382) aponta um paradoxo relacionado à efetividade dos *nudges*. Por definição, os *nudges* são políticas não coercitivas, que poderiam ser evitadas de maneira simples pelos destinatários da política. Apesar disso, quanto mais efetivos fossem os *nudges*, menos pessoas optariam por um rumo distinto daquele indicado pela política. No limite, a liberdade por empreender um curso de ação diverso (*opt-out*) seria apenas nominal. Na prática, então, se o *nudge* for muito efetivo, não haveria liberdade real e os *nudges*, de fato, reduziriam a liberdade de escolhas das pessoas. Eles estariam, assim, próximos de medidas coercitivas, ainda que sem coerção na sua estrutura clássica.

6.3.1.2. Desrespeito à racionalidade do agente e infantilização: nudges são manipuladores?

A segunda crítica que ora se apresenta admite que os *nudges* podem respeitar a liberdade de escolha das pessoas. No entanto, como os *policy-makers* exploram traços do comportamento que não são deliberativos, eles não estariam tratando as pessoas como seres humanos racionais e, por consequência, acabariam por infantilizá-las (SCHMIDT; ENGELEN, 2020, p. 5). O risco, nesse caso, não seria a coerção, mas a manipulação (VUGTS *et al.*, 2020, p. 9)[363].

Hausman e Welch (2010, p. 131), por exemplo, afirmam que *nudges* poderiam ser uma ameaça maior à liberdade do que a coerção em sentido estrito, uma vez que ameaçam o controle das pessoas sobre suas próprias avaliações e sobre a deliberação. Rebonato (2014, p. 379), por

[362] Grüne-Yanoff atribui essa falta de transparência a três motivos distintos. Em primeiro lugar, os resultados científicos relacionados aos vieses comportamentais não seriam completamente transparentes, já que os contextos em que as políticas são implementadas seriam muito mais complexos do que as situações experimentais. Em segundo lugar, não seria claro quando as pessoas estão realmente sujeitas a vieses comportamentais. Para Grüne-Yanoff, determinar se uma "falha" comportamental é ou não decorrente de um viés não é uma tarefa trivial. Por fim, a efetividade de tais políticas seria maior se elas não fossem transparentes para os indivíduos que são afetados pela intervenção.

[363] Sobre a relação entre *nudges* e manipulação, Cf.: (WILKINSON, 2012).

seu turno, define o paternalismo libertário, pela possibilidade de manipular deficiências cognitivas para empurrar as pessoas a melhores escolhas. Posner (1998, p. 1575), em sentido semelhante, afirma que a imagem de ser humano descrita por Jolls, Sunstein e Thaler (1998), seria caracterizada por preferências instáveis e por infinita manipulabilidade.

Nessa perspectiva, Jeremy Waldron (2014) pondera que a autonomia é um princípio sobre como as preferências de um indivíduo são buscadas. Não seria, portanto, algo meramente redutível ao bem-estar. Waldron afirma, então, que os *nudges* não ensinam o decisor a deixar de utilizar heurísticas que não são apropriadas ou a abandonar intuições irracionais. Não se buscaria educar as escolhas, mas, com base nas fraquezas dos indivíduos, manipular sua compreensão da situação a fim de produzir o resultado que seria obtido a partir da reflexão racional. Para Waldron, esse tratamento seria semelhante àquele direcionado a uma criança e os decisores, por essa razão, seriam tratados como meros meios para um fim almejado – ainda que este fim seja o seu próprio bem-estar. As capacidades individuais de pensar e descobrir as coisas por conta própria não seriam valorizadas pelos proponentes dos *nudges*.

Wright e Ginsburg (2015, p. 38), por outro lado, afirmam que, quando *nudges* reduzem o rol de opções a serem escolhidas pelos indivíduos, tendem a infantilizar o público, uma vez que a habilidade efetiva de tomada de decisões é desenvolvida por tentativa e erro. Se os indivíduos são dispensados de realizar escolhas importantes, a sua habilidade de decidir poderia se corroer, de modo que eles não atingiriam todo o seu potencial enquanto cidadãos. Como afirma Bovens (2009, p. 214-215), se os *nudges* tiverem apenas efeitos em curto-prazo, a consequência de tais políticas, em longo prazo, poderia ser a infantilização dos indivíduos, isto é, uma diminuição da responsabilidade em matérias relacionadas ao seu próprio bem-estar.[364]

Em sentido semelhante, Saint-Paul (2011, p. 153) afirma que o fato de ser tratado pela sociedade como alguém responsável e unitário gera, em contrapartida, comportamentos responsáveis e unitários por parte dos indivíduos. Em outras palavras, "[o]s incentivos para resolver meus próprios problemas comportamentais são muito maiores se eu esperar que a sociedade me considere responsável pelas consequências das minhas ações" (SAINT-PAUL, 2011, p. 153)[365]. Nesse caso, com as políticas elaboradas para empurrar as pessoas para escolhas

[364] Nesse sentido, Rebonato (2014, p. 387) especula que seria razoável supor que, se as pessoas se acostumam a raciocinar acerca dos termos de um problema decisório, os caminhos neurais associados a esse processo seriam reforçados e seriam mais facilmente utilizados em tarefas decisórias subsequentes. Ao invés disso, se as pessoas não são encorajadas a raciocinar – por exemplo, com o uso de *nudges* –, seriam reforçados os caminhos neurais relacionados às decisões mais automáticas, ou, falando livremente, às partes "emocionais" do cérebro. Rebonato menciona, como argumento para sua especulação, a famosa pesquisa de Maguire *et al.* (2000) que demonstrou mudanças nos padrões neurais de taxistas de Londres que deixaram de usar a memória e passaram a utilizar GPSs.

[365] Tradução livre. Texto original: "*The incentives to solve my own behavioral problems are much larger if I expect society to hold me responsible for the consequences of my actions*".

melhores, esses indivíduos seriam vistos como menos responsáveis pelas consequências das suas ações. Por essa razão, elas teriam menos incentivos para resolver os seus próprios problemas comportamentais.

6.3.1.3. Alienação do agente e substituição de preferências: nudges são doutrinadores?

Uma terceira preocupação decorrente da utilização de *nudges* envolve a possibilidade de essa estratégia de política pública moldar as preferências e valores dos indivíduos a ela submetidos. Em outros termos, quando empurrados, os indivíduos não seriam mais os "autores" das suas escolhas, que deixariam de refletir seus desejos autônomos (SCHMIDT; ENGELEN, 2020, p. 4). A ameaça, nesse caso, seria a doutrinação, que faria com que as pessoas endossassem ideias e valores que não são realmente seus (VUGTS *et al*., 2020, p. 11).

Sunstein (2018c, 2019), nesse sentido, pondera que, em algumas situações, pode ser que os decisores fiquem satisfeitos com os resultados produzidos por dois ou mais *nudges* ou, ainda, que as suas preferências tenham sido construídas pelos *nudges*. Se, por exemplo, um sujeito prefere não contratar um seguro para o seu veículo, mas depois de ser automaticamente inscrito a uma determinada franquia de seguro, fica satisfeito com o resultado, o *nudge* pode ter transformado as suas preferências. Por mais que o fato de o sujeito continuar a decidir na direção indicada pelo *nudge* pareça indicar que a política lhe favorece em termos de bem-estar, a legitimidade dessa alteração de preferências é questionável em termos de autonomia.

Bovens (2009, p. 212), por outro lado, afirma que os *nudges* criam uma situação aberrante de falha na integração da estrutura de preferências das pessoas. De fato, enquanto suas escolhas revelam um determinado rumo de ação, aquela preferência específica pode não ser realmente autêntica. Por exemplo, um indivíduo pode passar a comer alimentos saudáveis em razão de um determinado *nudge*, mas não adquirir realmente a preferência por consumir tais alimentos. Na ausência da política, então, o comportamento do indivíduo deixaria de refletir essa preferência.

Bovens (2009, p. 210-211) destaca, ainda, que diferentes tipos de agência podem estar sujeitos aos *nudges*. Esses tipos de agência podem ser divididos em dois grupos distintos, conforme a política restabeleça a coerência entre a preferência revelada pela escolha e a verdadeira preferência do indivíduo ou não se preocupe com tal harmonização.

Quatro tipos de agência fazem parte do primeiro grupo. O primeiro deles é a ignorância, quando a falta de conhecimento faz com que seja difícil que o indivíduo cumpra as etapas necessárias para atingir o seu objetivo. O segundo tipo de agência é a inércia, que também

impede que as decisões das pessoas se conformem às suas preferências. Bovens menciona ainda a *acrasia*, isto é, as ações do indivíduo que contrariam os seus objetivos pessoais. Nesse caso, a fraqueza da vontade é o que se coloca no caminho entre o indivíduo e os seus objetivos em longo prazo. O quarto tipo de agência que faz parte do primeiro grupo é o mal-estar (*queasiness*), que afasta o indivíduo de determinada decisão, embora ela se conforme às suas preferências[366]. Em relação a esses tipos de agência, não haveria maiores problemas em termos de autonomia, uma vez que as pessoas seriam empurradas para uma direção compatível com as suas estruturas de preferências (BOVENS, 2009, p. 213).

Quanto ao segundo grupo de tipos de agência, por outro lado, Bovens (2009, p. 211) aponta que há problemas éticos mais sensíveis, já que a pessoa poderia ser empurrada para uma direção que não se harmoniza com a sua estrutura de preferências. O primeiro tipo de agência desse grupo é a exceção. Nessa hipótese, o rumo de ação apontado pelo *nudge* se conforma, de maneira estatisticamente comprovada, com a maioria do seu grupo; alguns indivíduos específicos, contudo, teriam preferências distintas - *i.e.*, seriam exceções. Em outra perspectiva, Bovens afirma que os *nudges* poderiam ser elaborados para que um benefício social fosse atingido, ainda que a escolha sugerida pelo *nudge* não se harmonizasse com a estrutura de preferências das pessoas. O fim, aqui, não é propriamente paternalista. De todo modo, em relação a esses dois tipos de agência, a situação é mais problemática em termos de autonomia: os indivíduos podem estar tomando decisões que não refletem, de fato, a sua estrutura de preferências.

Hausman e Welch (2010), por seu turno, afirmam que os *nudges* podem empurrar indivíduos para escolhas que eles não fariam na ausência da política. Em vez de a escolha ser feita a partir da avaliação das alternativas, o rumo de ação escolhido refletiria as estratégias adotadas pelo arquiteto de escolhas. Na prática, então, o *policy-maker* estaria moldando as escolhas feitas pelos indivíduos. Assim como a coerção, essa estratégia ameaçaria a liberdade, uma vez que o ato de moldar escolhas envolve uma substituição do julgamento do indivíduo pelo do sujeito que elabora a política. Em um sentido semelhante, Baldwin (2014, p. 846) afirma que, quando *nudges* vão além da mera apresentação de informação ou de lembretes, o arquiteto de escolhas tenta, de forma oculta (*covert*) ou semioculta, estender a sua própria concepção de bem-estar aos destinatários da política. Nesse caso, pode haver o remodelamento da ideia que as pessoas têm sobre o seu próprio bem-estar, o que pode causar prejuízos em termos de autonomia.

[366] Por exemplo, um sujeito pode deixar de ser um doador de órgãos, uma vez que esse assunto lhe gera um mal-estar ao fazê-lo pensar sobre a morte.

6.3.2. Evitando as preocupações levantadas pelas críticas: mecanismos para afastar a coerção, a manipulação e a doutrinação

Sunstein empreendeu consideráveis esforços para responder às críticas aos *nudges* ligadas à autonomia.[367] Em linhas gerais, Sunstein (2015b, p. 424-426, 2015a, p. 104, 2016a, 2019a, p. 94) parece sempre sustentar que *nudges* têm por objetivo melhorar a "navegabilidade" (*navigability)*, no sentido de permitir que as pessoas cheguem às decisões que desejam, de forma a satisfazer suas preferências antecedentes. Afinal, como já mencionado, *nudges* paternalistas são exemplos de paternalismo fraco, "orientado para os meios" (*means-oriented*), já que não há uma tentativa de questionar ou alterar os fins dos indivíduos que têm suas escolhas influenciadas (SUNSTEIN, 2019a, p. 121).

O exemplo comumente mencionado por Sunstein para ilustrar esses pontos é o Sistema de Posicionamento Global (GPS): apenas se indica o caminho que levará o usuário ao destino por ele especificado. Ao mesmo tempo, trata-se de um "paternalismo leve" (*soft paternalismo*), na medida em que não envolve coerção ou incentivos materiais substanciais, de modo que a liberdade de escolha é mantida (SUNSTEIN, 2019a, p. 121). Por essas razões, não haveria objeções - inclusive em termos de autonomia – aos *nudges* em si, mas, apenas, a algumas espécies particulares de *nudge*.

Apesar do notável esforço de Sunstein para enfrentar essas questões, dois problemas específicos da sua argumentação fazem com que o debate chegue a um impasse. Em um primeiro aspecto, como já mencionado no Capítulo 3, Sunstein parte de um conceito de *nudge* mais amplo do que aquele adotado pelos críticos. Dentro desse conceito, existem medidas que, claramente, não geram qualquer problema ético. Exemplos disso são os lembretes ou a divulgação de informações relevantes para a tomada de decisões[368]. Nesse caso, evidentemente, não haverá objeção a *nudges* em si, uma vez que algumas modalidades de *nudges* não geram qualquer tipo de preocupação nem ameaçam a autonomia. Os críticos, no entanto, não se preocupam com esse tipo de política, mas com outras mais intrusivas, como regras *default*, alterações substanciais de enquadramento (*framing*) e outras alterações na estrutura das decisões. Assim, ao ressaltar que não existem objeções a *nudges* em si, Sunstein se esquiva do debate.

---

[367] Cf., nesse sentido: (SUNSTEIN, 2014a, 2015e, 2015b, 2015a, 2016a).
[368] Cf., nesse sentido, (SUNSTEIN, 2014d, 2019c).

Em outra perspectiva, como bem expuseram Engelen e Nys (2020), Sunstein baseia a sua defesa dos *nudges* na autocracia. Como mencionado no Capítulo 2, Kant (1997, p. 363) diferencia a autonomia - capacidade de determinar os próprios fins - da autocracia - capacidade de agir em conformidade com esses fins para realizá-los. Se os *nudges* paternalistas são apenas uma forma de paternalismo fraco, apenas contribuirão para que as pessoas tenham maior capacidade de realizar empiricamente os fins que almejam. A autonomia, entretanto, não se limita a essa realização empírica. Na definição de Engelen e Nys, seria necessário apurar se os meios indicados pelos nudges, podem conduzir as pessoas a caminhos que realizem os valores e preferências de terceiros, *i.e.*, os arquitetos de escolha.

Por outro lado, a crítica aos *nudges* centrada na autonomia não se limita às políticas paternalistas. Como mencionado na introdução deste Capítulo, mais do que os seus fundamentos e as suas justificativas, os *nudges* influenciam o comportamento de uma maneira específica que pode atrair preocupações: a partir da exploração de fatores supostamente irrelevantes, que pode passar despercebida pelos agentes decisores. Nesse caso, ainda que um *nudge* seja welfarista, a obscuridade potencialmente ligada à sua forma de atuação pode ser questionável.

Feitas essas considerações, é preciso, então, assumir uma posição nesse debate. Na prática, os *nudges* podem reduzir a autonomia das pessoas? A resposta, evidentemente, dependerá do caso concreto. Para começar a traçá-la, o critério da fácil resistibilidade, exposto por Saghai (2013) na sua tentativa de salvar o conceito de *nudge* e já apresentado no Capítulo 3, é um ponto de partida adequado, sobretudo por partir da mesma noção de autonomia adotada nesta dissertação. Como mencionado anteriormente, por mais que se rejeite a posição de Saghai em termos de definição das estratégias que seriam nudges, as suas considerações são relevantes para a avaliação normativa de tais intervenções, sobretudo no que diz respeito à autonomia.

Nesse contexto, retomando alguns aspectos conceituais expostos no Capítulo 3, Saghai (2013) sustenta que uma intervenção será facilmente resistível se um sujeito B que sofre a influência de A para agir de um determinado modo φ: a) tiver a capacidade de se tornar consciente da pressão de A para levá-lo a φ (capacidades de direcionar a atenção); b) tiver a capacidade de *inibir* a propensão à φ gerada pela influência de A (capacidades de inibição); e c) não estiver sujeito a uma influência ou colocado em circunstâncias que prejudicariam significativamente o exercício fácil (*effortless*) dessas capacidades.

Como, então, o critério da fácil resistibilidade se conecta às críticas aos *nudges* em termos de autonomia?

Inicialmente, as críticas ligadas à liberdade de escolhas se relacionam à possibilidade do sujeito de exercer facilmente as capacidades de atenção e de inibição. Em uma primeira visão, essa crítica pode parecer mal direcionada, quando se considera que *nudges* são essencialmente não coercitivos e não alteram substancialmente a estrutura de incentivos financeiros. Nesse caso, por definição, os *nudges* não seriam coercitivos e, em termos de autonomia, seriam muito mais palatáveis do que medidas de comando e controle tradicionais.[369]

Apesar disso, um aspecto que deve ser levado em consideração nessa análise, é a simetria *nudge/sludge*. Como já discutido nesta dissertação, os *nudges* aumentam a fricção relativa das opções contrárias àquela cuja escolha é facilitada pelo *nudge*. Mais do que isso, como a efetividade do *nudge* pode estar ligada às maneiras pelas quais o sujeito pode decidir de outra forma, como regras de *opt-out*, o arquiteto de escolhas pode ter incentivos para dificultar essa decisão, como já especulado na Subseção 6.2.4.

Nesse caso, se um *nudge*, além de reduzir as fricções associadas a uma determinada opção, ampliar excessivamente as fricções relativas associadas às opções contrárias, o resultado pode ser quase coercitivo. Na prática, decidir de outro modo pode ser inviável ou excessivamente custoso, em termos hedônicos, sociais ou cognitivos para o agente. Além de desvirtuar a sua natureza não coercitiva, um *nudge* acompanhado de um tamanho aumento de fricções pode ser substancialmente controlador e, assim, impedir que as decisões tomadas a partir da sua influência sejam atos autônomos. Não haveria, nesse caso, fácil resistibilidade, já que exercer as capacidades de direcionar a atenção e de inibição seria excessivamente custoso.

Essa preocupação não é essencialmente teórica. Um *nudge* brasileiro pode exemplificar esse problema. A Lei nº 13.183, de 4 de novembro de 2015, alterou a Lei n.º 12.618, de 30 de abril de 2012, para prever a adesão automática dos servidores públicos titulares de cargo efetivo da União, as suas autarquias e fundações ao plano com remuneração superior ao limite máximo estabelecido para os benefícios do Regime Geral de Previdência Social ao respectivo plano de previdência complementar. Em outras palavras, houve a alteração da regra *default*, para que esses servidores passassem a integrar o plano de previdência complementar da União.

Ao disciplinar essa adesão automática, a Orientação Normativa nº 9, de 19 de novembro de 2015, do Ministério de Planejamento, Orçamento e Gestão (2015) previu que os pedidos de desistência das inscrições automáticas dos servidores deveriam ser apreciados pela FUNPRESP-EXE, antes de serem deferidos (art. 4º, § 1º). Esse requisito, via de regra, traduz-se em uma barreira ao *opt-out*, que pode conduzir a uma manutenção excessiva da opção

[369] Em termos de autonomia, medidas coercitivas como aquelas propostas por Conly (2013), por sua vez, são rechaçadas nesta dissertação com base no critério da fácil resistibilidade.

default, incluindo para indivíduos que não gostariam dessa opção. A norma, nesse caso, poderia funcionar como um *sludge*: um excesso de fricções que pode fazer com que as pessoas tenham dificuldade de seguir o caminho que mais gostariam. Assim, um dos aspectos mais marcantes que caracterizam o *nudge* - a saber, a possibilidade de optar livremente por um caminho diverso - seria distorcido.

Evitar os nudges quase coercitivos passa, em primeiro lugar, por reconhecer que *nudges* ampliam as fricções relativas associadas às demais opções de escolha. Ao mesmo tempo, deve-se reconhecer que, para garantir a efetividade de políticas específicas, arquitetos de escolha terão incentivos para ampliar ainda mais essas fricções e garantir que as pessoas sigam na direção empurrada. Os *nudges* não controladores, no entanto, devem evitar essas armadilhas. Na análise de custos e benefícios, se um *nudge* precisa de fricções excessivas associadas às opções contrárias para funcionar bem, talvez seja mais razoável e transparente pensar em medidas tradicionais de comando e controle.

