

**UNIVERSIDADE FEDERAL DE PELOTAS**  
**Instituto de Filosofia, Sociologia e Política**  
**Programa de Pós-Graduação em Filosofia**



Tese

**A IRRACIONALIDADE DA RAZÃO NA  
TOMADA DE DECISÃO MORAL**

**Luca Nogueira Igansi**

Pelotas, 2023

**LUCA NOGUEIRA IGANSI**

**A IRRACIONALIDADE DA RAZÃO NA  
TOMADA DE DECISÃO MORAL**

Tese apresentada ao Programa de Pós-Graduação em Filosofia do Instituto de Filosofia, Sociologia e Política da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Doutor em Filosofia.

Orientadora: Flávia Carvalho Chagas

Pelotas, 2023

Universidade Federal de Pelotas / Sistema de Bibliotecas  
Catalogação na Publicação

I24i Igansi, Luca Nogueira

A irracionalidade da razão na tomada de decisão moral /  
Luca Nogueira Igansi ; Flávia de Carvalho Chagas,  
orientadora. — Pelotas, 2023.

201 f.

Tese (Doutorado) — Programa de Pós-Graduação em  
Filosofia, Instituto de Filosofia, Sociologia e Política,  
Universidade Federal de Pelotas, 2023.

1. Razão. 2. Racionalidade. 3. Juízo moral. 4. Tomada de  
decisão. 5. Filosofia experimental. I. Chagas, Flávia de  
Carvalho, orient. II. Título.

CDD : 173

Elaborada por Leda Cristina Peres Lopes CRB: 10/2064

LUCA NOGUEIRA IGANSI

A irracionalidade da razão na tomada de decisão moral

Tese aprovada, como requisito parcial, para obtenção de grau de Doutor em Filosofia no Programa de Pós-Graduação em Filosofia do Instituto de Filosofia, Sociologia e Política da Universidade Federal de Pelotas.

Data da Defesa: 06 de Março de 2023

Banca examinadora:

---

Profa. Dra. Flávia Carvalho Chagas (Orientadora)  
Doutora em Filosofia pela Universidade Federal do Rio Grande do Sul

---

Prof. Dr. Carlos Alberto Miraglia  
Doutor em Filosofia pela Universidade Federal de Pelotas

---

Prof. Dr. João Hobuss  
Doutor em Filosofia pela Universidade Federal do Rio Grande do Sul

---

Prof. Dra. Heloísa Allgayer  
Doutora em Biologia pela Universidade do Vale do Rio dos Sinos

---

Prof. Dr. Lucas Dalsotto  
Doutor em Filosofia pela Universidade Federal de Santa Maria



*Dedico este trabalho à minha avó Walda Nogueira. Como muitas mulheres, foi o cerne desta família a muito custo, e sem ela não teria chegado até aqui.*

## AGRADECIMENTOS

A pesquisa acadêmica em ciências humanas no Brasil, e especialmente em filosofia, é possivelmente um dos maiores e mais ingratos esforços que podemos executar. Não obstante, no período de meu doutoramento passei não apenas por uma das maiores pandemias globais da história, como também de um (des)governo que ameaçou cotidianamente a própria estrutura da educação e inclusive da democracia. Assim, os agradecimentos que seguem não devem ser tomados levianamente.

Agradeço ao PPGFil da UFPel pelo acolhimento, em especial à Mirela Moraes, João Hobuss, Carlos Miraglia e à orientadora querida que abraçou a causa e superou todas as expectativas – estabelecendo um novo patamar no que vou esperar de orientadores e de como pretendo orientar futuramente –, Flávia Chagas. Sem dúvidas também sou muito grato aos outros membros da banca e sua participação valiosíssima para o melhoramento deste trabalho em situação tão especial, o caro professor Lucas Dalsotto e a amiga Heloísa Allgayer.