Em outra perspectiva, os *nudges* manipulativos podem impedir que os agentes tenham a adequada compreensão da situação de escolha. Ao serem influenciados por técnicas que exploram fatores supostamente irrelevantes, os decisores devem ser capazes de perceber essa influência. No caso da alteração de uma regra *default*, o sujeito deve estar em condições de perceber que a sua omissão é juridicamente relevante e trará impactos à sua esfera jurídica pessoal. Se essa percepção não for possível, dificilmente o sujeito conseguiria exercer as suas capacidades de direcionar a atenção e de inibição.

O ato praticado sob a influência de um *nudge* obscuro e manipulativo seria heterônomo, na medida em que o sujeito deixaria de ter uma compreensão substancial da situação de escolha. Uma mensagem subliminar que estimula o consumo de alimentos saudáveis, por exemplo, ultrapassa as capacidades de percepção dos agentes. Do mesmo modo, a substituição inadvertida de pratos, talheres e copos por objetos menores, dificilmente, seria percebida pelos decisores. Em tais casos, o mecanismo adequado para suavizar as preocupações ligadas à autonomia é a transparência e a publicidade das medidas.

Nesse sentido, Thaler e Sunstein (2009, p. 247-248, 2021, p. 325-328) afirmam endossar o princípio da publicidade de John Rawls (1999). Na sua forma mais simples, esse princípio proíbe que um governo adote uma política que ele não conseguiria ou não desejaria defender publicamente perante os cidadãos. Nesse sentido, se o princípio fosse desrespeitado, o governo estaria sujeito ao embaraço ou a medidas ainda mais drásticas por parte da população, quando a política fosse divulgada. Desrespeitar o princípio da publicidade configura um desrespeito aos próprios cidadãos.

Apesar disso, como Hansen e Jespersen (2013, p. 16) ponderam, esse princípio não parece ser suficiente para evitar a manipulação em todos os casos. Mesmo Thaler e Sunstein (2009, p. 246), ao discutirem as mensagens subliminares, na edição original do *Nudge*, sustentam que elas seriam censuráveis já que são invisíveis e de impossível monitoramento. Embore esse trecho não conste na edição final do *Nudge*, Thaler e Sunstein (2021, p. 328) sustentam que uma declaração de direitos para nudges ("*a bill of rights for nudging*") deve proibir a publicidade subliminar, o que parece alinhado à sua posição original.

O que, então, deve ser visto como um *nudge* transparente? Na concepção epistêmica de Hansen e Jespersen (2013, p. 17), um *nudge* é transparente se existe uma expectativa razoável de que o decisor pode identificar de forma clara as intenções subjacentes ao *nudge* e os meios pelos quais a mudança de comportamento é buscada. A noção é próxima da transparência de cada elemento da intervenção (*token interference transparency*)[370], em que se pode reconhecer cada alteração específica na arquitetura de escolhas, sugerida por Bovens (2009, p. 216–217) como mecanismo de garantir a transparência de *nudges* legítimos.

Os *nudges* não transparentes, por seu turno, "funcionam de uma maneira que o cidadão na situação não consegue reconstruir seja a intenção ou os meios pelos quais a mudança de comportamento é buscada" (HANSEN; JESPERSEN, 2013, p. 18)[371]. Mensagens subliminares, alterações no tamanho de pratos e outros objetos, a mera inscrição automática dos decisores em programas do governo e outras estratégias similares são, a princípio, não transparentes segundo essa definição. Na abordagem de Bovens (2009, p. 216-217), estratégias como mensagens subliminares não conseguiriam obter transparência de cada elemento da intervenção, já que o seu funcionamento é baseado em passar ao largo da atenção e da consciência do decisor.

Hansen e Jespersen (2013) sugerem que a noção de transparência epistêmica seja utilizada como critério para avaliar a medida em que um *nudge* é manipulativo. A definição de *nudge* adotada nesta dissertação, como já mencionado no Capítulo 3, confunde-se com a categoria de *Nudges 1* utilizada por Hansen e Jespersen (2013). Com relação a essa categoria, em que há a exploração de fatores supostamente irrelevantes como parte integrante da técnica

[370] A transparência de cada elemento da intervenção (*token interference transparency*) é mais rigorosa do que a mera transparência da intervenção (*interference transparency*), em que o governo é transparente apenas a respeito de como tentará intervir nas escolhas. A tradução das expressões utilizadas por Bovens não é trivial. Opta-se por diferenciar entre a transparência da intervenção de maneira geral, em que apenas é descrita a maneira como será feita a intervenção, da transparência de cada elemento da intervenção, na qual o indivíduo atingido consegue identificar cada aspecto da intervenção a qual está submetido.

[371] Tradução livre. Texto original: "[...] *working in a way that the citizen in the situation cannot reconstruct either the intention or the means by which behavioural change is pursued*".

de intervenção, a falta de transparência epistêmica leva à manipulação do comportamento, segundo Hansen e Jespersen (2013).

Nesse caso, para afastar o risco de manipulação e garantir que um *nudge* promova atos substancialmente compreendidos pelos decisores, o arquiteto de escolhas deve buscar a transparência epistêmica antes, durante ou depois da intervenção. Na visão de Hansen e Jespersen, a utilização de tais tipos de *nudge* amplia as responsabilidades dos arquitetos de escolhas e exige que eles garantam que tais intervenções sejam baseadas em processos democráticos, tema que será retomado na Seção 6.4. A precaução adequada diante de tais intervenções pode envolver o fornecimento de informações claras aos decisores, que registrem os motivos que justificam uma intervenção e as medidas implementadas para direcionar a mudança de comportamento (HANSEN; JESPERSEN, 2013, p. 26).

Bovens (2009, p. 216-217) ressalta, porém, que, aparentemente, os *nudges* funcionariam melhor no escuro, de forma que, quanto mais transparência quanto aos elementos da intervenção fosse exigida, menos efetivos os *nudges* seriam. Por mais que experimentos que constatem a reatância psicológica, como o já mencionado experimento de Arad e Rubinstein (2018), citados no Capítulo 5, possam ilustrar essa preocupação, a literatura mais recente parece apontar em outra direção.

Loewenstein e outros autores (2015), por exemplo, demonstraram que a divulgação ampla de uma opção *default* não reduz de maneira significativa os seus efeitos. Steffel e outros autores (2016), por sua vez, demonstraram que a divulgação da intenção por trás de regras *default* não é suficiente para afastar a sua influência, embora a articulação ativa das preferências prévias dos decisores tenha reduzido a influência da regra *default*. Henrik Bruns e outros autores (2018) avaliaram em laboratório a medida em que diferentes tipos de transparência – conhecimento da potencial influência do *nudge*, do seu propósito ou de ambas as coisas – impactaria a efetividade de uma regra *default* que promovia contribuições para a redução da emissão de carbono. Os seus dados indicaram que nenhuma das formas de transparência reduzia a efetividade da medida e não indicaram evidências de reatância psicológica relacionada à transparência. Do mesmo modo, H. Min Bang e outros autores (2020) exploraram a medida em que a revelação de que uma determinada arquitetura de escolhas foi estruturada para influenciar decisões impacta a efetividade da medida.

Todos esses dados parecem indicar que é factível pensar em *nudges* e em alterações na arquitetura de escolhas de maneira transparente e evitando a manipulação. Essas medidas permitiriam que os decisores tivessem uma compreensão substancial das suas decisões, ao permitir o adequado exercício das suas capacidades de direcionar a atenção e de inibição. Uma

ressalva, contudo, é necessária. Ao longo do Capítulo 5, demonstrou-se que a aplicação prática de *nudges* em larga escala cria desafios adicionais. Testar a relação entre transparência e efetividade em experimentos de larga escala permitiria que esses dados fossem validados de maneira mais confiável. Essa é mais uma importante agenda de pesquisas a ser promovida.

Os problemas ligados à doutrinação, por outro lado, podem prejudicar o exercício fácil das capacidades de direcionar a atenção e de inibição. Mais do que isso, se o sujeito é guiado por preferências de terceiros para tomar uma determinada decisão, existe o risco de que suas ações sejam heterônomas, por não serem substancialmente intencionais.

Como se argumentou no Capítulo 2, na abordagem adotada nesta dissertação, a intencionalidade deve ser analisada em graus, a depender da medida em que as decisões são tomadas de maneira mais automática e mais distante das preferências de segunda ordem ou julgamentos diretos do decisor. Com relação a esse terceiro conjunto de críticas, a questão, então, é definir em que medida os *nudges* podem distanciar o sujeito das suas preferências de segunda ordem. A ideia, como discutido no Capítulo 2, está intimamente associada à noção de autenticidade.

Nesse contexto, um primeiro ponto que deve ser abordado diz respeito à possibilidade de os *nudges* corrigirem a estrutura de preferências dos indivíduos. Como apontado por Bovens (2009, p. 211–213), em diversas situações, os *nudges* podem simplesmente contribuir para que o indivíduo tome decisões que gostaria de tomar, mas não o faz por inércia, ignorância, acrasia ou mal-estar. Os fatores supostamente irrelevantes, aqui, contribuem para tal dificuldade. Em relação a esses casos, os *nudges* aprimoram a autocracia dos indivíduos e podem, de fato, contribuir para realização prática de atos substancialmente intencionais. Em outros termos, com relação a esses casos, *nudges* podem promover atos autônomos.

Quanto às pessoas que são empurradas a uma direção contrária às suas preferências, algumas situações podem ser delineadas.

Em primeiro lugar, como já discutido no Capítulo 5 e na Seção 6.2, os *nudges* contêm um mecanismo de autocontenção: diante de fortes preferências antecedentes, eles acabarão falhando (SUNSTEIN, 2019a, p. 99). Se um ateu convicto é empurrado para pagar o dízimo, é improvável que ele continue a seguir esse rumo de ação, ainda que fatores supostamente irrelevantes impactem as suas escolhas. Nessa hipótese, é improvável que ele tome decisões que não sejam substancialmente intencionais após ser influenciado por um *nudge*.

Por outro lado, há inúmeras situações em que as pessoas podem não ter preferências bem estabelecidas. A proteção de dados pessoais parece ser um exemplo que ilustra essa

afirmação[372]. Embora as pessoas valorizem a privacidade, pode ser que, por diversas razões, não tenham preferências formadas sobre o assunto. No caso concreto, estarão mais suscetíveis aos efeitos de *nudges*. Essa maior suscetibilidade de algumas pessoas aos *nudges* inviabilizaria a aplicação de tal estratégia de política pública? Se as preferências do indivíduo se modificarem após a influência da política específica, tal modificação de preferências seria legítima?

Quanto aos dois últimos casos, isto é, aqueles que não têm preferências antecedentes bem definidas ou que passam por uma mudança de preferências após a implementação da política, entender os perímetros estabelecidos pela intencionalidade, de maneira análoga ao exercício feito por Engelen e Nys (ENGELEN; NYS, 2020) com relação à sua definição de autonomia – aqui, incorporada à análise da intencionalidade – pode auxiliar a aliviar um pouco das preocupações.

Como afirmam Engelen e Nys (2020), mesmo se o indivíduo não tiver uma preferência bem estabelecida – por exemplo, sobre ser ou não um doador de órgãos –, existem valores com os quais ele se preocupa e que constituem quem ele é. Engelen e Nys argumentam que esses valores – que podem ser vistos como preferências de segunda ordem ou como julgamentos diretos, na expressão de Bernheim – estabelecem a base em que são formadas as preferências dos indivíduos em circunstâncias mais específicas. Em outras palavras, a intencionalidade estabelece um perímetro dentro do qual múltiplas preferências e decisões específicas podem ser vistas como substancialmente intencionais.

Os *nudges*, como ponderam Engelen e Nys, raramente ultrapassam os limites estabelecidos pelo perímetro da intencionalidade. Ainda que as preferências específicas ou julgamentos indiretos dos sujeitos sejam empurrados, é bastante razoável supor que ele permanecerá tomando decisões que se conformam aos valores, preocupações e interesses que constituem a sua personalidade; que fazem com que seus atos sejam no mínimo substancialmente intencionais.

Para ilustrar o seu argumento, Engelen e Nys (2020) trazem o exemplo de um sujeito ambientalista que valoriza a vida animal e se opõe à crueldade contra os animais. Em um supermercado, esse sujeito, vegetariano, poderia fazer diversas escolhas que estariam dentro dos perímetros de sua intencionalidade, como os mais diversos tipos de vegetais e frutas. Caso, no entanto, o sujeito fosse levado a comprar bifes *T-bone* ou costeletas de cordeiro, essa decisão deixaria de ser substancialmente intencional. Esse cenário dificilmente seria provocado somente pelos *nudges*.

---

[372] Vide nota 301 *supra*.

De todo modo, mais uma vez na linha de Engelen e Nys (2020), é mais plausível supor que, em caso de uma mudança de tal magnitude - em relação aos valores que constituem a sua própria personalidade - o indivíduo afetado pela política tenha simplesmente aceitado uma modificação nos seus valores. As pessoas mudam o tempo todo e, raramente, tais mudanças ocorrem por processos sobre os quais elas têm controle.

Também é preciso considerar que as pessoas estão submetidas a influências as mais diversas, na tentativa de moldar as suas preferências. Propagandas e anúncios são claros exemplos disso. O equilíbrio da pescaria, como já amplamente discutido nesta dissertação, faz com que as empresas tentem guiar o comportamento dos decisores pela exploração de fatores supostamente irrelevantes. Nesse caso, ainda que tenham dificuldades para gerar ganho de bem-estar nesses contextos, os *nudges* poderiam ser elaborados para facilitar que essas pessoas atinjam os fins que estabelecem de forma intencional, isto é, para aprimorar a sua autocracia.

Com base na análise de cada um dos elementos que compõem um ato autônomo na concepção adotada nesta dissertação, na medida em que forem transparentes e não ampliarem excessivamente as fricções associadas às opções contrárias, *nudges* podem influenciar o comportamento e manter atos substancialmente autônomos por parte dos decisores.[373] Quanto mais um *nudge* se afasta dessa decisão, maior será o grau de heteronomia do ato do decisor e, no limite, pode ser que ele deixe de ser substancialmente autônomo. Nessa hipótese, *nudges* seriam eticamente reprováveis, em termos de autonomia.

De qualquer forma, considerando que a arquitetura de escolhas é onipresente, os fatores supostamente irrelevantes podem reduzir o grau de autonomia de um ato praticado por um agente, ainda que sem a influência de um *nudge*. Nesse contexto, alterar a arquitetura de escolhas para promover atos autônomos pode ser um caminho adequado a ser seguido em termos de políticas públicas, de maneira a permitir que o sujeito tenha maior compreensão e maior controle sobre as suas próprias decisões e que possa tomá-las com maior grau de intencionalidade, isto é, alinhamento às suas preferências secundárias e aos seus valores.

Há que se considerar, ainda, como Sunstein (2015a) pondera e como já indicado na Seção 6.2, que algumas pessoas escolhem não escolher e podem fazê-lo por uma escolha ativa que delega suas próprias escolhas para um terceiro. Apesar disso, instituições públicas e privadas promoveriam a escolha ativa, sob o fundamento de que escolher é algo benéfico para as pessoas. Em tais hipóteses, poderia haver uma espécie de paternalismo que exige a escolha

[373] Em uma análise semelhante, Cf.: (MILLS, 2015).

(*choice-requiring paternalismo*), o que seria reprovável em termos de autonomia. Essa é mais uma razão pela qual os *nudges* podem promover atos autônomos.

6.4. Restrições institucionais e *nudges* ilegítimos: escrutínio democrático e controle de legalidade e inconstitucionalidade

Embora superável, como se argumentou na Seção 6.2, o problema do conhecimento é um fator que pode impedir a implementação adequada de *nudges*. Afinal, o arquiteto de escolhas empurra os decisores em uma direção que eles não desejam seguir. Do mesmo modo, embora possam promover a autonomia e serem compatíveis com atos substancialmente autônomos, os *nudges* podem ser instrumentalizados para a realização de fins ilícitos - *nudges* exploratórios - e para tentar burlar o processo decisório em detrimento do próprio indivíduo. Em outros termos, os *nudges* podem agir para garantir que os decisores escolham de maneira heterônoma.

Em linha com essas preocupações, algumas críticas aos *nudges* se voltam aos aspectos ligados aos limites de intervenção do Estado e às falhas de governo. Na prática, se os governos falham e elaboram leis e políticas públicas inadequadas ou mal-intencionadas, os problemas ligados à geração de bem-estar e à autonomia são maximizados.

Um primeiro conjunto de críticas envolve a ideia da ladeira escorregadia (*slippery slope*). O argumento da ladeira escorregadia, em geral, baseia-se na noção de aversão aos extremos: como as pessoas repelem extremismos, elas aceitariam de bom grado uma intervenção que fique no meio do caminho (RIZZO; WHITMAN, 2009). Nesse caso, quando uma política mais interventiva é aceita como novo *status quo*, gradativamente o novo meio do caminho, em comparação a extremos ainda mais interventivos, passa a ser aceitável, independentemente do seu mérito. Wright e Ginsburg (2015, p. 48) capturam bem essa crítica ao argumentar que os mesmos problemas teóricos e empíricos que fundamentam a aplicação de *nudges* aumentam a probabilidade de que, após a implementação dessas políticas, a regulamentação apresente uma tendência a se expandir. Segundo esses autores, caso a intervenção não fosse exitosa, o governo poderia vê-la como insuficiente. Assim, em vez de avaliar as causas do insucesso da intervenção arquitetos de escolha poderiam simplesmente

aumentar o grau de intervenção, na medida em que isso for necessário para atingir os resultados almejados pela política.[374]

Além disso, outro conjunto de críticas aponta que *nudges* podem ser utilizados para expandir as formas tradicionais de controle do governo sobre os indivíduos, de forma a disciplinã-los pelas suas próprias escolhas, aparentemente, livres.[375] A preocupação, nesse caso, é que os *nudges* se tornem uma ferramenta que facilite a dominação dos indivíduos pelo governo e escape dos controles individuais e democráticos, uma vez que tais estratégias de políticas públicas não seriam propriamente transparentes (SCHMIDT; ENGELEN, 2020, p. 7).

Por outro lado, há autores que criticam como os arquitetos de escolha, humanos como os decisores, poderiam evitar os fatores supostamente irrelevantes que prejudicam as decisões das pessoas. Posner (1998, p. 1975), por exemplo, argumenta que o *expert* também é um homem comportamental e, assim, também se comporta de maneira imprevisível. Como ele poderia "curar a irracionalidade no irracional"?[376] No mesmo sentido, Grüne-Yanoff e Hertwig (2016, p. 166-167) argumentam que, para que os *nudges* fossem bem-sucedidos, o arquiteto de escolhas precisaria estar bem-informado a respeito dos objetivos das pessoas. Essas informações, no entanto, estariam precisamente nos domínios em que a economia comportamental encontra fatores supostamente irrelevantes. Assim, para implementar *nudges*, o arquiteto de escolhas deveria ser imune aos vieses cognitivos ou, ao menos, menos sujeito aos seus efeitos. Essa não é a realidade, contudo, segundo o argumento de Grüne-Yanoff e Hertwig (2016, p. 172-173).

Por fim, também existe um temor de que o arquiteto de escolhas seja capturado por agentes privados no processo de definição de políticas públicas envolvendo *nudges*. Gigerenzer (2015, p. 376-377), por exemplo, utilizando como exemplo cuidados com a saúde, sustenta que muitas vezes os agentes públicos têm conflitos de interesse. Para ilustrar esse ponto, o autor descreve como muitas mulheres são "empurradas" para a realização de autoexames, embora não existam evidências científicas que apontem para a redução de mortalidade por câncer de mama diante de tal prática. Nas palavras de Gigerenzer (2015, p. 377): "Aqui, nudges servem aos interesses da indústria multibilionária de mamografia e não permite que mulheres façam

[374] Saint-Paul (2011, p. 149), por outro lado, sustenta que o reconhecimento de vieses comportamentais não deveria justificar um aumento da intervenção governamental em questões privadas, a partir, por exemplo, de *nudges*. Segundo o autor, políticas paternalistas disponibilizariam para o governo instrumentos que poderiam ser utilizados para perseguir a agenda de grupos de interesse privados em detrimento do eleitorado. Além disso, como o governo é composto por pessoas que também estão sujeitas a vieses comportamentais e poderiam cometer erros sistemáticos na sua tomada de decisões, delegar decisões para o governo não eliminaria a possibilidade de decisões inadequadas.

[375] Cf., *e.g.*, (JONES; PYKETT; WHITEHEAD, 2011; LEGGETT, 2014).

[376] Tradução livre. Texto original: "[...] *curing irrationality in the irrational?"*.

escolhas informadas"[377]. Rebonato (2014, p. 389-390), por sua vez, sustenta que Thaler e Sunstein supõem que os governantes seriam sempre benevolentes, no sentido de que teriam em vista apenas o bem-estar da população. No entanto, como argumentam James M. Buchanan e Gordon Tullock (1999, p. 282), no seu trabalho clássico, a pressão de grupos de interesse que buscam influenciar o resultado de assuntos específicos debatidos em assembleias legislativas lança dúvidas sobre a premissa de que o legislador possa priorizar o interesse público.

Responder a essas críticas é uma tarefa relativa à Teoria da Escolha Pública e à economia política de maneira geral. A Teoria da Escolha Pública Comportamental, inclusive, é uma nova de linha de pesquisas em economia comportamental que vem se desenvolvendo.[378] Evidentemente, em linha com o marco teórico adotado nesta dissertação e com o problema de pesquisa investigado na pesquisa, responder em detalhes todas essas críticas ultrapassa o escopo desta dissertação. De todo modo, posicionar-se minimamente sobre elas e indicar como elas parecem impactar a utilização de *nudges* é um caminho adequado para encerrar esta dissertação.

Inicialmente, com relação à ladeira escorregadia, Thaler e Sunstein (2021, p. 316) argumentam, acertadamente, que não há qualquer evidências de que tomar uma determinada medida X, como a adoção de um *nudge*, leve a uma maior probabilidade ou torne inevitável a adoção de medidas Y e Z, como técnicas de comando e controle coercitivas. Conly (2013, p. 114) acrescenta que o fato de não haver uma distinção clara entre A e B e entre B e C, e assim sucessivamente, não implica que A e Z são absolutamente diversos. Há uma variação clara entre *nudges* e técnicas coercitivas clássicas. Por mais que o arquiteto de escolhas possa ampliar as fricções associadas às decisões contrárias ao *nudge*, como já mencionado, essa técnica ainda é distante de intervenções coercitivas clássicas, baseadas na sanção e na coerção do Estado.

O risco de dominação e controle por parte de *nudges* também pode ser minimizado. Como sustenta Schmidt (2017), o poder de *nudge* não é uma fonte de *controle* se esse poder é transparente e democrático. Além de ser implementado por democracias republicanas, submetidas a um processo eleitoral, Schmidt aponta ser importante a existência de uma constituição com previsão de freios e contrapesos e a possibilidade de contestação de medidas públicas pela população. A transparência, nesse caso, é fundamental. Para além da própria intervenção em si, é essencial que o governo deixe um registro público dos *nudges* que implementa e dos seus resultados (IVANKOVIĆ; ENGELEN, 2019). Essa medida, afinal, potencializa a análise de custos e benefícios das medidas e a apuração dos seus eventuais

[377] Tradução livre. Texto original: "*Here, nudging serves the interests of the multibillion mammography industry and does not enable women to make informed decisions*".
[378] Cf., nesse sentido: (SCHNELLENBACH; SCHUBERT, 2014; SCHUBERT, 2017).

prejuízos, além de possibilitar um contato mais direto entre a academia e os *policy-makers*, o que contribui com a obtenção de evidências baseadas em políticas públicas, como argumentado no Capítulo 5.[379] Esse é um argumento que vale para todos os problemas discutidos nesta seção.

Quanto aos vieses dos arquitetos de escolha, é verdade que, como humanos, eles também estarão submetidos a fatores supostamente irrelevantes. Apesar disso, como Conly (2013, p. 116) menciona, é muito mais provável que o decisor seja tentado a agir de forma imprudente - ou automática - do que o arquiteto de escolhas.

Em um sentido semelhante, Sunstein (2014a, p. 120) pondera que o trabalho de agentes públicos é basear-se em processos deliberativos - Sistema 2 - e análises de custos e benefícios para tomar as suas decisões. Por mais que o Sistema 1 e os processos automáticos possam guiar essa tomada de decisões, a exigência - ao menos teórica ou legal[380] - de que agentes públicos se atentem aos custos e benefícios pode servir como medida de desenviesamento. Em outras palavras, embora esteja submetido a vieses, o arquiteto de escolhas decide em condições epistemologicamente mais apropriadas do que o decisor.

Por fim, quanto à possibilidade de captura do arquiteto de escolhas, é necessário frisar que esse é um risco inerente a qualquer decisão pública a ser tomada em democracias representativas. A questão é, portanto, eminentemente institucional e decorre da própria imperfeição do sistema. Qualquer legislação que for discutida e eventualmente aprovada estará sujeita a esse risco. Como Conly (2013) pondera, tal imperfeição não justifica, no entanto, o abandono da legislação É melhor legislar do que abandonar as leis em razão da imperfeição do sistema. A mesma afirmação vale à elaboração de *nudges*.

Aqui, valem, ainda, as considerações de Abhijit Banerjee e de Duflo a respeito do impacto da economia política na elaboração de políticas públicas. Não existe medida imune a erros. Como Abhijit Banerjee e Duflo (2021, p. 268) sustentam, "[a]té mesmo as políticas públicas mais bem-intencionadas e bem pensadas podem não causar impacto se não forem implementadas de forma adequada". Todas as críticas expostas acima evidenciam potenciais falhas no processo de tomada de decisões públicas, que podem levar a erros e políticas públicas inadequadas.

---

[379] A prática das *nudge units* dos Estados Unidos que foram objeto de análise por DellaVigna e Linos (2022), no experimento já mencionado no Capítulo 5, é um exemplo a ser seguido nesse sentido. Além dos *nudges* serem públicos, há um vasto banco de dados com seus efeitos, a população atingida e inúmeras outras variáveis que podem ser consideradas na análise de custos e benefícios da medida.

[380] Nesse sentido, o Decreto-Lei n.º 4.657, de 4 de setembro de 1942, a Lei de Introdução às normas do Direito Brasileiro, prescreve, no seu art. 20, que: "[n]as esferas administrativa, controladora e judicial, não se decidirá com base em valores jurídicos abstratos sem que sejam consideradas as consequências práticas da decisão" (BRASIL, 1942).

Esperar que o Brasil desenvolva instituições econômicas e políticas inclusivas[381] – para usar a terminologia da Nova Economia Institucional de Acemoglu e Robinson (ACEMOGLU; ROBINSON, 2012, 2019) –, para que os *nudges* possam ser bem aplicados não é razoável. Como ponderam Abhijit Banerjee e Duflo (2021, p. 270), "o progresso gradual e o acúmulo de pequenas mudanças podem, às vezes, terminar numa revolução silenciosa". A ideia de Abhijit Banerjee e Duflo, alinhada com o seu objeto de pesquisas - melhorar a vida das pessoas mais pobres - é partir de uma *"*visão de baixo". Por mais que as instituições sejam essenciais para garantir o desenvolvimento econômico, aprimoramentos graduais em termos de políticas públicas podem gerar ganhos de bem-estar. Entram, aqui, mudanças na margem e a descentralização e a democracia exercida na prática (BANERJEE; DUFLO, 2021). Os *nudges* são parte dessas mudanças.

De todo modo, para suavizar as críticas de caráter institucional, também é necessário pontuar que a utilização de *nudges* injustos pode ser coibida a partir do controle jurisdicional. Como Saraiva (2020, p. 38) pondera, por mais que se verifique uma tendência paternalista em Portugal – análise que também pode ser aplicada ao Brasil -, a Constituição portuguesa veda "soluções que promovam fins ilícitos e que sejam manipulativas, enganadoras, opacas e desproporcionadas". Embora o aprofundamento dessa análise ultrapasse o escopo desta dissertação, o fato é que os *nudges* podem ser submetidos ao teste da proporcionalidade.

O teste de proporcionalidade é o critério decisivo para identificar uma restrição ao exercício dos direitos fundamentais como compatível ou não com a Constituição (SILVA, 2021, p. 120). Na jurisprudência do STF, o teste da proporcionalidade é estruturado em três passos – adequação, necessidade e proporcionalidade em sentido estrito – que nada mais são do que "perguntas a serem respondidas no processo de controle da constitucionalidade das restrições aos direitos fundamentais" (SILVA, 2021, p. 120). Nesse caso, a submissão de um *nudge* ao teste da proporcionalidade envolverá a avaliação da medida em que: a) o objetivo pretendido pela política possa ser alcançado ou fomentado a partir da utilização de *nudges* (adequação); b) não há alternativas menos interventivas que possam ter efetividade semelhante

[381] Acemoglu e Robinson destacam o aprimoramento das instituições econômicas e políticas como o caminho que leva ao desenvolvimento econômico. As instituições econômicas são as "regras que regem o funcionamento da economia e dos incentivos que motivam a população" (ACEMOGLU; ROBINSON, 2012, p. 58). Exemplos são a proteção da propriedade privada e a segurança jurídica. As instituições políticas, por sua vez, são "as regras que regem os incentivos políticos" e "definem quem são os detentores de poder na sociedade e para que fins ele pode ser utilizado" (ACEMOGLU; ROBINSON, 2012, p. 62–63). Aqui, entram as ideias relacionadas à democracia, como o sufrágio universal. Tanto as instituições econômicas, quanto as instituições políticas podem ser inclusivas, quando estimulam a participação geral da população em atividades econômicas e são pluralistas, ou extrativistas, quando se voltam à extração de renda da população por uma elite.

aos *nudges*; e c) os motivos que fundamentam a adoção de um *nudge* têm peso suficiente para justificar a restrição que eles representam à autonomia e aos eventuais direitos fundamentais.

A análise empreendida nesta dissertação fornece subsídios que podem auxiliar nessa empreitada - como uma visão geral da efetividade dos *nudges*, a definição das estratégias alternativas aos *nudges*, a medida em que *nudges* podem gerar ganhos de bem-estar e a relação entre *nudges* e a autonomia. De qualquer forma, a constitucionalidade de um *nudge* deve ser apreciada à luz da política específica no caso concreto.

Apesar disso, como aponta Rute Saraiva (2020), em abstrato, essa constitucionalidade deve ser apreciada em três perspectivas, sempre à luz da proporcionalidade: a) em relação às pessoas diretamente afetadas pelo *nudge*; b) em relação aos agentes "racionais" que não precisam de empurrões; e c) quanto aos arquitetos de escolha.

A constitucionalidade dos *nudges* pode ser questionada, em primeiro lugar, sob o ponto de vista das pessoas que são empurradas. De fato, há que se avaliar se, pelo *nudge*, há uma negação da autodeterminação na fixação das suas preferências, uma manipulação das suas opções ou uma limitação da sua aprendizagem. A avaliação do *nudge*, à luz do princípio da proporcionalidade, envolveria a comparação, na situação concreta, entre os vários métodos de intervenção, o seu impacto na esfera privada e a sua efetividade, a fim de legitimar a intervenção menos intrusiva.

Uma avaliação interessante da constitucionalidade dos *nudges*, sob a perspectiva das pessoas que são empurradas, foi feita por Anne van Aaken (2015, 2016). Van Aaken (2016, p. 194-195) afirma ser necessário distinguir entre *nudges* que tentam mudar preferências e enfrentam, assim, a autonomia (paternalismo de fim) de *nudges* que promovem a cognição (paternalismo de meio). Para a autora, sempre que possível, um *nudge* deve ativar formas reflexivas de se pensar e, se o objetivo é aumentar a autonomia do agente decisor, manipulações invisíveis do ambiente de escolhas não devem ser consideradas medidas proporcionais. A avaliação da permissibilidade dos *nudges* à luz da proporcionalidade, nesse caso, aumentaria a transparência dos *nudges* - já que atribuiria ao regulador o ônus de provar a proporcionalidade da sua medida -, além de refinar e expandir a escolha de instrumentos regulatórios.

Em outra perspectiva, a constitucionalidade de *nudges* também pode ser questionada à luz das pessoas que não precisariam ser alvo de tais medidas. Afinal, há agentes racionais ou quase racionais que já decidiriam bem sem o *nudge*. Podemos pensar, por exemplo, no sujeito que é um *expert* em finanças e, sem qualquer intervenção, saberia tomar boas decisões no que diz respeito a poupança de recursos para o futuro. Esse sujeito, não precisaria de um *nudge*. Nesse caso, a avaliação à luz do princípio da proporcionalidade, como ensina Saraiva (2020),

deve passar pela avaliação da existência de custos desnecessários, inadequados ou excessivos, especialmente em comparação com medidas menos onerosas e restritivas, a eficácia da medida e a sua aceitabilidade. A mesma afirmação vale para a avaliação da constitucionalidade dos *nudges* em relação aos arquitetos de escolha privados, no caso de *nudges* impuros.

Por fim, para encerrar esta breve revisão bibliográfica da constitucionalidade dos *nudges*, convém mencionar ainda a obra de Mark Schweizer (2016). Em primeiro lugar, Schweizer afirma que os *nudges*, por si só, não interferem na autonomia, já que não existe algo como um contexto *desenviesado*. Apesar disso e na linha já mencionada acima, os *nudges* podem envolver medidas coercitivas contra pessoas que vão além do decisor e, por consequência, podem originar questionamentos constitucionais por interferir na liberdade desses terceiros. Ao mesmo tempo, embora Schweizer afirme que *nudges* paternalistas não parecem servir a um fim legítimo[382] que justificaria a interferência em um direito fundamental, a não ser que o dano prevenido sejam sérias lesões corporais ou a morte, o autor reconhece que a Corte Constitucional da Alemanha costuma permitir intervenções paternalistas no domínio da proteção ao consumidor e aos empregados. Por fim, em linhas gerais, Schweizer afirma que os *nudges* seriam considerados medidas adequadas e necessárias. Quanto à proporcionalidade em sentido estrito (*balancing*), o autor ressalta que a avaliação em abstrato é mais complexa, mas afirma que, em geral, os *nudges* devem ser bem-sucedidos nessa etapa.

O controle constitucional – difuso ou concreto – de *nudges* específicos não é uma questão meramente teórica. A forma de questionar a constitucionalidade de um *nudge* dependerá, evidentemente, da maneira como ele é inserido na ordem jurídica. Como se demonstrou no Capítulo 4, os *nudges impuros*, em geral, são implementados a partir de determinações legais, regras regulatórias, editais de licitação ou sanções. Os *nudges puros*, por seu turno, costumam tomar a forma de determinações ou atos administrativos típicos.

Embora não seja o objetivo desta dissertação aprofundar-se nos desdobramentos dessa questão, há casos já julgados pelo Supremo Tribunal Federal envolvendo a utilização de *nudges* e há casos pendentes de julgamento.

No julgamento do Recurso Extraordinário 1.070.522/PE, por exemplo, o Ministro Luiz Fux, relator do recurso, avaliou a constitucionalidade de procedimentos licitatórios que exigem percentuais mínimos e máximos a serem observados por emissoras de rádio na produção e

[382] Embora não seja uma etapa do teste da proporcionalidade adotado pelo STF, a legitimidade dos fins que a medida pretende atingir é identificada como um elemento adicional do teste da proporcionalidade – precedente à adequação, à necessidade e à proporcionalidade em sentido estrito - por autores ligados à aplicação desta regra perante a Corte Europeia de Direitos Humanos (AFONSO DA SILVA, 2002, p. 35).

transmissão de programas culturais, artísticos e jornalísticos locais. Na sua visão, o aumento da oferta de programas locais seria um *nudge* e, portanto, não seria uma intervenção exagerada na vida privada. Não houve, porém, uma avaliação da medida em que os *nudges* podem violar a autonomia em determinadas circunstâncias e podem ser um fardo exagerado ao arquiteto de escolhas privado - a emissora de rádio -, sobretudo à luz das críticas aos *nudges* já discutidas nesta dissertação.

Por outro lado, o Partido Socialismo e Liberdade (PSOL), representado por seus advogados André Maimoni, Alberto Maimoni e Álvaro Maimoni, ajuizou a Ação Direta de Inconstitucionalidade n.º 5502/DF, em trâmite no Supremo Tribunal Federal, para sustentar que a presunção do consentimento dos servidores para a adesão aos regimes de previdência complementar da União - tema já discutido na Seção 6.3 - viola o art. 202 da Constituição da República[383], uma vez que desnatura o caráter facultativo dessa adesão. De acordo com a petição inicial, facultativo seria "aquilo que se permite que seja feito, que depende da vontade, que não é obrigatório, ou seja, tudo o que não for imposto, porque depende de uma conduta comissiva (manifestação de vontade) daquele indivíduo a quem se atribuiu a faculdade" (BRASIL, 2016).

Em linha com o que já foi discutido nesta dissertação, a facultatividade de uma determinada opção não se confunde com a necessidade de que a adesão a ela tenha sido expressa. Em termos legislativos, não há qualquer óbice para que os servidores optassem por sair do regime. O problema, como mencionado na Seção 6.3, é a criação de fricções desnecessárias para o exercício dessa opção de saída. Embora essa questão não possa ser discutida pela via do controle concentrado de constitucionalidade, já que essas fricções são determinadas por uma orientação normativa, a exigência pode ser desmedida e pode tornar o *nudge* ilegítimo em termos normativos[384]. De todo modo, a discussão empreendida nesta dissertação auxilia a delimitar quais os problemas da presunção do consentimento e como é possível elaborar políticas públicas com maior legitimidade a partir do reconhecimento de tais problemas.

---

[383] "Art. 202. O regime de previdência privada, de caráter complementar e organizado de forma autônoma em relação ao regime geral de previdência social, será facultativo, baseado na constituição de reservas que garantam o benefício contratado, e regulado por lei complementar. (Redação dada pela Emenda Constitucional nº 20, de 1998)" (BRASIL, 1988)

[384] Essa é, no entanto, apenas uma especulação. No caso concreto, a exigência de apreciação prévia para o *opt-out* pode ter justificativas legítimas, uma vez que a desistência não é simplesmente o *opt-out*. De fato, de acordo com a legislação aplicável, o sujeito pode desistir da adesão automática em até 90 dias após essa adesão e ter a restituição das contribuições feitas nesse período. Sem a apreciação por um órgão público, seria mais difícil aferir se a desistência foi feita no prazo correto e apurar o montante a ser restituído.

6.5. Principais conclusões do Capítulo 6

Este Capítulo enfrentou a discussão normativa que envolve a aplicação de *nudges*. Basicamente, foram discutidos dois conjuntos mais amplos de críticas à utilização dessa técnica de intervenção, ligados, respectivamente, ao bem-estar e à autonomia. Inicialmente, foram enfrentadas questões ligadas aos fundamentos que justificam a implementação de um *nudge*. Em um primeiro aspecto, foi discutida a necessidade de utilização de *nudges* em um contexto de racionalidade ecológica e de problemas econômicos estruturais. Em seguida, discutiu-se como é possível aplicar os *nudges*, apesar do problema do conhecimento e da dificuldade inevitável de se compreender quais opções ampliam o bem-estar das pessoas. Feita essa exposição, foram enfrentadas as críticas ligadas a uma possível heteronomia de decisões influenciadas por um *nudge*. Por fim, foram tecidas breves considerações a respeito das questões institucionais que envolvem a aplicação de *nudges*, que estão, em geral, ligadas à possibilidade de que os arquitetos de escolha errem ou estejam mal-intencionados. As principais conclusões deste Capítulo foram as seguintes:

1. Os *nudges* são compatíveis com a racionalidade ecológica e suas justificativas não dependem do *Homo economicus*. Problemas estruturais devem ser enfrentados com as ferramentas tradicionalmente disponíveis ao regulador para lidar com essas questões. Os nudges não alteram essa conjuntura. No entanto, a racionalidade ecológica deve ser entendida em termos procedimentais. A partir dessa concepção, verifica-se que nem sempre as pessoas empregam os procedimentos adequados à tomada de decisão em um determinado ambiente. A contribuição ecológica dos *nudges*, nesse caso, envolve a adaptação dos ambientes de escolha aos procedimentos de tomada de decisão efetivamente empregados pelos decisores, o que permite a prevenção, a correção ou a reação a arquiteturas de escolha que gerem correspondências nocivas;
2. O problema do conhecimento ilustra dificuldades que podem gerar falhas na aplicação de *nudges*, mas não inviabiliza normativamente a utilização dessa técnica de intervenção. Escolhas ativas – aprimoradas por *boosts* ou por técnicas de desenviesamento – nem sempre serão uma alternativa viável, sobretudo em um contexto de limitada largura de banda. Ao mesmo tempo, na prática, é questionável a medida em que o paternalismo libertário enfrentaria verdadeiros questionamentos. Além de medidas paternalistas coercitivas serem comuns no Brasil, as pessoas parecem gostar de *nudges*, o que pode indicar a aceitabilidade de tais medidas. De todo modo, é preciso

destacar que, enquanto julgamentos indiretos são consistentemente alterados por fatores supostamente irrelevantes, os julgamentos diretos são mais constantes. Em condições epistemologicamente favoráveis, o arquiteto de escolhas pode perguntar às pessoas a forma pela qual elas gostariam de decidir e, assim, inferir quais seriam as escolhas de agentes informados e desenviesados. Nesse caso, análises de custos e benefícios podem indicar se o problema do conhecimento inviabiliza ou não um *nudge* concreto;

3. Os nudges podem promover atos substancialmente autônomos, mas também podem ser incompatíveis com a autonomia à luz da ideia de fácil resistibilidade. Se *nudges* ampliarem excessivamente as fricções relativas associadas às opções contrárias, o resultado pode ser quase coercitivo e os *nudges* podem ser medidas controladoras. A falta de transparência, por sua vez, impede que as pessoas tenham substancial compreensão da decisão influenciada por um *nudge* e leva à manipulação do comportamento. A transparência, ainda, não parece ser incompatível com a efetividade dos *nudges*. Além disso, dificilmente os *nudges* levariam a decisões substancialmente incompatíveis com as preferências de segunda ordem das pessoas. Por mais que possam criar uma falha na integração da estrutura de preferências das pessoas, *nudges* dificilmente fariam com que elas escolhessem opções que lhes fossem absolutamente repugnantes;

4. A falibilidade do arquiteto de escolhas - seja em razão de seus próprios vieses cognitivos, seja pela influência do poder econômico na tomada de decisões políticas – limita os níveis de sucesso que uma política baseada em *nudges* pode atingir. No entanto, a imperfeição do sistema não inviabiliza a utilização dessa técnica de intervenção. As restrições institucionais da democracia, a submissão de *nudges* ao escrutínio público e a possibilidade do controle jurisdicional de *nudges* ilegítimos, a partir da regra da proporcionalidade, são amarras que contribuem para evitar esses problemas.