Sou muito grato aos amigos Werner Michel, Lisi Gastal, Lúcia Badia, Roberto Pirs, Pedro Giusti, Leonardo Camacho, Daniel Añaña, Ana Konradt, Raphael Balboni, Augusto Schmidt, Rafaela Casarin, Esmute Farias, Lauro Quadrado, Paula Luersen, Eduardo Luersen, Bya de Paula, Mathaus Krüger, Thaís Barbieri, José Bordin, Paulo Momento, Marcinha e Robertinho Neves, Lucas Misu, Lucas Matos, Guilherme Magalhães, Raíza Alves, Lara Azevedo, Matias Camargo, Karina Peres, Daniel Blotta, Aroldo dos Anjos, Juliano de Lima, Bruna Oliveira, Lucas Matos, Sandrine Knopp, Rafael Cristino, Lu Nunes e Nati Carivalis. Aos parceiros de jogo Pedro Monsalve, Bruno Souza, Lucas Giusti, Gabriela Silveira, Rodrigo Neves, Paulo Tessmer, Matheus de Oliveira. Aos Kroktopus Dielon Mendes, Luis Barbosa Júnior e Douglas Bierhalz. Aos Postmortem Inc. Douglas Veiga, Mou Machado, Juliano Pacheco e Bruno Añaña. Aos TechMahinders Luis Filipe, Daniela Vignon, Wildson de Paula, Laura Melo, Thayna Santos, Mauricio Rego, César Rinaldi, Paulo Almeida, Jennifer Silva, Anna Karenina Cavalheiro, Nelson Stoppa, Ka Polete e em especial à chefinha Luísa Pogozelski. Da minha família, minha amiga e terceira avó Lucy Grupelli, minha amada e pefeita irmã Maria Joana Igansi, meu pai e Luciane Silva.

Aos amados seu José de Souza, Rosângela Carivalis, Carolina Carivalis e Scott. Aos restianos Davi Sarubbi, Luana Marmitt, Marcos Cordeiro (em especial por me apresentar o Zotero, sem o qual esta tese não teria sido concluída a tempo), Luana Alt, Gustavo Silveira, André Luzzardi, Lilian Bandeira, João Scaini e Karen Radünz. E demais aos FoFilers queridos Gustavo Teles, Guilherme Meneghello, Bruna Schneid e Jade Arbo por todo apoio e carinho.

Além de todos estes, há três figuras principais cuja ausência definitivamente impossibilitaria elaborar esta pesquisa e tese. A primeira é Alexandra Elbakyan, criadora da maior rede de compartilhamento gratuito de textos acadêmicos e científicos e indubitavelmente a maior responsável pela comunicação e avanço de pesquisas pelo mundo todo por sua rebelação contra os confinamentos capitalistas que excluem – quando não deliberadamente oneram significativamente – os próprios produtores de ciência e quem auxilia seu financiamento, ou seja, a população em geral dos frutos de seu trabalho e investimento. A segunda, nas presidências que mais investiram não apenas em educação mas na estrutura básica da sociedade brasileira a fim de que não apenas houvesse maior acesso ao ensino superior e à pesquisa acadêmica no geral, como também maior estabilidade e segurança nas classes mais baixas, permitindo a busca pela educação e profissionalização sem necessitar se preocupar com a fome, Luis Inácio “Lula” da Silva e Dilma Rouseff – que apesar de serem passíveis de diversas críticas, ainda assim estavam anos-luz à frente em progresso do que as tragédias que os sucederam. Por fim, mas a mais importante, minha melhor amiga e mãe Valéria de Oliveira Nogueira, que me apoiou ininterruptamente de infinitas maneiras desde antes de meu nascimento e sem a qual não teria a estrutura familiar e psicológica para tudo que fiz até o momento.

*Métodos “racionais” de tomada de decisão [...] lógica, matemática, teoria da probabilidade [...] são computacionalmente fracas: incapazes de resolver os problemas naturais adaptativos que nossos ancestrais tiveram que consistentemente resolver a fim de se reproduzir. [...] Esta performance pobre na maioria de problemas naturais é a razão primária da qual especializações de resolução de problemas foram favorecidas pela seleção natural ao invés de solucionadores de problemas de propósito geral. Apesar de afirmações amplas ao contrário, a mente humana não é pior do que racional [...], mas pode frequentemente ser melhor que racional.*

*(COSMIDES; TOOBY, 1994, p. 324, tradução nossa)*

## RESUMO

IGANSI, Luca Nogueira. **A irracionalidade da razão na tomada de decisão moral**. 2023. 196f. Tese (Doutorado em Filosofia) – Programa de Pós graduação em Filosofia, Instituto de Filosofia, Sociologia e Política, Universidade Federal de Pelotas, Pelotas, 2023.

Se analisará o quanto a racionalidade na agência humana se abstém de influências emocionais em diálogo com ciências empíricas através do conceito de tomada de decisão. Inicia-se com um panorama da importância da razão na racionalidade ética na história da filosofia, aprofundando-se posteriormente em teorias contemporâneas racionalistas ou dependentes de um conceito forte de razão. Estabeleceremos então as bases metodológicas filosóficas para a análise do problema a partir do naturalismo filosófico, disciplina metafilosófica que analisa o mundo de maneira científica ausente de elementos sobrenaturais. A filosofia experimental é sua subárea que dialoga e guia diretamente pesquisas na área empírica. Prosseguiremos então para as bases metodológicas interdisciplinares procurando estabelecer o estado-da-arte nas ciências da biologia evolutiva, que compreende pessoas humanas como animais com traços fisiológicos, cognitivos e genéticos selecionados a partir de contextos históricos naturais específicos; da economia comportamental, que compreende a razão do agente a tomar decisões como limitada em contraste com o viés clássico da racionalidade perfeita; da psicologia cognitiva, que compreende experimentalmente o estudo da mente humana e dos processos envolvidos em nossa cognição; e das neurociências comportamentais, que oferecem um sistema unificado destas outras disciplinas através da análise detalhada da neuroanatomia em nossas ações; utilizando “tomada de decisão” como palavra-chave interdisciplinar que equivale ao problema filosófico em questão e apresentando a teoria do processo duplo como fio condutor didático para a compreensão integrada dos temas. Uma vez tendo estabelecido este paradigma metodológico geral, partimos para as evidências empíricas acerca de nossas faculdades cognitivas relacionadas à tomada de decisão, investigando desde exemplos de influência de fatores externos como nojo e fome; neuropatologias como danos cerebrais, calosotomia e psicopatia; e do preconceito como manifestado no racismo e no sexismo. Partiremos posteriormente para debater estes dados com leituras filosóficas, assim como trazer outros elaborados por filósofos experimentais. Investigamos então como se dá o processo de geração do juízo moral a nível das neurociências comportamentais, debatemos sobre o fato do desacordo moral e a ética das virtudes, sobre experimentos de pensamento clássicos sob estas novas metodologias e evidências, aprofundando-nos no debate filosófico sobre neutralidade que vem a partir dos dados sobre preconceito, para então inferir acerca da importância de intuições na deliberação moral. Por fim, realizamos um panorama geral da discussão e concluímos que, apesar da centralidade da importância de influências emocionais e inconscientes no processo deliberativo, ainda assim a razão possui um papel orientador importante – todavia apenas se entendida como uma habilidade constantemente passível de erro e necessitando de manutenção.

**Palavras-Chave:** razão, racionalidade, juízo moral, tomada de decisão, filosofia experimental.

## ABSTRACT

IGANSI, Luca Nogueira. **The irrationality of reason in moral decision making.** 2023. 196f. Thesis (PhD degree in Philosophy) – Programa de Pós graduação em Filosofia, Instituto de Filosofia, Sociologia e Política, Universidade Federal de Pelotas, Pelotas, 2023.