Em suma, os *nudges* são uma estratégia promissora à implementação de políticas públicas. Atentar-se às suas limitações permite que a sua efetividade e os seus ganhos de bem-estar sejam potencializados. Por mais que seus efeitos sejam modestos, os *nudges* podem gerar efetiva mudança de comportamento e ampliar o bem-estar individual e coletivo. Os riscos de que *nudges* gerem decisões heterônomas, por sua vez, devem ser tidos em consideração para evitar políticas ilegítimas que imponham valores para terceiros e que passem ao largo da sua consciência. Além de garantir a legitimidade dos *nudges*, essas preocupações evitam a rejeição

popular da medida que pode levar à reatância psicológica. Essas foram as principais questões evidenciadas nesta dissertação.

CONSIDERAÇÕES FINAIS

Ainda que com diversas dúvidas que apontam caminhos de pesquisa na literatura das ciências comportamentais, sobretudo associadas à economia, à filosofia política e ao direito, o problema proposto para esta dissertação foi respondido. A hipótese, de certa forma, foi confirmada, o que, no entanto, deve ser visto com parcimônia. A implementação de *nudges* em políticas públicas *pode* gerar ganhos de bem-estar, sem riscos à autonomia. Retomando a pergunta que inicia esta dissertação, o governo pode promover a mudança de comportamento pelos *nudges*.

Apesar disso, na contramão do *hype* em torno de tal medida, as situações em que *nudges* podem ser utilizados com ganhos de bem-estar e sem riscos à autonomia são mais limitadas do que pareciam à época em que a hipótese foi imaginada. Há diversas lacunas nas pesquisas que envolvem a aplicação e a elaboração de *nudges* que devem ser preenchidas para que se chegue a conclusões mais assertivas, que permitam um maior grau de previsibilidade ao arquiteto de escolhas. Antes mesmo de expor as conclusões desta dissertação, convém enumerar estes caminhos de pesquisa:

a) Os dados existentes hoje na literatura não permitem uma comparação adequada e ampla entre *nudges* e estratégias alternativas, como *boosts*, *thinks*, *nudges plus* e *sludges*. Entender as diferenças de efetividade entre essas medidas e os desafios para implementá-las em larga escala é uma etapa essencial para que o arquiteto de escolhas possa avaliar os custos e benefícios de cada uma dessas estratégias;

b) Como qualquer ideia ou política pública, os *nudges* encontram limitações para implementação em larga escala. Políticas públicas reais não são feitas em placas de petri. Testar diferentes tipos de *nudges*, em diferentes contextos e em larga escala é uma etapa essencial para compreender as verdadeiras medidas de ganho de bem-estar dos *nudges*;

c) Do mesmo modo, mais avaliações da persistência dos efeitos dos *nudges* são fundamentais. Os dados parecem indicar que a persistência é possível, a partir de uma combinação de mudança de hábitos e de adaptações tecnológicas, mas não é exatamente claro como gerar esses efeitos de transbordamento. Assim, também é essencial entender os efeitos de *nudges* em longo prazo;

d) A reatância psicológica pode ser uma dura barreira à implementação de *nudges*, sobretudo em contextos polarizados e de descrença governamental, como o que se identifica hoje. Verificar em que medida *nudges* podem desencadear a reatância é outra

tarefa que pode ser mais bem explorada na literatura que discute a efetividade dos *nudges*;

e) Entender, em termos práticos, o equilíbrio da pescaria também é outra tarefa essencial para a compreensão dos limites dos *nudges*. No estudo já citado de Willis (2013), parece claro que a conduta dos bancos dos Estados Unidos e as suas múltiplas alterações da arquitetura de escolhas em detrimento do *nudge* ligado aos limites de cheque especial foi essencial para a inefetividade da medida. Explorar e mensurar os efeitos de outros *nudges* impuros em cenários de desalinhamento de incentivos entre arquiteto de escolhas e governos é outro caminho de pesquisas a ser explorado;

f) Há poucos experimentos que exploram os efeitos de *nudges* combinados com outras estratégias – comportamentais ou tradicionais. Por mais que os efeitos autônomos de *nudges* sejam modestos, essa técnica de intervenção pode potencializar os efeitos de outras estratégias, em um papel complementar. Preencher essa lacuna é outra tarefa importante à compreensão da efetividade dos *nudges*;

g) As pesquisas de opiniões das pessoas a respeito dos *nudges* parecem indicar que há ampla aprovação para a maior das medidas. O conceito de *nudge* adotado em tais pesquisas, no entanto, é excessivamente amplo. Replicá-las a partir de um conceito estrito, como o adotado nesta dissertação, permite pensar se *nudges* mais interventivos - e mais efetivos - são realmente aprovados pela população;

h) Identificar as situações epistêmicas em que os julgamentos diretos das pessoas podem ser mais bem apurados também é outro caminho a ser seguido. Preferências teóricas, fora do momento de escolhas, podem ser suscitadas e explorar os contextos em que elas mais se conformam aos verdadeiros interesses dos agentes também é essencial. Na análise de custos e benefícios, o governo deve ser capaz de estimar, ao menos, o volume de pessoas cujo interesse é contrário a um *nudge* específico;

i) A literatura parece indicar que não há um conflito entre *nudges* e transparência, como Bovens (2009) sugeriu. Os estudos indicam, de fato, que tornar um *nudge* transparente não reduz a efetividade da medida. Apesar disso, esses experimentos ainda são limitados e em pequena escala. Ampliá-los e testar *nudges* obscuros e transparentes com diferentes motivações, incluindo motivações vistas como ilegítimas, pode permitir o aprofundamento dessa questão.

Feitas essas ponderações, é necessário, então, expor as conclusões desta pesquisa.

Em primeiro lugar, por mais que Gigerenzer e os seus seguidores defendam a racionalidade ecológica de processos heurísticos, em muitas situações do mundo real, esses processos decisórios levarão a escolhas inadequadas. Para muitos contextos, sobretudo decisões com implicações mais concretas em termos jurídicos e econômicos, estratégias de otimização podem ser úteis e recomendadas. Se o sujeito não consegue agir em conformidade a elas, o caminho mais adequado pode ser alterar o ambiente de escolhas e torná-lo compatível aos processos decisórios que as pessoas efetivamente utilizam. Em outras palavras, os *nudges* podem ser uma medida de promoção da racionalidade ecológica.

A literatura parece indicar, por outro lado, que, em várias situações, *nudges* conseguem provocar efetiva de comportamento, ainda que modesta. Esse efeito pode ser ainda maior se considerados: a) a possibilidade de *nudges* potencializarem os efeitos de outras estratégias e serem uma boa técnica acessória na elaboração de políticas públicas; e b) possíveis efeitos de transbordamento que sejam provocados pela utilização de *nudges*. Se um *nudge* específico cria hábitos louváveis ou faz com que o sujeito adquira novas tecnologias para implementar o comportamento influenciado de forma menos custosa, os resultados poderão ser ainda mais significativos.

Apesar disso, o problema do conhecimento incorpora um elemento de incerteza na aplicação de um *nudge*, que deve ser considerado na sua análise de custos e benefícios. Por mais que os *nudges* gerem mudança efetiva de comportamento, nem sempre o arquiteto de escolhas saberá para onde empurrar. De todo modo, a prática política de países como o Brasil - marcado por várias medidas paternalistas - parece indicar que, em concreto, os *nudges* e outras medidas de paternalismo suave talvez não gerem preocupações concretas. A não coercitividade dos *nudges*, que permite que aqueles que querem decidir de outro modo optem por outra direção, assim como as pesquisas que parecem demonstrar ampla aprovação popular à maioria dos *nudges* avaliados, são aspectos que reforçam esse argumento.

O problema do conhecimento também pode ser superado a partir da busca de condições epistemologicamente favoráveis para aferir os julgamentos diretos das pessoas. Em certas situações, sobretudo longe do momento de escolhas, os agentes conseguem expressar suas preferências teóricas. Perguntá-las pode permitir que o governo estime como um agente informado e livre de vieses agiria diante de uma determinada decisão.

Também é necessário se considerar, na análise de custos e benefícios, que *nudges* aparentemente paternalistas podem gerar benefícios sociais indiretos - são *nudges* quase welfaristas. Assim, considerando que a arquitetura de escolhas é onipresente e sempre influencia as decisões e que existe uma possibilidade - que pode ser esclarecida naquelas

condições epistemologicamente favoráveis – de que as pessoas estejam criando prejuízos a si mesmas em razão de fatores supostamente irrelevantes, há fundamentos normativos para aplicação de *nudges*, em termos de geração de bem-estar.

A postura adotada nesta dissertação, no entanto, é cética em muitos contextos. A maioria dos contextos em que *nudges* podem ser aplicados envolvem um arquiteto de escolhas privado; são *nudges* impuros. Em tais casos, se houver um desalinhamento de incentivos entre o arquiteto de escolhas e o governo que impõe a medida, o seu resultado pode ser semelhante ao *nudge* implementado nos Estados Unidos com relação ao limite de crédito especial. Se os arquitetos de escolhas privados tiverem a possibilidade de alterar cada aspecto não regulado da arquitetura de escolhas para garantir a inefetividade do *nudge*, dois resultados alternativos parecem mais prováveis: ou o *nudge* não terá qualquer efetividade em mudar o comportamento – e, logo, não irá gerar ganhos de bem-estar –, ou a política talvez fique excessivamente custosa, pois pode depender de medidas mais interventivas por parte do Estado.

Essa análise, no entanto, não é normativa, mas empírica; não abala propriamente a legitimidade dos *nudges* como técnica capaz de gerar bem-estar. Se os objetivos promovidos pela medida forem muito relevantes, talvez valha a pena ampliar o seu custo, a fim de garantir a sua efetividade. Em muitos casos, mudar o equilíbrio da pescaria não é trivial; além do *nudge*, o governo precisaria abusar da cenoura e do porrete. No caso da regulação de cigarros, por exemplo, sem regras de comando e controle – como a proibição de propagandas – ou alterações dos incentivos monetários – como impostos extrafiscais –, é improvável que alertas chamativos nos maços de cigarro tivessem alguma efetividade. As empresas, afinal, trabalhariam intensamente para expressar uma mensagem diversa em outros meios. Os fins almejados pela política, aqui, parecem justificar um maior grau de intervenção e de custos da medida.

Por outro lado, os *nudges* em si não geram riscos à autonomia. A forma de implementá-los, no entanto, pode afastar atos praticados sob a influência de um *nudge* dos limites aceitáveis de autonomia, considerando, aqui, decisões substancialmente compreendidas, intencionais e não submetidas a influências controladoras.

Como demonstrado, os *nudges* podem ser implementados com um aumento excessivo das fricções associadas às opções contrárias. No caso da inscrição automática de servidores da União ao plano de previdência complementar, por exemplo, a opção de sair do plano depende da aprovação prévia do FUNPRESP-EXE. No limite, se o aumento das fricções for excessivo, os *nudges* podem ser quase coercitivos e ser uma influência substancialmente controladora. A decisão de deixar de seguir o *nudge* não pode ser custosa como a espera do Coronel de Garcia Márquez pelo recebimento da sua pensão.

Por outro lado, se os *nudges* não forem implementados de forma transparente, existe o risco de que eles sejam manipulativos. Como técnicas que exploram fatores supostamente irrelevantes, os *nudges* podem passar despercebidos. Evitar a manipulação – e a falta de compreensão das decisões que estão sendo tomadas – passa por transparência efetiva. Não basta vincular-se ao princípio da publicidade de Rawls. Governos, afinal, conseguem defender toda sorte de medida independentemente de sua legitimidade, como o espírito dos tempos atuais não cansa de evidenciar. A transparência efetiva – epistêmica, nos termos de Hansen e Jespersen (2013) – passa por garantir que as pessoas tenham a possibilidade de compreender os fundamentos da medida e os mecanismos utilizados para mudar o comportamento.

Parece improvável, ainda, que *nudges* gerem decisões substancialmente não intencionais, ressalvados os casos de *nudges* quase coercitivos e de *nudges* puramente manipulativos; nestas últimas hipóteses, o sujeito pode acabar vinculado aos interesses de um terceiro; a questão, aqui, porém, não é exatamente a intenção, mas a existência influência controladora ou a impossibilidade de se ter compreensão de uma medida. De todo modo, em termos exclusivamente de intencionalidade, *nudges* parecem muito sutis para fazerem com que as pessoas decidam de maneira contrária às suas preferências secundárias. Como argumentado acima, a intencionalidade apresenta um perímetro dentro do qual múltiplas opções são aceitáveis – isto é, alinhadas às preferências secundárias do agente. Um *nudge* não parece ser suficiente para que um agente tome uma decisão que ele reprovaria profundamente.

Por fim, ainda que a possibilidade de implementação de *nudges* ilícitos exista, a garantia de transparência cumulada com as instituições democráticas e com o possível controle jurisdicional de *nudges* ilegítimos mitiga parte das preocupações. Não se pretende argumentar que as instituições estão funcionando plenamente; este não é o papel desta dissertação. De todo modo, defende-se que mudanças marginais e graduais são possíveis e podem gerar benefícios concretos às pessoas. Os nudges, afinal, parecem ser capazes de, em alguns casos, gerar ganhos de bem-estar e promover a autonomia.

REFERÊNCIAS

ABDUKADIROV, S. (ed.). **Nudge theory in action**: Behavioral design in policy and markets. London: Palgrave Macmillan, 2016.

ACEMOGLU, D.; ROBINSON, J. A. **Por que as nações fracassam**: as origens do poder, da prosperidade e da pobreza. Tradução de Cristiana Serra. Rio de Janeiro: Elsevier, 2012.

ACEMOGLU, D.; ROBINSON, J. A. **The narrow corridor**: States, societies, and the fate of liberty. New York: Penguin Press, 2019.

ACHEN, C. H.; LARRY, M. **Democracy for realists**: Why elections do not produce responsive government. Princeton: Princeton University Press, 2016.

ACQUISTI, A.; BRANDIMARTE, L.; LOEWENSTEIN, G. Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. **Journal of Consumer Psychology**, Fort Worth, v. 30, n. 4, p. 736-758, Oct. 2020. DOI: https://doi.org/10.1002/jcpy.1191.

AFONSO DA SILVA, V. O proporcional e o razoável. **Revista dos Tribunais**, São Paulo, v. 798, p. 23-50, 2002. Disponível em: https://constituicao.direito.usp.br/wp-content/uploads/2002-RT798-Proporcionalidade.pdf. Acesso em: 9 jul. 2022.

AKERLOF, G. A. The market for "lemons": Quality uncertainty and the market mechanism. **The Quarterly Journal of Economics**, Oxford, v. 84, n. 3, p. 488-500, Aug. 1970. DOI: https://doi.org/10.2307/1879431.

AKERLOF, G.; SHILLER, R. **Phishing for phools**: The economics of manipulation and deception. Princeton: Princeton University Press, 2015.

ALLCOTT, H.; ROGERS, T. The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. **American Economic Review**, Nashville, v. 104, n. 10, p. 3003-3037, Oct. 2014. DOI: https://doi.org/10.1257/aer.104.10.3003.

ARAD, A.; RUBINSTEIN, A. The people's perspective on libertarian-paternalistic policies. **Journal of Law and Economics**, Chicago, v. 61, n. 2, p. 311-333, May 2018. DOI: https://doi.org/10.1086/698608.

ARARIPE, C. M. R. A.; MACHADO, R. C. R. Os programas de integridade para contratação com a administração pública estadual: *nudge* ou obrigação legal?: um olhar sobre as duas perspectivas. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 386-405, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5332.

ARIELY, D. **Positivamente irracional**: os benefícios inesperados de desafiar a lógica em todos os aspectos de nossas vidas. Tradução de Afonso Celso da Cunha Serra. Rio de Janeiro: Elsevier, 2010.

ARIELY, D. **Previsivelmente irracional**. Tradução de Jussara Simões. Rio de Janeiro: Elsevier, 2008.

ARNESON, R. J. Nudge and shove. **Social Theory and Practice**, Tallahassee, v. 41, n. 4, p. 668-691, Oct. 2015. DOI: https://doi.org/10.5840/soctheorpract201541436.

ASCH, S. E. Opinions and social pressure. **Scientific American**, New York, v. 193, n. 5, p. 31-35, Nov. 1955. DOI: https://doi.org/10.1038/scientificamerican1155-31.

BABCOCK, L.; LOEWENSTEIN, G.; ISSACHAROFF, S. Creating convergence: Debiasing biased litigants. **Law and Social Inquiry**, Cambridge, v. 22, n. 4, p. 913-925, 1997. DOI: https://doi.org/10.1111/j.1747-4469.1997.tb01092.x.

BALDWIN, R. From regulation to behaviour change: Giving nudge the third degree. **Modern Law Review**, London, v. 77, n. 6, p. 831-857, Nov. 2014. DOI: https://doi.org/10.1111/1468-2230.12094.

BANCO CENTRAL DO BRASIL. **Resolução nº 3.919, de 25 de novembro de 2010**. Altera e consolida as normas sobre cobrança de tarifas pela prestação de serviços por parte das instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil e dá outras providências. Brasília, DF: Banco Central do Brasil, 2010. Disponível em: https://www.bcb.gov.br/pre/normativos/res/2010/pdf/res_3919_v4_P.pdf. Acesso em: 9 jul. 2022.

BANCO CENTRAL DO BRASIL. **Resolução nº 4.549, de 26 de janeiro de 2017**. Dispõe sobre o financiamento do saldo devedor da fatura de cartão de crédito e de demais instrumentos de pagamento pós-pagos. Brasília, DF: Banco Central do Brasil, 2017. Disponível em: https://normativos.bcb.gov.br/Lists/Normativos/Attachments/50330/Res_4549_v1_O.pdf. Acesso em: 9 jul. 2022.

BANERJEE, A. *et al*. **Selecting the most effective nudge**: Evidence from a large-scale experiment on immunization. Cambridge: National Bureau of Economic Research, 2021a. (NBER Working Paper Series, n. 28726). Disponível em: https://www.nber.org/system/files/working_papers/w28726/w28726.pdf. Acesso em: 9 jul. 2022.

BANERJEE, A. V.; DUFLO, E. **A economia dos pobres**: uma nova visão sobre a desigualdade. Tradução de Pedro Maia Soares. Rio de Janeiro: Zahar, 2021.