We will analyze to what degree rationality in human agency abstains itself from emotional influences in dialogue with the empirical sciences through the concept of decision making. Beginning with a panorama of the importance of reason on ethical reasoning in the history of philosophy, developing afterwards in contemporary rationalist theories or ones dependant of a strong notion of reason. The philosophical methodological basis will then be established for the analysis of the problem through philosophical naturalism, a metaphilosophical discipline which analyzes the world in a scientific manner abstained from supernatural elements. Experimental philosophy is its subarea, which directly dialogues with and guides research in the empirical realm. We will proceed then to the interdisciplinary methodological bases aiming to establish the state-of-the-art in the sciences of evolutionary biology, which comprehends human persons as animals with physiological, cognitive and genetic traits selected through specific natural history contexts; of behavioral economics, which understands reason in the agent while making decisions as bounded, in contrast with the classical perspective of perfect rationality; of cognitive psychology, comprehending experimentally the study of human mind and the processes involved in our cognition; and of behavioral neurosciences, which offer an unified system of these previous disciplines through the detailed analysis of neuroanatomy in our actions; utilizing "decision making" as the interdisciplinary keyword which equates to the philosophical issue at hand, and presenting the theory of dual process as an educational common thread toward the integrated understanding of these themes. Once this general methodological paradigm is established, we proceed to the empirical evidences regarding our cognitive faculties related to decision making, investigating from the influences of external factors such as disgust and hunger; neuropathologies as brain damage, callosotomy and psychopathy; and prejudice as manifested in racism and sexism. Afterwards we proceed to debate these data with philosophical readings, as well as bringing more of them from experimental philosophers. We then investigate how the process of generating moral judgements takes place at the behavioral neuroscientific level, debate on the fact of moral disagreement and virtue ethics, on thought experiments under these new methodologies and evidences, furthering the philosophical debate on neutrality from the data on prejudice, to then infer on the importance of intuitions in moral deliberation. Finally, we provide a general overview of the discussion and conclude that, despite the centrality of the importance of emotional and unconscious influences in the deliberative process, still reason holds an important guiding role - however only if understood as skill constantly prone to error and requiring maintenance.

**Keywords:** reason, reasoning, moral judgement, decision making, experimental philosophy

## SUMÁRIO

<b>1.</b>	<b>UMA DICOTOMIA FICTÍCIA.....</b>	<b>14</b>
<b>2.</b>	<b>RAZÃO NA TRADIÇÃO FILOSÓFICA.....</b>	<b>25</b>
	2.1 Dos pré-socráticos à modernidade .....	25
	2.2 Racionalidade contemporânea e desacordo moral.....	38
<b>3.</b>	<b>METODOLOGIAS CIENTÍFICAS NA FILOSOFIA .....</b>	<b>51</b>
	3.1 Naturalismo filosófico.....	51
	3.2 Filosofia experimental .....	63
<b>4.</b>	<b>A ESTRUTURA CIENTÍFICA DA TOMADA DE DECISÃO.....</b>	<b>69</b>
	4.1 Um paradigma evolutivo.....	71
	4.2 Razão Limitada e Economia Comportamental.....	78
	4.3 Psicologia Cognitiva e Vieses.....	87
	4.3.1 <i>Psicologia e cognição.....</i>	<i>87</i>
	4.3.2 <i>Dissonância e outros vieses cognitivos.....</i>	<i>91</i>
	4.3.3 <i>Dois sistemas, dois processos .....</i>	<i>99</i>
	4.4 Neurociências.....	103
	4.4.1 <i>De buracos no crânio à ressonâncias eletromagnéticas ..</i>	<i>105</i>
	4.4.2 <i>Neuroanatomia.....</i>	<i>108</i>
	4.4.3 <i>Integrando conceitos .....</i>	<i>113</i>
<b>5.</b>	<b>IRRACIONALIDADE EM DADOS.....</b>	<b>122</b>
	5.1 Nojo e outros fatores externos .....	122
	5.2 Neuropatologias cognitivas .....	126
	5.2.1 <i>Atravessando o cérebro de Gage.....</i>	<i>126</i>
	5.2.2 <i>Síndrome do Cérebro Dividido .....</i>	<i>127</i>
	5.2.3 <i>Psicopatia.....</i>	<i>130</i>
	5.3 Preconceito e cognição .....	135
<b>6.</b>	<b>A AMORALIDADE DA MORAL.....</b>	<b>142</b>
	6.1 Situando a moralidade.....	144
	6.2 Desacordo, virtudes e situacionismo.....	149
	6.3 Experimentos e Pensamento .....	159
	6.4 Preconceito e Epistemologia .....	163
	6.4.1 <i>Racismo, ciência e filosofia .....</i>	<i>165</i>

6.4.2	<i>Localizando a epistemologia</i> .....	169
6.5	<b>O papel das intuições</b> .....	172
7.	<b>É POSSÍVEL REFUTAR KANT?</b> .....	176
7.1	<b>Como, afinal, ocorre o juízo moral?</b> .....	177
7.2	<b>A racionalidade da razão</b> .....	180
8.	<b>CONSIDERAÇÕES FINAIS</b> .....	184
8.1	<b>Limitações da pesquisa</b> .....	186
	<b>REFERÊNCIAS</b> .....	188

## 1. UMA DICOTOMIA FICTÍCIA

“Qualquer coisa levada a seu extremo lógico torna-se deprimente, quando não cancerígena.”

Ursula K. Le Guin

Uma vez que, como veremos, a filosofia experimental surge do naturalismo filosófico que, por sua vez, surge graças à filosofia analítica advinda da virada linguística, é importante iniciarmos a introdução deste trabalho através da análise do próprio título. Três substantivos são evidentes: “razão”, “irracionalidade” e “tomada de decisão moral”.

A “razão” é tida na tradição em filosofia predominantemente como uma faculdade de todos os indivíduos minimamente sãos e saudáveis, ou como coloca Korsgaard “pode ser usada para indicar sanidade, como quando pessoas dizem que alguém mentalmente transtornado ‘perdeu sua razão’” (2018, p. 294, tradução nossa)<sup>1</sup>. Em termos gerais, tende a ser considerada nossa capacidade intelectual de análise e processamento de informações, frequentemente associada à lógica. Em muitos casos, atribui-se à racionalidade humana um *status* único, como se nossa espécie fosse singularmente dotada desta capacidade neste nível. “Irracionalidade” é um contraste à “racionalidade”, que engendra uma aplicação desta razão sóbria; ou seja, pressupõe ordem, calculismo e frieza emocional a fim de não afetar a retidão lógica do tópico em questão. Assim, a “irracionalidade” indica seu oposto, uma ausência da clareza racional e presença de desordem, da confusão cega das paixões.

Evidentemente, este contraste é uma provocação: não irei propôr que há alguma “parte” da razão capaz de irracionalidade. Todavia, nossa estrutura cognitiva associada com aquilo que chamamos de “razão” não apenas é passível de falha, mas como esta tese pretende demonstrar, isto é a norma e não a exceção. E é aí que entra a próxima expressão: “tomada de decisão moral” une “tomada de decisão”, termo técnico interdisciplinar que iremos nos adentrar a seguir, e “moral”, termo de praxe no âmbito filosófico referente a análise do bom e do mau ético dentro de ações, costumes e juízos. Em suma, a junção destes

---

<sup>1</sup> No original: “*may be used to denote sanity, as when people say that a mentally deranged person has ‘lost his reason’*”.

dois termos demonstra que iremos focar na deliberação humana de maneira interdisciplinar para tratarmos de tópicos caros à ética e metaética.

É importante frisar aqui que este é, por definição, um empreendimento descritivo. O naturalismo filosófico como entendido aqui é por excelência uma posição metaética, e portanto não pretende normatividade. Em outras palavras, o intuito deste trabalho é aprofundar o conhecimento acerca do tópico em questão e estabelecer uma imagem do que pode ser dito de maneira cientificamente informada acerca do processo de ação de agentes. Não haverá aqui a tentativa de estipular leis, diretrizes ou quaisquer normas morais atreladas ou não a algum posicionamento ético maior. Haverá sim a crítica ou elogio de determinadas posições de acordo com o que os dados nos ilustrarão, mas nenhum comprometimento, refutação ou sugestão de nenhuma teoria. É um trabalho investigativo, portanto não vou adotar uma posição moral substantiva a fim de evitar viéses.

Quanto à razão, há frequentemente uma dicotomização com “emoção” que perpassa a totalidade da história humana – especialmente nas tradições filosóficas que tratam da agência individual. A razão, como sugere sua definição inicial, tende a ser prioritária por tratar-se de reger processos objetivamente analisáveis, de caráter computacional cuja proficiência é inquestionavelmente correta, verdadeira e infalível. Nossas emoções, por outro lado, tendem a ser definidas como processos secundários, inferiores. Seriam análogas a uma primitividade instintiva e por si só apenas úteis à arte ou à esfera subjetiva, ou à razão apenas quando em papel qualificador ou catalisador. De acordo com Francisco de Lima e Nythamar de Oliveira,

[e]sse tipo de tradição perpassou grande parte do pensamento ético ocidental e está presente em Sócrates, Platão, Aristóteles, Agostinho, Tomás de Aquino, Descartes, Kant, Hegel, para citar os mais emblemáticos. (2019, p. 2)

Assim, predominantemente, a razão possui primazia sobre as emoções no que tange a escolha moral nas teorias éticas, metaéticas e epistemológicas até o momento.