BANERJEE, S. *et al*. Choice architecture 2.0?: Embedding reflection in behavioural cues to promote sustainable diets. *In*: IAREP SABE 2021 Virtual Conference. [*S. l.*]: IAREP SABE, 2021b.

BANERJEE, S.; JOHN, P. Nudge plus: Incorporating reflection into behavioral public policy. **Behavioural Public Policy**, Cambridge, p. 1-16, 16 Apr. 2021. DOI: https://doi.org/10.1017/bpp.2021.6.

BANG, H. M.; SHU, S. B.; WEBER, E. U. The role of perceived effectiveness on the acceptability of choice architecture. **Behavioural Public Policy**, Cambridge, v. 4, n. 1, p. 50-70, 2020. DOI: https://doi.org/10.1017/bpp.2018.1.

BARTLING, B.; FEHR, E.; HERZ, H. The intrinsic value of decision rights. **Econometrica**, Cleveland, v. 82, n. 6, p. 2005-2039, Nov. 2014. DOI: https://doi.org/10.3982/ECTA11573.

BARTON, A.; GRÜNE-YANOFF, T. From libertarian paternalism to nudging: And beyond. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 341-359, Sept. 2015. DOI: https://doi.org/10.1007/s13164-015-0268-x.

BECKER, G. S. **The economic approach to human behavior**. Chicago: The University of Chicago Press, 1976.

BEGGS, J. N. Private-sector nudging: The good, the bad, and the uncertain. *In*: ABDUKADIROV, S. (ed.). **Nudge theory in action**: Behavioral design in policy and markets. London: Palgrave Macmillan, 2016. p. 125-158.

BENARTZI, S. *et al*. Should governments invest more in nudging?. **Psychological Science**, New York, v. 28, n. 8, p. 1041-1055, Aug. 2017. DOI: https://doi.org/10.1177/0956797617702501.

BENARTZI, S. *et al*. The law and economics of company stock in 401 (k) plans. **Journal of Law and Economics**, Chicago, v. 50, n. 1, p. 45-79, Feb. 2007. DOI: https://doi.org/10.1086/508312.

BENHABIB, J.; BISIN, A.; SCHOTTER, A. Present-bias, quasi-hyperbolic discounting, and fixed costs. **Games and Economic Behavior**, [Amsterdam], v. 69, n. 2, p. 205-223, July 2010. DOI: https://doi.org/10.1016/j.geb.2009.11.003.

BERNHEIM, B. D. The good, the bad, and the ugly: A unified approach to behavioral welfare economics. **Journal of Benefit-Cost Analysis**, Cambridge, v. 7, n. 1, p. 12-68, May 2016. DOI: https://doi.org/10.1017/bca.2016.5.

BERNHEIM, B. D.; ANTONIO, R. Choice-theoretic foundations for behavioral welfare economics. *In*: CAPLIN, A.; SCHOTTER, A. (ed.). **The foundations of positive and normative economics**: A hand book. Oxford: Oxford University Press, 2011. DOI: https://doi.org/10.1093/acprof:oso/9780195328318.003.0007.

BERTHET, V.; OUVRARD, B. Nudge: Towards a consensus view?. **Psychology and Cognitive Sciences Open Journal**, Fairfax, v. 5, n. 1, p. 1-5, Aug. 2019. DOI: https://doi.org/10.17140/PCSOJ-5-143.

BESHEARS, J. *et al*. Creating exercise habits using incentives: The trade-off between flexibility and routinization. **Management Science**, Providence, v. 67, n. 7, p. 4139-4171, July 2021. DOI: https://doi.org/10.1287/mnsc.2020.3706.

BESHEARS, J.; KOSOWSKY, H. Nudging: Progress to date and future directions. **Organizational Behavior and Human Decision Processes**, San Diego, v. 161, suppl., p. 3-19, Nov. 2020. DOI: https://doi.org/10.1016/j.obhdp.2020.09.001.

BIJOS, L. M. D. J.; PROLA JÚNIOR, C. H.; TABAK, B. M. Decisões judiciais sobre os direitos dos índios: um olhar da análise econômico-comportamental do direito. **Revista Quaestio Iuris**, Rio de Janeiro, v. 10, n. 4, p. 2458-2481, 2017. DOI: https://doi.org/10.12957/rqi.2017.26868.

BIRD, K. A. *et al*. Nudging at scale: Experimental evidence from FAFSA completion campaigns. **Journal of Economic Behavior & Organization**, Amsterdam, v. 183, p. 105-128, Mar. 2021. DOI: https://doi.org/10.1016/j.jebo.2020.12.022.

BORDALO, P.; GENNAIOLI, N.; SHLEIFER, A. Salience theory of choice under risk. **The Quarterly Journal of Economics**, Oxford, v. 127, n. 3, p. 1243-1285, Aug. 2012. DOI: https://doi.org/10.1093/qje/qjs018.

BOVENS, L. The ethics of nudge. *In*: GRÜNE-YANOFF, T.; HANSSON, S. O. (ed.). **Preference change**: Approaches from philosophy, economics and psychology. [London]: Springer, 2009. p. 207-219.

BOWLES, S. **The moral economy**: Why good incentives are no substitute for good citizens. New Haven: Yale University Press, 2016.

BRACCINI NETO, M. O *nudge* como panacéia: os limites da economia comportamental a partir de sua aplicação em fundos de pensão e suas consequências ao caso brasileiro. *In*: ENCONTRO ANUAL DO ECONOMICS AND POLITICS RESEARCH GROUP (EPRG), 7., 2019, Brasília, DF. **Anais** [...]. Brasília, DF: EPRG, 2019. Disponível em: https://econpolrg.files.wordpress.com/2019/12/s2.2-paper_mariobraccinineto.pdf. Acesso em: 11 jul. 2022.

BRADT, J. Comparing the effects of behaviorally informed interventions on flood insurance demand: An experimental analysis of 'boosts' and 'nudges'. **Behavioural Public Policy**, Cambridge, v. 6, n. 3, p. 485-515, 2022. DOI: https://doi.org/10.1017/bpp.2019.31.

BRANDON, A. *et al*. **Do the effects of social nudges persist?**: Theory and evidence. Cambridge: National Bureau of Economic Research, 2022. (NBER Working Paper Series, n. 23277). DOI: https://doi.org/10.3386/w23277.

BRANSON, C. *et al*. **Acceptable behaviour**: Public opinion on behavior change policy. London: Ipsos MORI, [2012]. Disponível em: https://www.ipsos.com/sites/default/files/publication/1970-01/sri-ipsos-mori-acceptable-behaviour-january-2012.pdf. Acesso em: 10 jul. 2022.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil**. Brasília, DF: Presidência da República, 1988. Disponível em: http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 10 jul. 2022.

BRASIL. **Decreto-Lei nº 4.657, de 4 de setembro de 1942**. Lei de Introdução ao Código Civil Brasileiro. Rio de Janeiro, RJ: Presidência da República, 1942. Disponível em: http://www.planalto.gov.br/ccivil_03/decreto-lei/del4657.htm. Acesso em: 10 jul. 2022.

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Brasília, DF: Presidência da República, 2002. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/2002/l10406compilada.htm. Acesso em: 10 jul. 2022.

BRASIL. **Lei nº 12.529, de 30 de novembro de 2011**. Estrutura o Sistema Brasileiro de Defesa da Concorrência; dispõe sobre a prevenção e repressão às infrações contra a ordem econômica;

altera a Lei nº 8.137, de 27 de dezembro de 1990, o Decreto-Lei nº 3.689, de 3 de outubro de 1941 - Código de Processo Penal, e a Lei nº 7.347, de 24 de julho de 1985; revoga dispositivos da Lei nº 8.884, de 11 de junho de 1994, e a Lei nº 9.781, de 19 de janeiro de 1999; e dá outras providências. Brasília, DF: Presidência da República, 2011. Disponível em: http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12529.htm. Acesso em: 10 jul. 2022.

BRASIL. **Lei nº 6.194, de 19 de dezembro de 1974**. Dispõe sobre Seguro Obrigatório de Danos Pessoais causados por veículos automotores de via terrestre, ou por sua carga, a pessoas transportadas ou não. Brasília, DF: Presidência da República, 1974. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l6194.htm. Acesso em: 10 jul. 2022.

BRASIL. **Lei nº 6.404, de 15 de dezembro de 1976**. Dispõe sobre as Sociedades por Ações. Brasília, DF: Presidência da República, 1976. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l6404consol.htm. Acesso em: 10 jul. 2022.

BRASIL. **Lei nº 8.036, de 11 de maio de 1990**. Dispõe sobre o Fundo de Garantia do Tempo de Serviço, e dá outras providências. Brasília, DF: Presidência da República, 1990. Disponível em: http://www.planalto.gov.br/ccivil_03/leis/l8036consol.htm. Acesso em: 10 jul. 2022.

BRASIL. Ministério do Planejamento, Orçamento e Gestão. Secretaria de Gestão Pública. **Orientação Normativa nº 9, de 19 de novembro de 2015**. Estabelece orientações quanto à inscrição automática de servidores públicos da Administração Pública federal direta, suas autarquias e fundações, no plano de benefícios Execprev, da Funpresp-Exe. Brasília, DF: Secretaria de Gestão Pública, 2015. Disponível em: https://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/33339086/do1-2015-11-23-orientacao-normativa-n-9-de-19-de-novembro-de-2015-33339080. Acesso em: 10 jul. 2022.

BRASIL. Supremo Tribunal Federal. **Petição inicial da Ação Direta de Institucionalidade nº 5502/DF**. Requerente: Partido Socialismo e Liberdade (PSOL). Intdo: Presidente da República. Relator: Min. Nunes Marques. Brasília, DF, 2016. Disponível em: http://portal.stf.jus.br/processos/detalhe.asp?incidente=4967555. Acesso em: 11 jul. 2022.

BREHM, S. S.; BREHM, J. W. **Psychological reactance**: A theory of freedom and control. New York: Academic Press, 1981.

BRONCHETTI, E. T. *et al*. When a nudge isn't enough: Defaults and saving among low-income tax filers. **National Tax Journal**, Chicago, v. 66, n. 3, p. 609-634, Sept. 2013. DOI: https://doi.org/10.17310/ntj.2013.3.04.

BRUNS, H. *et al*. Can nudges be transparent and yet effective?. **Journal of Economic Psychology**, Leeds, v. 65, p. 41-59, Apr. 2018. DOI: https://doi.org/10.1016/j.joep.2018.02.002.

BUBB, R.; PILDES, R. H. How behavioral economics trims its sails and why. **Harvard Law Review**, Cambridge, v. 127, n. 6, p. 1593-1678, 2014.

BUCHANAN, J. M.; TULLOCK, G. **The calculus of consent**: Logical foundations of constitutional democracy. Indianopolis: Liberty Fund, 1999.

BURGESS, A. “Nudging” healthy lifestyles: The UK experiments with the behavioural alternative to regulation and the market. **European Journal of Risk Regulation**, Cambridge, v. 3, n. 1, p. 3-16, 2012. DOI: https://doi.org/10.1017/S1867299X00001756.

CADARIO, R.; CHANDON, P. Which healthy eating nudges work best?: A meta-analysis of field experiments. **Marketing Science**, Catonsville, v. 39, n. 3, p. 465-486, May/June 2020. DOI: https://doi.org/10.1287/mksc.2018.1128.

CALLAWAY, E. ‘Randomistas’ who used controlled trials to fight poverty win economics Nobel. **Nature**, London, 14 Oct. 2019. DOI: https://doi.org/10.1038/d41586-019-03125-y.

CAMERER, C. *et al*. Regulation for conservatives: Behavioral economics and the case for “asymmetric paternalism”. **University of Pennsylvania Law Review**, Philadelphia, v. 151, n. 3, p. 1211-1254, Jan. 2003. DOI: https://doi.org/10.2307/3312889.

CAMERER, C. F. Prospect theory in the wild: Evidence from the field. *In*: KAHNEMAN, D.; TVERSKY, A. (ed.). **Choices, values, and frames**. Cambridge: Cambridge University Press, 2000. p. 288-300.

CAMERER, C. F.; LOEWENSTEIN, G. Behavioral economics: Past, present and future. *In*: CAMERER, C. F.; LOEWENSTEIN, G.; RABIN, M. (ed.). A**dvances in behavioral economics**. Princeton: Princeton University Press, 2004. p. 3-51.

CAMERER, C. F.; LOEWENSTEIN, G.; RABIN, M. (ed.). **Advances in behavioral economics**. Princeton: Princeton University Press, 2004.

CARVALHO, R. B. Exercício de posição dominante em mercados digitais: uma análise do caso Google Shopping. *In*: SILVA, L. N. E.; TEIXEIRA, L. F. D. (ed.). **Mercados digitais**: o livro da disciplina. Belo Horizonte: Editora Expert, 2022. p. 191-250.

CLAVIEN, C. Ethics of nudges: A general framework with a focus on shared preference justifications. **Journal of Moral Education**, [Abingdon], v. 47, n. 3, p. 366-382, 2018. DOI: https://doi.org/10.1080/03057240.2017.1408577.

COASE, R. H. The problem of social cost. **The Journal of Law and Economics**, Chicago, v. 3, p. 1-44, Oct. 1960. Disponível em: https://www.law.uchicago.edu/files/file/coase-problem.pdf. Acesso em: 10 jul. 2022.

COMPETITION & MARKET AUTHORITY. **Online choice architecture**: How digital design can harm: Discussion paper. London: Crown, 2022. Disponível em: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1066524/Online_choice_architecture_discussion_paper.pdf. Acesso em: 10 jul. 2022.

CONGIU, L.; MOSCATI, I. A review of nudges: Definitions, justifications, effectiveness. **Journal of Economic Surveys**, Hoboken, June 2021. DOI: http://dx.doi.org/10.2139/ssrn.3728512.

CONGIU, L.; MOSCATI, I. Message and environment: A framework for nudges and choice architecture. **Behavioural Public Policy**, Cambridge, v. 4, n. 1, p. 71-87, Mar. 2020. DOI: https://doi.org/10.1017/bpp.2018.29.

CONLY, S. **Against autonomy**: Justifying coercive paternalism. New York: Cambridge Univeristy Press, 2013.

COSMIDES, L.; TOOBY, J. Are humans good intuitive statisticians after all?: Rethinking some conclusions from the literature on judgment under uncertainty. **Cognition**, Hague, v. 58, n. 1, p. 1-73, Jan. 1996. DOI: https://doi.org/10.1016/0010-0277(95)00664-8.

COSTA, D. L.; KAHN, M. E. Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. **Journal of the European Economic Association**, Tessenderlo, v. 11, n. 3, p. 680-702, June 2013. DOI: https://doi.org/10.1111/jeea.12011.

COSTA, E. J. F. **Levando a imparcialidade a sério**: proposta de um modelo interseccional entre direito processual, economia e psicologia. Salvador: JusPodivm, 2018.

COSTA, N. L. M. "*Nudge*" como abordagem regulatória de prevenção à corrupção pública no Brasil. **Revista de Informação Legislativa: RIL**, Brasília, DF, v. 54, n. 214, p. 91-111, abr./jun. 2017. Disponível em: https://www12.senado.leg.br/ril/edicoes/54/214/ril_v54_n214_p91. Acesso em: 10 jul. 2022.

COWEN, T. Daniel Kahneman on Cutting Through the Noise. *In*: CONVERSATIONS with Tyler, ep. 56. [*S. l.*], 19 Dec. 2018. Disponível em: https://conversationswithtyler.com/episodes/daniel-kahneman/. Acesso em: 10 jul. 2022.

CRONQVIST, H.; THALER, R. H.; YU, F. When nudges are forever: Inertia in the Swedish premium pension plan. **AEA Papers and Proceedings**, Nashville, v. 108, p. 153-158, May 2018. DOI: https://doi.org/10.1257/pandp.20181096.

DAURA, S. A. *Behavioral economics* e direito do consumidor: novas perspectivas para o enfrentamento do superendividamento. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 568-599, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5330.

DELLAVIGNA, S.; LINOS, E. RCTs to scale: Comprehensive evidence from two nudge units. **Econometrica**, Cleveland, v. 90, n. 1, p. 81-116, Jan. 2022. DOI: https://doi.org/10.3982/ECTA18709.

DHAMI, S. **The foundations of behavioral economic analysis**. Oxford: Oxford University Press, 2016.

DHAMI, S.; AL-NOWAIHI, A.; SUNSTEIN, C. R. Heuristics and public policy: Decision-making under bounded rationality. **Studies in Microeconomics**, [London], v. 7, n. 1, p. 7-58, June 2019. DOI: https://doi.org/10.1177/2321022219832148.

DINNER, I. *et al*. Partitioning default effects: Why people choose not to choose. **Journal of Experimental Psychology: Applied**, Washington, DC, v. 17, n. 4, p. 332-341, Dec. 2011. DOI: https://doi.org/10.1037/a0024354.

DJIMEU, E. W.; HOUNDOLO, D. G. Power calculation for causal inference in social science: Sample size and minimum detectable effect determination. **Journal of Development**

Effectiveness, Abingdon, v. 8, n. 4, p. 508-527, 2016. DOI: https://doi.org/10.1080/19439342.2016.1244555.

DUNNE, N. **Competition law and economic regulation**: Making and managing markets. Cambridge: Cambridge University Press, 2015.

DWORKIN, G. Paternalism. *In*: ZALTA, E. N. (ed.). **The Stanford encyclopedia of philosophy**. Stanford: Metaphysics Research Lab, 2020. Disponível em: https://plato.stanford.edu/entries/paternalism/. Acesso em: 10 jul. 2022.

DWORKIN, G. **The theory and practice of autonomy**. New York: Cambridge Univeristy Press, 1988.

EAGLEMAN, D. **Incógnito**: as vidas secretas do cérebro. Rio de Janeiro: Rocco, 2012.

EMENS, E. F. Changing name changing: Framing rules and the future of marital names. **The University of Chicago Law Review**, Chicago, v. 74, n. 3, p. 761-863, 2007. Disponível em: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=5414&context=uclrev. Acesso em: 10 jul. 2022.

ENGELEN, B. Nudging and rationality: What is there to worry?. **Rationality and Society**, [Thousand Oaks], v. 31, n. 2, p. 204-232, 2019. DOI: https://doi.org/10.1177/1043463119846743.

ENGELEN, B.; NYS, T. Nudging and autonomy: Analyzing and alleviating the worries. **Review of Philosophy and Psychology**, London, v. 11, p. 137-156, Fev. 2020. DOI: https://doi.org/10.1007/s13164-019-00450-z.

FADEN, R. R.; BEAUCHAMP, T. L. **A history and theory of informed consent**. Oxford: Oxford University Press, 1986.

FEAST and Thugs: Nudging 2020, com Cass Sunstein. [*S. l.*: *s. n.*], 2021. 1 vídeo (57 min). Publicado pelo canal GEEC Grupo de Estudos em Economia Comportamental. Disponível em: https://www.youtube.com/watch?v=Q_6HSdcUlkw. Acesso em: 12 jul. 2022.

FEINBERG, J. **Harm to self**. Oxford: Oxford University Press, 1986.

FEINBERG, J. **Harmless wrongdoing**. Oxford: Oxford University Press, 1988.

FELSEN, G.; CASTELO, N.; REINER, P. B. Decisional enhancement and autonomy: Public attitudes towards overt and covert nudges. **Judgment and Decision Making**, [*s. l.*], v. 8, n. 3, p. 202-213, 2013. DOI: https://doi.org/10.1037/t31564-000.

FISCHHOFF, B. Debiasing. *In*: KAHNEMAN, D.; SLOVIC, P.; TVERSKY, A. (ed.). **Judgment under uncertainty**: Heuristics and biases. New York: Cambridge Univeristy Press, 1982. p. 422-444.

FISCHHOFF, B.; BEYTH, R. I knew it would happen. **Organizational Behavior and Human Performance**, New York, v. 13, n. 1, p. 1-16, Feb. 1975. DOI: https://doi.org/10.1016/0030-5073(75)90002-1.

FRANKFURT, H. G. Freedom of the will and the concept of a person. **The Journal of Philosophy**, Charlottesville, v. 68, n. 1, p. 5-20, Jan. 1971. DOI: https://doi.org/10.2307/2024717.

FREDERICK, S. Measuring intergenerational time preference: Are future lives valued Less?. **Journal of Risk and Uncertainty**, Boston, v. 26, n. 1, p. 39-53, Jan. 2003. DOI: https://doi.org/10.1023/A:1022298223127.

FREDRICKSON, B. L.; KAHNEMAN, D. Duration neglect in retrospective evaluations of affective episodes. **Journal of Personality and Social Psychology**, Washington, DC, v. 65, n. 1, p. 45-55, July 1993. DOI: https://doi.org/10.1037//0022-3514.65.1.45.

FRENCH, J. Why nudging is not enough. **Journal of Social Marketing**, [Bingley], v. 1, n. 2, p. 154-62, July 2011. DOI: https://doi.org/10.1108/20426761111141896.

FREY, E.; ROGERS, T. Persistence: How treatment effects persist after interventions stop. **Policy Insights from the Behavioral and Brain Sciences**, Thousand Oaks, v. 1, n. 1, p. 172-179, 2014. DOI: https://doi.org/10.1177/2372732214550405.

FRIEDMAN, M. The methodology of positive economics. *In*: FRIEDMAN, M. (ed.). **Essays in positive economics**. Chicago: University of Chicago Press, 1966. p. 3-16.

FURNHAM, A.; BOO, H. C. A literature review of the anchoring effect. **Journal of Socio-Economics**, Amsterdam, v. 40, n. 1, p. 35-42, Feb. 2011. DOI: https://doi.org/10.1016/j.socec.2010.10.008.

GIGERENZER, G. Gigerenzer's law of indispensable ignorance. *In*: EDGE. London, 2004. Disponível em: https://www.edge.org/response-detail/10224. Acesso em: 10 jul. 2022.

GIGERENZER, G. How to make cognitive illusions disappear: Beyond "heuristics and biases". *In*: STROEBE, W.; HEWSTONE, M. (ed.). **European review of social psychology**. Hoboken: John Wiley & Sons, 1991. v. 2, p. 83-115.

GIGERENZER, G. On narrow norms and vague heuristics: A reply to Kahneman and Tversky. **Psychological Review**, Washington, DC, v. 103, n. 3, p. 592-596, 1996. DOI: https://doi.org/10.1037/0033-295X.103.3.592.

GIGERENZER, G. On the supposed evidence for libertarian paternalism. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 361-383, Sept. 2015. DOI: https://doi.org/10.1007/s13164-015-0248-1.

GIGERENZER, G. Towards a rational theory of heuristics. *In*: FRANTZ, R.; MARSH, L. (ed.). **Minds, models and milieux**: Commemorating the centennial of the birth of Herbert Simon: Archival insights into the evolution of economics. London: Palgrave Macmillan, 2016. p. 34-59.

GIGERENZER, G.; GOLDSTEIN, D. G. Reasoning the fast and frugal way: Models of bounded rationality. **Psychological Review**, Washington, DC, v. 103, n. 4, p. 650-669, Oct. 1996. DOI: https://doi.org/10.1037/0033-295x.103.4.650.

GIGERENZER, G.; SELTEN, R. (ed.). **Bounded rationality**: The adaptive toolbox. Cambridge: MIT Press, 2002.

GIGERENZER, G.; TODD, P. M. Fast and frugal heuristics: The adaptive toolbox. *In*: GIGERENZER, G.; TODD, P. M.; ABC RESEARCH GROUP (ed.). **Simple heuristics that make us smart**. Oxford: Oxford University Press, 1999. p. 3-34.

GIGERENZER, G.; TODD, P. M.; ABC RESEARCH GROUP (ed.). **Simple heuristics that make us smart**. Oxford: Oxford University Press, 1999.

GILOVICH, T.; GRIFFIN, D.; KAHNEMAN, D. (ed.). **Heuristics and biases**: The psychology of intuitive judgment. New York: Cambridge Univeristy Press, 2002.

GNEEZY, U.; RUSTICHINI, A. A fine is a price. **The Journal of Legal Studies**, Chicago, v. 29, n. 1, p. 1-17, Jan. 2000. DOI: https://doi.org/10.1086/468061.

GOLDIN, J. Which way to nudge?: Uncovering preferences in the behavioral age. **Yale Law Journal**, New Haven, v. 125, n. 1, p. 226-270, Oct. 2015. Disponível em: https://law.stanford.edu/wp-content/uploads/2016/06/Goldin-Which-Way-to-Nudge.pdf. Acesso em: 10 jul. 2022.

GOLDSTEIN, D. G.; GIGERENZER, G. The recognition heuristic: How ignorance makes us smart. *In*: GIGERENZER, G.; TODD, P. M.; ABC RESEARCH GROUP (ed.). **Simple heuristics that make us smart**. Oxford: Oxford University Press, 1999. p. 37-58.

GRAVERT, C.; OLSSON COLLENTINE, L. When nudges aren't enough: Norms, incentives and habit formation in public transport usage. **Journal of Economic Behavior and Organization**, [Amsterdam], v. 190, p. 1-14, Oct. 2021. DOI: https://doi.org/10.1016/j.jebo.2021.07.012.

GRAYOT, J. D. Dual process theories in behavioral economics and neuroeconomics: A critical review. **Review of Philosophy and Psychology**, London, v. 11, n. 1, p. 105-136, Feb. 2020. DOI: https://doi.org/10.1007/s13164-019-00446-9.

GREENE, J. D. Dual-process morality and the personal/impersonal distinction: A reply to McGuire, Langdon, Coltheart, and Mackenzie. **Journal of Experimental Social Psychology**, New York, v. 45, n. 3, p. 581-584, May 2009. DOI: https://doi.org/10.1016/j.jesp.2009.01.003.

GRÜNE-YANOFF, T. Old wine in new casks: Libertarian paternalism still violates liberal principles. **Social Choice and Welfare**, [Berlin], v. 38, n. 4, p. 635-645, Apr. 2012. DOI: https://doi.org/10.1007/s00355-011-0636-0.

GRÜNE-YANOFF, T.; HERTWIG, R. Nudge versus boost: How coherent are policy and theory?. **Minds and Machines**, [Berlin], v. 26, p. 149-183, Mar. 2016. DOI: https://doi.org/10.1007/s11023-015-9367-9.

GUALA, F.; MITTONE, L. A political justification of nudging. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 385-395, Sept. 2015. DOI: https://doi.org/10.1007/s13164-015-0241-8.

GUSTIN, M. B. S.; DIAS, M. T. F.; NICÁCIO, C. S. **(Re)pensando a pesquisa jurídica**: teoria e prática. 5. ed. rev. São Paulo: Almedina, 2020.

GUYER, P. Kant on the theory and the practice of autonomy. **Social Philosophy and Policy**, Cambridge, v. 20, n. 2, p. 70-98, July 2003. DOI: https://doi.org/10.1017/S026505250320203X.

HAGMAN, W. *et al*. Public views on policies involving nudges. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 439-453, Sept. 2015a. DOI: https://doi.org/10.1007/s13164-015-0263-2.

HAIDT, J. The emotional dog and its rational tail: A social intuitionist approach to moral judgment. **Psychological Review**, Washington, DC, v. 108, n. 4, p. 814-834, Oct. 2001. DOI: https://doi.org/10.1037/0033-295x.108.4.814.

HAIDT, J. **The righteous mind**: Why good people are divided by politics and religion. New York: Pantheon, 2012.

HALPERN, D. **Inside the nudge unit**: How small changes can make a big difference. London: WH Allen, 2015.

HANSEN, P. G. The definition of nudge and libertarian paternalism: Does the hand fit the glove?. **European Journal of Risk Regulation**, Cambridge, v. 7, n. 1, p. 155-174, 2016. DOI: https://doi.org/10.1017/S1867299X00005468.

HANSEN, P. G.; JESPERSEN, A. M. Nudge and the manipulation of choice. **European Journal of Risk Regulation**, Cambridge, v. 4, n. 1, p. 3-28, 2013. DOI: https://doi.org/10.1017/S1867299X00002762.

HARDIN, G. The tragedy of the commons. **Science, New Series**, New York, v. 162, n. 3859, p. 1243-1248, Dec. 1968. Disponível em: https://pages.mtu.edu/~asmayer/rural_sustain/governance/Hardin%201968.pdf. Acesso em: 10 jul. 2022.

HARDISTY, D. J.; JOHNSON, E. J.; WEBER, E. U. A dirty word or a dirty world?: Attribute framing, political affiliation, and query theory. **Psychological Science**, New York, v. 21, n. 1, p. 86-92, Jan. 2010. DOI: https://doi.org/10.1177/0956797609355572.

HAUSHOFER, J.; FEHR, E. On the psychology of poverty. **Science**, New York, v. 344, n. 6186, p. 862-867, May 2014. DOI: https://doi.org/10.1126/science.1232491.

HAUSMAN, D. M. Nudging and other ways of steering choices. **Intereconomics**, Hamburg, v. 53, n. 1, p. 17-20, 2018. Disponível em: https://www.intereconomics.eu/contents/year/2018/number/1/article/nudging-and-other-ways-of-steering-choices.html. Acesso em: 10 jul. 2022.

HAUSMAN, D. M.; WELCH, B. Debate: To nudge or not to nudge. **Journal of Political Philosophy**, Cambridge, v. 18, n. 1, p. 123-136, Mar. 2010. DOI: https://doi.org/10.1111/j.1467-9760.2009.00351.x.

HAYEK, F. A. **The market and other orders**. Chicago: The University of Chicago Press, 2014. (The collected works of F. A. Hayekv, n. 15).

HAYEK, F. A. The use of knowledge in society. **American Economic Review**, Nashville, v. 35, n. 4, Sept. 1945.

HENRICH, J.; HEINE, S. J.; NORENZAYAN, A. The weirdest people in the world?. **The Behavioral and Brain Sciences**, Cambridge, v. 33, n. 2/3, p. 61-83, June 2010. DOI: https://doi.org/10.1017/S0140525X0999152X.

HERTWIG, R. When to consider boosting: Some rules for policy-makers. **Behavioural Public Policy**, Cambridge, v. 1, n. 2, p. 143-161, Nov. 2017. DOI: https://doi.org/10.1017/bpp.2016.14.

HERTWIG, R.; GRÜNE-YANOFF, T. Nudging and boosting: Steering or empowering good decisions. **Perspectives on Psychological Science**, Washington, DC, v. 12, n. 6, p. 973-986, Nov. 2017. DOI: https://doi.org/10.1177/1745691617702496.

HIGGINS, J. P. T. *et al*. (ed.). **Cochrane handbook for systematic reviews of interventions**. 2nd. ed. Hoboken: Wiley Blackwell, 2019.

HILL, C. A. Anti-Anti-Anti-Paternalism. **NYU Journal of Law & Liberty**, New York, v. 2, p. 444–454, 2007. DOI:

HOMONOFF, T. A. Can small incentives have large effects?: The impact of taxes versus bonuses on disposable bag use. **American Economic Journal: Economic Policy**, Washington, DC, v. 10, n. 4, p. 177-210, Nov. 2018. DOI: https://doi.org/10.1257/pol.20150261.

HOSSAIN, T.; MORGAN, J. ...Plus shipping and handling: Revenue (non) equivalence in field experiments on eBay. **The B.E. Journal of Economic Analysis & Policy**, [*s. l.*], v. 6, n. 2, p. 0102202153806371429, Jan. 2006. DOI: https://doi.org/10.2202/1538-0637.1429.

HUMMEL, D.; MAEDCHE, A. How effective is nudging?: A quantitative review on the effect sizes and limits of empirical nudging studies. **Journal of Behavioral and Experimental Economics**, Amsterdam, v. 80, p. 47-58, June. 2019. DOI: https://doi.org/10.1016/j.socec.2019.03.005.

IVANKOVIĆ, V.; ENGELEN, B. Nudging, transparency, and watchfulness. **Social Theory and Practice**, Tallahassee, v. 45, n. 1, p. 43-73, Jan. 2019. DOI: https://doi.org/10.5840/soctheorpract20191751.

JACHIMOWICZ, J. M. *et al*. When and why defaults influence decisions: A meta-analysis of default effects. **Behavioural Public Policy**, Cambridge, v. 3, n. 2, p. 159-186, 2019. DOI: https://doi.org/10.1017/bpp.2018.43.

JOHN, P. *et al*. **Nudge, nudge, think, think**: Experimenting with ways to change civic behaviour. London: Bloomsbury Academic, 2011.

JOHN, P. **How far to nudge?**: Assessing behavioural public policy. Cheltenham: Edward Elgar Publishing, 2018.

JOHN, P.; SMITH, G.; STOKER, G. Nudge nudge, think think: Two strategies for changing civic behaviour. **The Political Quarterly**, Oxford, v. 80, n. 3, p. 361-370, July/Sept. 2009. DOI: https://doi.org/10.1111/j.1467-923X.2009.02001.x.

JOHN, P.; STOKER, G. Rethinking the role of experts and expertise in behavioural public policy. **Policy & Politics**, Bristol, v. 47, n. 2, p. 209-226, Apr. 2019. DOI: https://doi.org/10.1332/030557319X15526371698257.

JOHNSON, E. J. *et al*. Beyond nudges: Tools of a choice architecture. **Marketing Letters**, Boston, v. 23, n. 2, p. 487-504, June 2012. DOI: https://doi.org/10.1007/s11002-012-9186-1.

JOHNSON, E. J. *et al*. Framing, probability distortions, and insurance decisions. **Journal of Risk and Uncertainty**, Boston, v. 7, n. 1, p. 35-51, Aug. 1993. DOI: https://doi.org/10.1007/BF01065313.

JOHNSON, E. J.; GOLDSTEIN, D. Do defaults save lives?. **Science**, New York, v. 302, n. 5649, p. 1338-1339, Nov. 2003. DOI: https://doi.org/10.1126/science.1091721.

JOLLS, C.; SUNSTEIN, C. R. Debiasing through law. **Journal of Legal Studies**, Chicago, v. 35, n. 1, p. 199-242, Jan. 2006. DOI: https://doi.org/10.1086/500096.

JOLLS, C.; SUNSTEIN, C. R.; THALER, R. A behavioral approach to law and economics. **Stanford Law Review**, Stanford, v. 50, n. 5, p. 1471-1550, 1998. DOI: https://doi.org/10.2307/1229304.

JONES, R.; PYKETT, J.; WHITEHEAD, M. Governing temptation: Changing behaviour in an age of libertarian paternalism. **Progress in Human Geography**, London, v. 35, n. 4, p. 483-501, 2011. DOI: https://doi.org/10.1177/0309132510385741.

JORDÃO, E. F.; BOGOSSIAN, A. M. Três ressalvas à regulação paternalista. **Revista Brasileira de Estudos Políticos**, Belo Horizonte, n. 122, p. 241-267, jan./jun. 2021. Disponível em: https://pos.direito.ufmg.br/rbep/index.php/rbep/article/view/830. Acesso em: 11 jul. 2022.

JUNG, J. Y.; MELLERS, B. A. American attitudes toward nudges. **Judgment and Decision Making**, [*s. l.*], v. 11, n. 1, p. 62-74, Jan. 2016. Disponível em: https://journal.sjdm.org/15/15824a/jdm15824a.pdf. Acesso em: 13 jul. 2022.

KAHNEMAN, D. A perspective on judgment and choice: Mapping bounded rationality. **The American Psychologist**, Washington, DC, v. 58, n. 9, p. 697-720, Sept. 2003. DOI: https://doi.org/10.1037/0003-066X.58.9.697.

KAHNEMAN, D. A psychological point of view: Violations of rational rules as a diagnostic of mental processes. **The Behavioral and Brain Sciences**, Cambridge, v. 23, n. 5, p. 681-683, Oct. 2000. DOI: https://doi.org/10.1017/S0140525X00403432.

KAHNEMAN, D. **Attention and effort**. Englewood Cliffs: Prentice-Hall, 1973.

KAHNEMAN, D. **Rápido e devagar**: duas formas de pensar. Tradução de Cássio de Arantes Leite. Rio de Janeiro: Objetiva, 2012.

KAHNEMAN, D.; FREDERICK, S. Representativeness revisited: Attribute substitution in intuitive judgment. *In*: GILOVICH, T.; GRIFFIN, D.; KAHNEMAN, D. (ed.). **The psychology of intuitive judgement**. New York: Cambridge University Press, 2002. p. 49-81.

KAHNEMAN, D.; KNETSCH, J. L.; THALER, R. H. Anomalies: The endowment effect, loss aversion, and status quo bias. **Experiments in Environmental Economics**, Washington, DC, v. 5, n. 1, p. 193-206, 1991. DOI: https://doi.org/10.1257/jep.5.1.193.

KAHNEMAN, D.; RITOV, I.; SCHKADE, D. Economic preferences or attitude expressions?: An analysis of dollar responses to public issues. **Journal of Risk and Uncertainty**, Boston, v. 19, p. 203-235, Dec. 1999. DOI: https://doi.org/10.1023/A:1007835629236.

KAHNEMAN, D.; SIBONY, O.; SUNSTEIN, C. R. **Noise**: A flaw in human judgment. New York: Little Brown Spark, 2021.

KAHNEMAN, D.; SLOVIC, P.; TVERSKY, A. (ed.). **Judgment under uncertainty**: Heuristics and biases. New York: Cambridge Univeristy Press, 1982.

KAHNEMAN, D.; TVERSKY, A. Choices, values, and frames. **The American Psychologist**, Washington, DC, v. 39, n. 4, p. 341-350, 1984. DOI: https://doi.org/10.1037/0003-066X.39.4.341.

KAHNEMAN, D.; TVERSKY, A. On the psychology of prediction. **Psychological Review**, Washington, DC, v. 80, n. 4, p. 237-251, July 1973. DOI: https://doi.org/10.1037/h0034747.

KAHNEMAN, D.; TVERSKY, A. On the reality of cognitive illusions. **Psychological Review**, Washington, DC, v. 103, n. 3, p. 582-591, July 1996. DOI: https://doi.org/10.1037/0033-295X.103.3.582.

KAHNEMAN, D.; TVERSKY, A. On the study of statistical intuitions. **Cognition**, Hague, v. 11, n. 2, p. 123-141, Mar. 1982. DOI: https://doi.org/10.1016/0010-0277(82)90022-1.

KAHNEMAN, D.; WAKKER, P. P.; SARIN, R. Back to Bentham?: Explorations of experienced utility. **The Quarterly Journal of Economics**, Oxford, v. 112, n. 2, p. 375-405, May 1997. DOI: https://doi.org/10.1162/003355397555235.

KANT, I. **Lectures on ethics**. Cambridge: Cambridge University Press, 1997.

KANT, I. **Metafísica dos costumes**. Tradução de Nathalie Bressiani *et al*. Petrópolis: Vozes, 2013.

KELMAN, M. **The heuristics debate**. Oxford: Oxford Scholarship Online, 2011. DOI: https://doi.org/10.1093/acprof:oso/9780199755608.001.0001.

KEYSAR, B.; HAYAKAWA, S. L.; AN, S. G. The foreign-language effect. **Psychological Science**, New York, v. 23, n. 6, p. 661-668, June 2012. DOI: https://doi.org/10.1177/0956797611432178.

KOROBKIN, R. Libertarian welfarism. **California Law Review**, Berkeley, v. 97, n. 6, p. 1651-1686, Dec. 2009. Disponível em: https://www.californialawreview.org/print/1libertarian-welfarism/. Acesso em: 11 jul. 2022.

KRAAKMAN, R. *et al*. **The anatomy of corporate law**: A comparative and functional approach. 3rd. ed. Oxford: Oxford University Press, 2017.

LAIBSON, D. Golden eggs and hyperbolic discounting. **The Quarterly Journal of Economics**, Oxford, v. 112, n. 2, p. 443-478, May 1997. DOI: https://doi.org/10.1162/003355397555253.

LAKENS, D. Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. **Frontiers in Psychology**, Pully, v. 4, article 863, Nov. 2013. DOI: https://doi.org/10.3389/fpsyg.2013.00863.

LAMBERTON, C. Nudging in a sludge-filled world. *In*: HUFFPOST. [*S. l.*], 30 Sept. 2016. Disponível em: https://www.huffpost.com/entry/nudging-in-a-sludgefilled_b_12087688. Acesso em: 11 jul. 2022.

LANCIERI, F. Narrowing data protection's enforcement gap. **Maine Law Review**, Portland, v. 74, n. 1, p. 1-68, Mar. 2021. DOI: http://dx.doi.org/10.2139/ssrn.3806880.

LAPLACE, P. S. M. **A philosophical essay on probabilities**. New York: Dover Publications, 1951.

LEAL, F.; RIBEIRO, L. M. Heurística de ancoragem e fixação de danos morais em juizados especiais cíveis no Rio de Janeiro: uma nova análise. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 778-800, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5295.

LEGGETT, W. The politics of behaviour change: Nudge, neoliberalism and the state. **Policy & Politics**, Bristol, v. 42, n. 1, p. 3-19, 2014. DOI: https://doi.org/10.1332/030557312X655576.

LENHARD, W.; LENHARD, A. Computation of effect sizes. *In*: PSYCHOMETRICA. [*S. l.*], 2016. Disponível em: https://www.psychometrica.de/effect_size.html. Acesso em: 11 jul. 2022.

LEWIS, M. **The undoing project**: A friendship that changed our minds. New York: W. W. Norton & Company, 2016.

LIN, Y.; OSMAN, M.; ASHCROFT, R. Nudge: Concept, effectiveness, and ethics. **Basic and Applied Social Psychology**, Hillsdale, v. 39, n. 6, p. 293-306, 2017. DOI: https://doi.org/10.1080/01973533.2017.1356304.

LINS HORTA, R. Arquitetura de escolhas, direito e liberdade: notas sobre o "paternalismo libertário". **Pensar Revista de Ciências Jurídicas**, Fortaleza, v. 22, n. 2, p. 651-664, maio/ago. 2017. DOI: https://doi.org/10.5020/2317-2150.2017.5602.

LIST, J. A. **The voltage effect**. New York: Currency, 2022.

LOEWENSTEIN, G. *et al*. Warning: you are about to be nudged. **Behavioral Science & Policy**, Durham, v. 1, n. 1, p. 35-42, 2015. DOI: http://dx.doi.org/10.2139/ssrn.2417383.

LOEWENSTEIN, G. F.; HAISLEY, E. C. The economist as therapist: Methodological ramifications of "light" paternalism. **SSRN Electronic Journal**, Rochester, Feb. 2007. DOI: http://dx.doi.org/10.2139/ssrn.962472.

LOEWENSTEIN, G.; CHATER, N. Putting nudges in perspective. **Behavioural Public Policy**, Cambridge, v. 1, n. 1, p. 26-53, 2017. DOI: http://dx.doi.org/10.1017/bpp.2016.7.

LOEWENSTEIN, G.; HAISLEY, E. The economist as therapist: Methodological ramifications of "light" paternalism. *In*: CAPLIN, A.; SCHOTTER, A. (ed.). **The foundation of positive and normative economics**: A handbook. Oxford: Oxford University Press, 2008. (The Handbooks in Economics Methodologies Series). p. 210-248.

LOEWENSTEIN, G. Hot-cold empathy gaps and medical decision making. **Health Psychology**, Washington, DC, Jul. 2005, v. 24, n. 4S, p. S49-S56, 2005. DOI: https://doi.org/10.1037/0278-6133.24.4.S49.

LUO, Y. *et al*. **A meta-analytic cognitive framework of nudge and sludge**. [*S. l.*: *s. n.*], 2021. Disponível em: https://psyarxiv.com/dbmu3/. Acesso em: 11 jul. 2022.

MACHADO SEGUNDO, H. D. B. Ciência do direito tributário, economia comportamental e extrafiscalidade. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 640-659, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5252.

MADRIAN, B. C.; SHEA, D. F. The power of suggestion: Inertia in 401(k) participation and savings behavior. **The Quarterly Journal of Economics**, Oxford, v. 116, n. 4, p. 1149-1187, Nov. 2001. DOI: https://doi.org/10.1162/003355301753265543.

MAGUIRE, E. A. *et al*. Navigation-related structural change in the hippocampi of taxi drivers. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, DC, v. 97, n. 8, p. 4398-4403, Mar. 2000. DOI: https://doi.org/10.1073/pnas.070039597.

MANI, A. *et al*. Poverty impedes cognitive function. **Science**, New York, v. 341, n. 6149, p. 976-680, Aug. 2013. DOI: https://doi.org/10.1126/science.1238041.

MANKIW, N. G. **Introdução à economia**. Tradução de Elisete Paes e Lima *et al*. 6. ed. São Paulo: Cengage Learning, 2013.

MARCUS, A. Highly criticized paper on dishonesty retracted. *In*: RETRACTION Watch. [*S. l.*], 14 Sept. 2021. Disponível em: https://retractionwatch.com/2021/09/14/highly-criticized-paper-on-dishonesty-retracted/. Acesso em: 11 jul. 2022.

MARKS, M. Biosupremacy: Data, antitrust, and monopolistic power over human behavior. **U. C. Davis Law Review**, Davis, v. 55, p. 513-589, 2021. Disponível em: https://lawreview.law.ucdavis.edu/issues/55/1/articles/files/55-1_Marks.pdf. Acesso em: 11 jul. 2022.

MAXIMILIANO, C. **Hermenêutica e aplicação do direito**. 21. ed. Rio de Janeiro: Forense, 2017.

MELLO, Thiago de. **Poesia comprometida com a minha e a tua vida**. Rio de Janeiro: Editora Civilização Brasileira, 1989.

MERTENS, S. *et al*. The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, DC, v. 119, n. 1, p. e2107346118, Jan. 2022. DOI: https://doi.org/10.1073/pnas.2107346118.

MILANOVIC, B. **A desigualdade no mundo**: uma nova abordagem para a era da globalização. Coimbra: Actual, 2017.

MILANOVIC, B. **Capitalismo sem rivais**: o futuro do sistema que domina o mundo. Tradução de Bernardo Ajzenberg. São Paulo: Todavia, 2020.

MILL, J. S. **Sobre a liberdade e a sujeição das mulheres**. Tradução de Paulo Geiger. São Paulo: Companhia das Letras, 2017.

MILLS, C. The heteronomy of choice architecture. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 495-509, Sept. 2015. DOI: https://doi.org/10.1007/s13164-015-0242-7.

MILLS, S. Nudge/sludge symmetry: On the relationship between nudge and sludge and the resulting ontological, normative and transparency implications. **Behavioural Public Policy**, Cambridge, p. 1-24, Dec. 2020. DOI: https://doi.org/10.1017/bpp.2020.61.

MITCHELL, G. Libertarian paternalism is an oxymoron. **Northwestern University Law Review**, Champaign, v. 99, n. 3, p. 1245-1277, 2005.

MLODINOW, L. **Subliminar**: como o inconsciente influencia nossas vidas. Tradução de Claudio Carina. São Paulo: Zahar, 2013.

MONGIN, P.; COZIC, M. Rethinking nudge: Not one but three concepts. **Behavioural Public Policy**, Cambridge, v. 2, n. 1, p. 107-124, 2018. DOI: https://doi.org/10.1017/bpp.2016.16.

MORTON, J. M. Toward an ecological theory of the norms of practical deliberation. **European Journal of Philosophy**, Oxford, v. 19, n. 4, p. 561-584, Dec. 2011. DOI: https://doi.org/10.1111/j.1468-0378.2010.00400.x.

MULLAINATHAN, S.; SHAFIR, E. **Escassez**: uma nova forma de pensar a falta de recursos na vida das pessoas e nas organizações. Tradução de Bruno Casotti. Rio de Janeiro: Best Business, 2016.

MÜNSCHER, R.; VETTER, M.; SCHEUERLE, T. A review and taxonomy of choice architecture techniques. **Journal of Behavioral Decision Making**, Chichester, v. 29, n. 5, p. 511-524, Dec. 2016. DOI: https://doi.org/10.1002/bdm.1897.

MURAMATSU, R.; FONSECA, P. Economia e psicologia na explicação da escolha intertemporal. **Revista de Economia Mackenzie**, São Paulo, v. 6, n. 1, p. 87-112, 2008.

Disponível em: https://biblat.unam.mx/hevila/RevistadeeconomiaMackenzie/2008/vol6/no1/4.pdf. Acesso em: 11 jul. 2022.

NISBETT, R.; ROSS, L. **Human inference**: Strategies and shortcomings of social judgment. New Jersey: Prentice-Hall, 1980.

NUNES, D.; LUD, N.; PEDRON, F. Q. **Desconfiando da imparcialidade dos sujeitos processuais**: um estudo sobre os vieses cognitivos, a mitigação de seus efeitos e o debiasing. Salvador: JusPodivm, 2018.

NUSSBAUM, M. C. **Creating capabilities**: The human development approach. Cambridge: The Belknap Press, 2011.

O'DONOGHUE, T.; RABIN, M. Choice and procrastination. **The Quarterly Journal of Economics**, Oxford, v. 116, n. 1, p. 121-160, Feb. 2001. DOI: https://doi.org/10.1162/003355301556365.

OGUS, A. **Regulation**: Legal form and economic theory. Oxford: Hart Publishing, 2004.

OLIVEIRA, L. J. D.; STANCIOLI, B. S. *Nudge* e informação: a tomada de decisão e o "homem médio". **Revista Direito GV**, São Paulo, v. 17, n. 1, p. e2114, 2021. DOI: https://doi.org/10.1590/2317-6172202114.

OLIVEIRA, M. **Racionalidade quebrada**: a racionalidade ecológica das heurísticas rápidas e frugais. 2005. 504 f. Tese (Doutoramento em Psicologia) – Universidade de Coimbra, Coimbra, 2005.

OLIVER, A. Nudging, shoving, and budging: Behavioural economic-informed policy. **Public Administration**, [London], v. 93, n. 3, p. 700-714, 2015. DOI: https://doi.org/10.1111/padm.12165.

OPPENHEIMER, D. M. Not so fast! (and not so frugal!): Rethinking the recognition heuristic. **Cognition**, Hague, v. 90, n. 1, p. B1-B9, Nov. 2003. DOI: https://doi.org/10.1016/S0010-0277(03)00141-0.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD). **Behavioural insights and public policy**: Lessons from around the World. Paris: OECD Publishing, 2017. Disponível em: https://www.oecd.org/gov/regulatory-policy/behavioural-insights-and-public-policy-9789264270480-en.htm. Acesso em: 11 jul. 2022.

ORWELL, G. **1984**. Tradução de Heloisa Jahn e Alexandre Hubner. São Paulo: Companhia das Letras, 2019.

OSMAN, M. **Backfiring, reactance, boomerang, spillovers, and rebound effects**: Can we learn anything from examples where nudges do the opposite of what they intended?. [*S. l.*: *s. n.*], 2020. DOI: https://doi.org/10.31234/osf.io/ae756.

OSTROM, E. **Governing the commons**: The evolution of institutions for collective action. Cambridge: Cambridge University Press, 2015.

PASTERNAK, B. **Doutor Jivago**. Tradução de Sonia Branco. São Paulo: Companhia das Letras, 2017.

PEARL, J.; MACKENZIE, D. **The book of why**: The new science of cause and effect. New York: Basic Books, 2018.

PEETERS, R. Manufacturing responsibility: The governmentality of behavioural power in social policies. **Social Policy and Society**, Cambridge, v. 18, n. 1, p. 51-65, 2019. DOI: https://doi.org/10.1017/S147474641700046X.

PELTZMAN, S. The effects of automobile safety regulation. **Journal of Political Economy**, Chicago, v. 83, n. 4, p. 677-725, Aug. 1975. DOI: https://doi.org/10.1086/260352.

PEREIRA RIBEIRO, M. C.; DOMINGUES, V. H. Paternalismo libertário e políticas públicas: intervenção e transparência. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 11, n. 1, p. 105-120, abr. 2021. DOI: https://doi.org/10.5102/rbpp.v11i1.6561.

PETRESCU, D. C. *et al.* Public acceptability in the UK and USA of nudging to reduce obesity: The example of reducing sugar-sweetened beverages consumption. **PLoS One**, San Francisco, v. 11, n. 6, p. e0155995, June 2016. DOI: https://doi.org/10.1371/journal.pone.0155995.

PETTICREW, M. *et al*. Dark nudges and sludge in big alcohol: Behavioral economics, cognitive biases, and alcohol industry corporate social responsibility. **The Milbank Quarterly**, New York, v. 98, n. 4, p. 1290-1328, Dec. 2020. DOI: https://doi.org/10.1111/1468-0009.12475.

PIKETTY, T. **O capital no século XXI**. Tradução de Monica Baumgarten De Bolle. Rio de Janeiro: Intrínseca, 2014.

POSNER, R. A. Rational choice, behavioral economics, and the law. **Stanford Law Review**, Stanford, v. 50, n. 5, p. 1551, 1998.

POSNER, R. **Economic analysis of law**. 3rd. ed. New York: Aspen Publishers, 1986.

RABIN, M.; SCHRAG, J. L. First impressions matter: A model of confirmatory bias. **The Quarterly Journal of Economics**, Oxford, v. 114, n. 1, p. 37-82, Feb. 1999. DOI: https://doi.org/10.1162/003355399555945.

RAGHUBIR, P.; SRIVASTAVA, J. Monopoly money: The effect of payment coupling and form on spending behavior. **Journal of Experimental Psychology: Applied**, Washington, DC, v. 14, n. 3, p. 213-225, 2008. DOI: https://doi.org/10.1037/1076-898X.14.3.213.

RAWLS, J. **A theory of justice**. Cambridge: The Belknap Press, 1999.

RAZ, J. **The morality of freedom**. Oxford: Oxford University Press, 1986.

REBONATO, R. A critical assessment of libertarian paternalism. **Journal of Consumer Policy**, Dordrecht, v. 37, n. 3, p. 357-396, Sept. 2014. DOI: https://doi.org/10.1007/s10603-014-9265-1.

REISCH, L. A.; SUNSTEIN, C. R. Do Europeans like nudges?. **Judgment and Decision Making**, [*s. l.*], v. 11, n. 4, p. 310-325, July 2016. DOI: https://doi.org/10.2139/ssrn.2739118.

REISCH, L. A.; SUNSTEIN, C. R.; GWOZDZ, W. Beyond carrots and sticks: Europeans support health nudges. **Food Policy**, Guildford, v. 69, p. 1-10, May 2017. DOI: https://doi.org/10.1016/j.foodpol.2017.01.007.

REYNOLDS, J. P. *et al.* Public acceptability of nudging and taxing to reduce consumption of alcohol, tobacco, and food: A population-based survey experiment. **Social Science & Medicine**, Oxford, v. 236, p. 112395, Sept. 2019. DOI: https://doi.org/10.1016/j.socscimed.2019.112395.

RICK, S.; LOEWENSTEIN, G. Intangibilidade na escolha intertemporal. *In*: ÁVILA, F.; BIANCHI, A. M. (ed.). **Guia de economia comportamental e experimental**. São Paulo: EconomiaComportamental.org, 2015. p. 76-97.

RITOV, I.; BARON, J. Reluctance to vaccinate: Omission bias and ambiguity. **Journal of Behavioral Decision Making**, Chichester, v. 3, n. 4, p. 263-277, Oct./Dec.1990. DOI: https://doi.org/10.1002/bdm.3960030404.

RIZZO, M. J.; WHITMAN, D. G. Little brother is watching you: New paternalism on the slippery slopes. **Arizona Law Review**, Tucson, v. 51, n. 685, p. 685-739, 2009. Disponível em: https://arizonalawreview.org/pdf/51-3/51arizlrev685.pdf. Acesso em: 11 jul. 2022.

RIZZO, M. J.; WHITMAN, D. G. The knowledge problem of new paternalism. **SSRN Electronic Journal**, Rochester, v. 2009, n. 4, p. 905-968, 2008.

SAGHAI, Y. Salvaging the concept of nudge. **Journal of Medical Ethics**, London, v. 39, n. 8, p. 487-493, Aug. 2013. DOI: https://doi.org/10.1136/medethics-2012-100727.

SAINT-PAUL, G. **The tyranny of utility**: Behavioral social science and the rise of paternalism. Princeton: Princeton University Press, 2011.

SAMUELSON, P. A. A note on measurement of utility. **The Review of Economic Studies**, Oxford, v. 4, n. 2, p. 155-161, Feb. 1937. DOI: https://doi.org/10.2307/2967612.

SAMUELSON, P. A. Consumption theory in terms of revealed preference. **Economica**, [Jakarta], v. 15, n. 60, p. 243-253, Nov. 1948. DOI: https://doi.org/10.2307/2549561.

SAMUELSON, W.; ZECKHAUSER, R. Status quo bias in decision making. **Journal of Risk and Uncertainty**, Boston, v. 1, n. 1, p. 7-59, Mar. 1988. DOI: https://doi.org/10.1007/BF00055564.

SARAIVA, R. Entre o paternalismo e o estado libertário: a ética e constitucionalidade de uma nova regulação comportamental. *In*: GOMES, C. A. *et al.* (ed.). **Garantia de direitos e regulação**: perspectivas de direito administrativo. Lisboa: AAFDL Editora, 2020. p. 19-48.

SCHMIDT, A. T. Getting real on rationality: Behavioral science, nudging, and public policy. **Ethics**, Chicago, v. 129, n. 4, p. 511-543, July 2019. DOI: https://doi.org/10.1086/702970.

SCHMIDT, A. T. The power to nudge. **American Political Science Review**, Cambridge, v. 111, n. 2, p. 404-417, 2017. DOI: https://doi.org/10.1017/S0003055417000028.

SCHMIDT, A. T.; ENGELEN, B. The ethics of nudging: An overview. **Philosophy Compass**, [London], v. 15, n. 4, p. e12658, Apr. 2020. DOI: https://doi.org/10.1111/phc3.12658.

SCHNELLENBACH, J.; SCHUBERT, C. **Behavioral public choice**: A survey. Freiburg: Walter Eucken Institut, 2014. (Freiburg Discussion Papers on Constitutional Economics, n. 14/03). Disponível em: http://hdl.handle.net/10419/92975. Acesso em: 11 jul. 2022.

SCHUBERT, C. Exploring the (behavioural) political economy of nudging. **Journal of Institutional Economics**, Cambridge, v. 13, n. 3, p. 499-522, 2017. DOI: https://doi.org/10.1017/S1744137416000448.

SCHULTZ, P. W. *et al*. The constructive, destructive, and reconstructive power of social norms: Reprise. **Perspectives on Psychological Science**, Washington, DC, v. 13, n. 2, p. 249-254, Mar. 2018. DOI: https://doi.org/10.1177/1745691617693325.

SCHWEIZER, M. Nudging and the principle of proportionality. *In*: MATHIS, K.; TOR, A. (ed.). **Nudging**: Possibilities, limitations and applications in European Law and Economics. [London]: Springer, 2016. p. 93-120.

SELINGER, E.; WHYTE, K. Is there a right way to nudge?: The practice and ethics of choice architecture. **Sociology Compass**, [Oxford], v. 5, n. 10, p. 923-935, Oct. 2011. DOI: https://doi.org/10.1111/j.1751-9020.2011.00413.x.

SELINGER, E.; WHYTE, K. P. Nudging cannot solve complex policy problems. **European Journal of Risk Regulation**, Cambridge, v. 3, n. 1, p. 26-31, 2012. DOI: https://doi.org/10.1017/S1867299X0000177X.

SEN, A. Behaviour and the concept of preference. **Economica**, [Jakarta], v. 40, n. 159, p. 241-259, Aug. 1973. DOI: https://doi.org/10.2307/2552796.

SHAH, A. K.; MULLAINATHAN, S.; SHAFIR, E. Some consequences of having too little. **Science**, New York, v. 338, n. 6107, p. 682-685, Nov. 2012. DOI: https://doi.org/10.1126/science.1222426.

SHAROT, T. **The optimism bias**: A tour of the irrationally positive brain. New York: Pantheon Books, 2011.

SHU, L. L. *et al.* Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end. **Proceedings of the National Academy of Sciences of the United States of America**, Washington, DC, v. 109, n. 38, p. 15197-15200, Sept. 2012. DOI: https://doi.org/10.1073/pnas.1209746109.

SILVA, A. C. S.; RODRIGUES, D. R. N.; TIBALDI, S. D. *Nudges* e políticas públicas: um mecanismo de combate ao trabalho em condição análoga à de escravo. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 267-287, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5317.

SILVA, L. N. *et al.* Entre o porrete, a cenoura e os empurrões: como medidas regulatórias híbridas podem aprimorar a renegociação das faturas de cartão de crédito. **Economic Analysis of Law Review**, Brasília, DF, v. 10, n. 2, p. 70-96, maio/ago. 2019. Disponível em: https://portalrevistas.ucb.br/index.php/EALR/article/view/9675. Acesso em: 11 jul. 2022.

SILVA, L. N. *et al*. Reduzindo a tributação cognitiva: lições comportamentais para a diminuição dos efeitos psicológicos adversos da pobreza. **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 288-326, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5331.

SILVA, L. N. **Regulação e concorrência no mercado doméstico de aviação no Brasil**: a análise econômica do direito e a comparação com o processo de desregulação nos Estados Unidos. São Paulo: Editora Singular, 2013.

SILVA, L. N. Regulação, concorrência e o setor bancário: reflexões. **Revista de Informação Legislativa**, Brasília, DF, v. 40, n. 158, p. 171-193, abr./jun. 2003. Disponível em: https://www2.senado.leg.br/bdsf/item/id/852. Acesso em: 11 jul. 2022.

SILVA, V. A. **Direito constitucional brasileiro**. São Paulo: Editora da Universidade de São Paulo, 2021.

SIMON, H. A. A behavioral model of rational choice. **The Quarterly Journal of Economics**, Oxford, v. 69, n. 1, p. 99-118, Feb. 1955. DOI: https://doi.org/10.2307/1884852.

SIMON, H. A. Rational choice and the structure of the environment. **Psychological Review**, Washington, DC, v. 63, n. 2, p. 129-138, 1956. DOI: https://doi.org/10.1037/h0042769.

SIMONS, D. J.; CHABRIS, C. F. Gorillas in our midst: Sustained inattentional blindness for dynamic events. **Perception**, London, v. 28, n. 9, p. 1059-1074, Sept. 1999. DOI: https://doi.org/10.1068/p281059.

SINNOTT-ARMSTRONG, W. Consequentialism. *In*: ZALTA, E. N. (ed.). **The Stanford encyclopedia of philosophy**. Stanford: Metaphysics Research Lab, 2019. Disponível em: https://plato.stanford.edu/entries/consequentialism/. Acesso em: 11 jul. 2022.

SLOVIC, P. The construction of preference. **The American Psychologist**, Washington, DC, v. 50, n. 5, p. 364-371, 1995. DOI: https://doi.org/10.1037/0003-066X.50.5.364.

SMITH, A. **A riqueza das nações**: investigação sobre sua natureza e suas causas. Tradução de Luiz João Baraúna. São Paulo: Editora Nova Cultural, 1996.

SMITH, A. **Teoria dos sentimentos morais ou ensaio para uma análise dos princípios pelos quais os homens naturalmente julgam a conduta e o caráter, primeiro de seus próximos, depois de si mesmos**. Tradução de Lya Luft. São Paulo: Martins Fontes, 2015.

SOMAN, D. **The last mile**: Creating social and economic value from behavioral insights. Toronto: University of Toronto Press, 2015.

SOMAN, D.; HOSSAIN, T. Successfully scaled solutions need not be homogenous. **Behavioural Public Policy**, Cambridge, v. 5, n. 1, p. 80-89, Jan. 2021. DOI: https://doi.org/10.1017/bpp.2020.24.

STANCIOLI, B. *et al.* O sistema nacional de transplantes: saúde e autonomia em discussão. **Revista de Direito Sanitário**, São Paulo, v. 11, n. 3, p. 123-154, fev. 2011. DOI: https://doi.org/10.11606/issn.2316-9044.v11i3p123-154.

STANCIOLI, B. **Renúncia ao exercício de direitos da personalidade ou como alguém se torna o que quiser**. Belo Horizonte: Editora D'Plácido, 2017.

STANCIOLI, B. S. **Relação jurídica médico-paciente**. Belo Horizonte: Del Rey, 2004.

STANOVICH, K. E.; WEST, R. F. Individual differences in reasoning: Implications for the rationality debate?. **The Behavioral and Brain Sciences**, Cambridge, v. 23, n. 5, p. 645-665, 2000. DOI: https://doi.org/10.1017/S0140525X00003435.

STEEL, P. The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. **Psychological Bulletin**, Washington, DC, v. 133, n. 1, p. 65-94, Jan. 2007. DOI: https://doi.org/10.1037/0033-2909.133.1.65.

STEFFEL, M.; WILLIAMS, E. F.; POGACAR, R. Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. **Journal of Marketing Research**, [Petersburg], v. 53, n. 5, p. 865-880, Oct. 2016. DOI: https://doi.org/10.1509/jmr.14.0421.

SUNSTEIN, C. R. **Libertarian paternalism is not an oxymoron (nudge before nudge, kind of)**. [*S. l.*], 29 Mar. 2022. Twitter: @CassSunstein. Disponível em: https://twitter.com/CassSunstein/status/1508762754659897344?s=20&t=TPzqFXZSKn2V-CLK05qeKQ. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. "Better off, as judged by themselves": A comment on evaluating nudges. **International Review of Economics**, [Berlin], v. 65, n. 1, p. 1-8, Mar. 2018c. DOI: https://doi.org/10.1007/s12232-017-0280-9.

SUNSTEIN, C. R. Behavioral analysis of law. **University of Chicago Law Review**, Chicago, v. 64, n. 4, p. 1175-1195, 1997. Disponível em: https://chicagounbound.uchicago.edu/uclrev/vol64/iss4/4/. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. Behavioral welfare economics. **Journal of Benefit-Cost Analysis**, Cambridge, v. 11, n. 2, p. 196-220, June 2020c. DOI: https://doi.org/10.1017/bca.2020.14.

SUNSTEIN, C. R. **Choosing not to choose**: Understanding the value of choice. New York: Oxford University Press, 2015a.

SUNSTEIN, C. R. Do people like nudges?. **Administrative Law Review**, [Saratoga Springs], v. 68, n. 2, p. 177-232, 2016b.

SUNSTEIN, C. R. Hayekian behavioral economics. **Behavioural Public Policy**, Cambridge, p. 1-19, Mar. 2021b. DOI: https://doi.org/10.1017/bpp.2021.3.

SUNSTEIN, C. R. **How change happens**. Cambridge: MIT Press, 2019a.

SUNSTEIN, C. R. Human behavior and the law of work. **Virginia Law Review**, [Charlottesville], v. 87, n. 2, p. 205-276, Apr. 2001. DOI: https://doi.org/10.2307/1073888.

SUNSTEIN, C. R. Incompletely theorized agreements. **Harvard Law Review**, Cambridge, v. 108, n. 7, p. 1733-1772, May 1995. DOI: https://doi.org/10.2307/1341816.

SUNSTEIN, C. R. Is deontology a heuristic?: On psychology, neuroscience, ethics, and law. . **SSRN**, Rochester, 1 Aug. 2013b. DOI: http://dx.doi.org/10.2139/ssrn.2304760.

SUNSTEIN, C. R. **Legal reasoning and political conflict**. 2. ed. New York: Oxford University Press, 2018a.

SUNSTEIN, C. R. Misconceptions about nudges. **Journal of Behavioral Economics for Policy**, Honolulu, v. 2, n. 1, p. 61-67, Mar. 2018b. Disponível em: https://sabeconomics.org/wordpress/wp-content/uploads/JBEP-2-1-9.pdf. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. *Nudge*: um guia bem breve. *In*: ÁVILA, F.; BIANCHI, A. M. (ed.). **Guia de economia comportamental e experimental**. São Paulo: EconomiaComportamental.org, 2015d. p. 110-115.

SUNSTEIN, C. R. Nudges do not undermine human agency. **Journal of Consumer Policy**, Dordrecht, v. 38, n. 3, p. 207-210, Sept. 2015e. DOI: https://doi.org/10.1007/s10603-015-9289-1.

SUNSTEIN, C. R. Nudges vs. shoves. **Harvard Law Review**, Cambridge, v. 127, n. 210, 2014b. Disponível em: https://dash.harvard.edu/handle/1/13031644. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. Nudging: A very short guide. **Business Economics**, [Cleveland], v. 54, n. 2, p. 127-129, Apr. 2019c. DOI: https://doi.org/10.1057/s11369-018-00104-5.

SUNSTEIN, C. R. Nudging: A very short guide. **Journal of Consumer Policy**, Dordrecht, v. 37, n. 4, p. 583-588, Dec. 2014d. DOI: https://doi.org/10.1007/s10603-014-9273-1.

SUNSTEIN, C. R. *Nudging*: um guia bem breve. *In*: ÁVILA, F.; BIANCHI, A. M. (ed.). **Guia de economia comportamental e experimental**. São Paulo: EconomiaComportamental.org, 2015b. p. 109-114.

SUNSTEIN, C. R. **Simpler**: The future of government. New York: Simon & Schuster Paperbacks, 2014c.

SUNSTEIN, C. R. Sludge and ordeals. **Duke Law Journal**, Durham, v. 68, n. 8, p. 1844-1883, 2019b. Disponível em: https://dlj.law.duke.edu/article/sludge-and-ordeals-sunstein-vol68-iss8/. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. Sludge audits. **Behavioural Public Policy**, Cambridge, p. 1-20, 2020a. DOI: https://doi.org/10.1017/bpp.2019.32.

SUNSTEIN, C. R. **Sludge**: What stops us from getting things done and what to do about it. Cambridge: MIT Press, 2021a.

SUNSTEIN, C. R. Switching the default rule. **New York University Law Review**, New York, v. 77, n. 1, p. 106-134, Apr. 2002. Disponível em: https://www.nyulawreview.org/issues/volume-77-number-1/switching-the-default-rule/. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. **The ethics of influence**: Government in the age of behavioral science. New York: Cambridge Univeristy Press, 2016a.

SUNSTEIN, C. R. The ethics of nudging. **Yale Journal on Regulation**, New Haven, v. 32, n. 2, p. 413-450, 2015c. Disponível em: https://www.yalejreg.com/print/the-ethics-of-nudging/. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. The Storrs lectures: Behavioral economics and paternalism. **The Yale Law Journal**, New Haven, v. 122, n. 7, p. 1826-1899, 2013a. Disponível em: https://core.ac.uk/download/pdf/157779433.pdf. Acesso em: 12 jul. 2022.

SUNSTEIN, C. R. **Too much information**: Understanding what you don't want to know. Cambridge: MIT Press, 2020b.

SUNSTEIN, C. R. **Why nudge?**: The politics of libertarian paternalism. New Haven: Yale University Press, 2014a.

SUNSTEIN, C. R.; REISCH, L. A. **Trusting nudges**: Toward a bill of rights for nudging. New York: Routledge, 2019.

TANNENBAUM, D.; FOX, C. R.; ROGERS, T. On the misplaced politics of behavioural policy interventions. **Nature Human Behaviour**, London, v. 1, n. 7, p. 1-7, July 2017. DOI: https://doi.org/10.1038/s41562-017-0130.

TAYLOR, C. **The ethics of authenticity**. Cambridge: Harvard University Press, 1991.

TAYLOR, S. E. The availability bias in social perception and interaction. *In*: KAHNEMAN, D.; SLOVIC, P.; TVERSKY, A. (ed.). **Judgment under uncertainty**: Heuristics and biases. New York: Cambridge Univeristy Press, 1982. p. 190-200.

TEIXEIRA, R. N. Quem é Gerd Gigerenzer?: entendendo as críticas do "proeminente psicólogo alemão". *In*: ECONOMIA Comportamental. [*S. l.*], 25 set. 2017. Disponível em: http://www.economiacomportamental.org/nacionais/quem-e-gerd-gigerenzer-entendendo-as-criticas-do-proeminente-psicologo-alemao/. Acesso em: 12 jul. 2022.

TESSIN, J. **Representation and government performance**. Dissertation presented to the Faculty of Princeton University in candidacy for the degree of Doctor of Philosophy. Princeton, 2009.

THALER, R. H. From cashews to nudges: The evolution of behavioral economics. **American Economic Review**, Nashville, v. 108, n. 6, p. 1265-1287, June 2018b. DOI: https://doi.org/10.1257/aer.108.6.1265.

THALER, R. H. **Misbehaving**: The making of behavioral economics. New York: W. W. Norton & Company, 2015.

THALER, R. H. Much ado about nudging. *In*: BEHAVIOURAL Public Policy Blog. [Cambridge], 2 June 2017. Disponível em: https://bppblog.com/2017/06/02/much-ado-about-nudging/. Acesso em: 12 jul. 2022.

THALER, R. H. Nudge, not sludge. **Science**, New York, v. 361, n. 6401, p. 431, Aug. 2018a. DOI: https://doi.org/10.1126/science.aau9241.

THALER, R. H.; BENARTZI, S. Save more tomorrow: Using behavioral economics to increase employee saving. **Journal of Political Economy**, Chicago, v. 112, n. S1, p. S164-S187, Feb. 2004. DOI: https://doi.org/10.1086/380085.

THALER, R. H.; SUNSTEIN, C. R. Libertarian paternalism is not an oxymoron. **University of Chicago Law Review**, Chicago, n. 43, p. 1159-1202, 2003b.

THALER, R. H.; SUNSTEIN, C. R. Libertarian paternalism. **American Economic Review**, Nashville, v. 93, n. 2, p. 175-179, May 2003a. DOI: https://doi.org/10.1257/000282803321947001.

THALER, R. H.; SUNSTEIN, C. R. **Nudge**: Improving decisions about health, wealth, and happiness. 2nd. ed. New York: Penguin Books, 2009.

THALER, R. H.; SUNSTEIN, C. R. **Nudge**: The final edition. New York: Penguin Books, 2021.

TIROLE, J. **Economia do bem comum**. Tradução de André Telles. Rio de Janeiro: Zahar, 2020.

TIROLE, J. **Economics for the common good**. Princeton: Princeton University Press, 2016.

TODD, P. M.; GIGERENZER, G. What is ecological rationality?. *In*: TODD, P. M.; GIGERENZER, G. **Ecological rationality**: Intelligence in the World. Oxford: Oxford University Press, 2012. p. 3-33.

TOR, A. Nudges that should fail?. **Behavioural Public Policy**, Cambridge, v. 4, n. 3, p. 316-342, Nov. 2020. DOI: https://doi.org/10.1017/bpp.2019.5.

TOR, A. The critical and problematic role of bounded rationality in nudging. *In*: MATHIS, K.; TOR, A. (ed.). **Nudging**: Possibilities, limitations and applications in European Law and Economics. [London]: Springer, 2016. p. 3-10.

TOR, A. Understanding behavioral antitrust. **Texas Law Review**, Austin, v. 92, n. 3, p. 573-667, 2014. Disponível em: http://texaslawreview.org/wp-content/uploads/2015/08/Tor.pdf. Acesso em: 13 jul. 2022.

TVERSKY, A.; KAHNEMAN, D. Evidential impact of base rates. *In*: KAHNEMAN, D.; SLOVIC, P.; TVERSKY, A. (ed.). **Judgment under uncertainty**: Heuristics and biases. New York: Cambridge Univeristy Press, 1982b. p. 153-160.

TVERSKY, A.; KAHNEMAN, D. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. **Psychological Review**, Washington, DC, v. 90, n. 4, p. 293-315, 1983. DOI: https://doi.org/10.1037/0033-295X.90.4.293.

TVERSKY, A.; KAHNEMAN, D. Judgment under uncertainty: Heuristics and biases. **Science**, New York, v. 185, n. 4157, p. 1124-1131, Sept. 1974. DOI: https://doi.org/10.1126/science.185.4157.1124.

TVERSKY, A.; KAHNEMAN, D. Judgments of and by representativeness. *In*: KAHNEMAN, D.; SLOVIC, P.; TVERSKY, A. (ed.). **Judgment under uncertainty**: Heuristics and biases. New York: Cambridge Univeristy Press, 1982a. p. 84-98.

VAN AAKEN, A. Constitutional limits to paternalistic nudging: A proportionality assessment. *In*: KEMMERER, A. *et al.* (ed.). **Choice architecture in demcoracies**: Exploring the legitimacy of nudging: Recht im kontext. Oxford: Hart Publishing, 2016. p. 161-195.

VAN AAKEN, A. Judge the nudge: In search of the legal limits of paternalistic nudging in the EU. *In*: ALEMANNO, A.; SIBONY, A. L. (ed.). **Nudge and the law**: A European perspective. Oxford: Hart Publishing, 2015. (Modern studies in European Law). p. 83-112.

VAN ROEKEL, H.; REINHARD, J.; GRIMMELIKHUIJSEN, S. Improving hand hygiene in hospitals: Comparing the effect of a nudge and a boost on protocol compliance. **Behavioural Public Policy**, Cambridge, v. 6, n. 1, p. 52-74, 2022. DOI: https://doi.org/10.1017/bpp.2021.15.

VICTOR COELHO, M. C.; AYALA, P. D. A. Paternalismo libertário e proteção jurídica do ambiente: porque proteger o ambiente deve ser, antes de tudo, não eliminar as liberdades? **Revista Brasileira de Políticas Públicas**, Brasília, DF, v. 8, n. 2, p. 406-428, ago. 2018. DOI: https://doi.org/10.5102/rbpp.v8i2.5335.

VRANAS, P. B. M. Gigerenzer's normative critique of Kahneman and Tversky. **Cognition**, Hague, v. 76, n. 3, p. 179-193, Sept. 2000. DOI: https://doi.org/10.1016/s0010-0277(99)00084-0.

VUGTS, A. *et al.* How autonomy is understood in discussions on the ethics of nudging. **Behavioural Public Policy**, Cambridge, v. 4, n. 1, p. 108-123, Mar. 2020. DOI: https://doi.org/10.1017/bpp.2018.5.

WACHNER, J.; ADRIAANSE, M. A.; DE RIDDER, D. T. D. And how would that make you feel?: How people expect nudges to influence their sense of autonomy. **Frontiers in Psychology**, Pully, v. 11, p. 607894, Dec. 2020. DOI: https://doi.org/10.3389/fpsyg.2020.607894.

WALDRON, J. It's all for your own good. *In*: THE NEW York Times Review of Books. New York, 9 Oct. 2014. Disponível em: https://www.nybooks.com/articles/2014/10/09/cass-sunstein-its-all-your-own-good/?insrc=hpma. Acesso em: 12 jul. 2022.

WALLACE, D. F. Isto é água. *In*: WALLACE, D. F. **Ficando longe do fato de já estar meio que longe de tudo**. Tradução de Daniel Galera e Daniel Pellizzari. São Paulo: Companhia das Letras, 2012. p. 263-275.

WEINSTEIN, N. D. Unrealistic optimism about future life events. **Journal of Personality and Social Psychology**, Washington, DC, v. 39, n. 5, p. 806-820, 1980. DOI: https://doi.org/10.1037/0022-3514.39.5.806.

WHITE, M. D. Behavioral law and economics: The assault on consent, will, and dignity. *In*: FAVOR, C.; GAUS, G.; LAMONT, J. (ed.). **Essays on philosophy, politics & economics**: Integration & common research projects. Staford: Stanford University Press, 2010. p. 202-223.

WHITE, M. D. Overview of behavioral economics and policy. *In*: ABDUKADIROV, S. (ed.). **Nudge theory in action**: Behavioral design in policy and markets. London: Palgrave Macmillan, 2016. p. 15-37.

WHITMAN, D. G.; RIZZO, M. J. The problematic welfare standards of behavioral paternalism. **Review of Philosophy and Psychology**, London, v. 6, n. 3, p. 409-425, Sept. 2015. DOI: https://doi.org/10.1007/s13164-015-0244-5.

WILKINSON, T. M. Nudging and manipulation. **Political Studies**, Guildford, v. 61, n. 2, p. 341-355, 2012. DOI: https://doi.org/10.1111/j.1467-9248.2012.00974.x.

WILLIS, L. E. When nudges fail: Slippery defaults. **University of Chicago Law Review**, Chicago, v. 80, n. 3, p. 1155-1229, 2013. Disponível em: https://chicagounbound.uchicago.edu/cgi/viewcontent.cgi?article=5773&context=uclrev. Acesso em: 12 jul. 2022.

WOLKART, E. N. **Análise econômica do processo civil**: como a economia, o direito e a psicologia podem vencer a tragédia da justiça. São Paulo: Revista dos Tribunais, 2019.

WRIGHT, J. D.; GINSBURG, D. H. Behavioral law and economics: Its origins, fatal flaws, and implications for liberty. **Northwestern Univeristy Law Review**, Evanston, v. 106, n. 3, p. 12-63, 2015. Disponível em: https://scholarlycommons.law.northwestern.edu/nulr/vol106/iss3/2/. Acesso em: 12 jul. 2022.

ZAJONC, R. B. Feeling and thinking: Preferences need no inferences. **The American Psychologist**, Washington, DC, v. 35, n. 2, p. 151-175, 1980. DOI: https://doi.org/10.1037/0003-066X.35.2.151.

ZAMIÁTIN, I. **Nós**. Tradução de Gabriela Soares. São Paulo: Aleph, 2017.

ZAMIR, E.; TEICHMAN, D. (ed.). **The Oxford handbook of behavioral economics and the law**. New York: Oxford University Press, 2014.

ZELLERMAYER, O. **The pain of paying**. 1996. Dissertation (PhD) – Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, 1996.