Certamente, como é comum na filosofia, existem muitas variações e exceções a estas perspectivas. Todavia, em termos gerais ela é tão incorporada

em nossos paradigmas que mesmo nas mídias ficcionais é um tema constantemente revisitado. No universo fictício de *Star Trek*, por exemplo, este contraste manifesta-se no povo alienígena *Vulcano*, origem do popular personagem *Spock*, que, ao contrário dos humanos, reprimem suas emoções e servem-se unicamente da lógica e da razão para a ação. Esta característica é tomada como virtude de seu povo em um culto semi-religioso predominante, e seu desvio por questões de passionalidades é causa frequente de represália tanto entre si como de outros povos.

No entanto, esta é apenas uma faceta deste povo icônico da ficção científica. Ao aprofundarmo-nos em suas origens históricas neste universo ficcional, torna-se clara sua necessidade de exaltação da comedido racional: antes de suas doutrinas racionais, os vulcanos eram passionais, violentos, destrutivos a ponto de beirar a sua própria aniquilação enquanto espécie. Independente da época em que as diferentes iterações da série se passa, a mera aproximação a estados emotivos ou carência de disciplina nos seus ritos meditativos e procedimentais pode abalá-los e levá-los a episódios de descontrole e imprudência exacerbados.

Evidentemente, não somos vulcanos, e muito menos vivemos no universo fictício de *Star Trek*. Mas como diria Le Guin,<sup>2</sup> o papel deste tipo de obra não é extrapolar ou fazer previsões – “[a] ficção científica não prevê; descreve.” (2014, p. 4) Descreve como tendemos a ver o universo e como tentamos projetar nossa imagem de nós mesmos e nosso mundo. No eixo “razão/emoção”, os vulcanos são a epítome da visão racionalista: manejam os impulsos passionais a fim do uso mais otimizado da lógica no processo deliberativo. Todavia, em múltiplas instâncias deste seriado ficcional, esta perspectiva falha em compreender a plenitude da ação humana (e de outras espécies) em todos os seus fatores, tornando-se mais um empecilho do que uma virtude. Após a realização desta pesquisa, acredito que o mesmo ocorre no viés filosófico, pela miríade de evidências que demonstram a irracionalidade da razão humana; ou como coloca Jonathan Baron, a “[r]acionalidade [...] não é o mesmo que acurácia, e irracionalidade não é o mesmo que erro” (BARON, 2007, p. 63). E, como diz Spock, “fatos insuficientes sempre convidam ao perigo” – neste caso, o perigo

---

<sup>2</sup> Agradeço à Natália C. de Souza e Jade Arbo por me apresentarem esta obra e me enriquecerem com suas discussões.

de focar a epistemologia e a ética em apenas um dos múltiplos fatores envolvidos em nossa ação e criar uma demanda normativa para outrem a partir disto.

O objetivo geral deste trabalho, portanto, é tornar clara a importância de uma compreensão total dos fatores envolvidos em certas partes da investigação filosófica, especificamente na epistemologia moral, no que tange o processo deliberativo da ação humana. Inúmeros avanços tecnológicos e metodológicos do último século permitiram que áreas outrora íntimas – ou mesmo parte – da filosofia como sociologia, economia, psicologia, antropologia e outras obtivessem acesso sem precedentes à novas informações sobre nosso funcionamento enquanto espécie. Uma miríade de dados sobre nosso comportamento rastreáveis evolutivamente até minúcias de nossa neurofisiologia aplicados em testes controlados nos provêem constantemente uma imagem mais completa em prol de nosso próprio entendimento sobre nós mesmos.

Embora o diálogo interdisciplinar nestas áreas seja a norma, ainda no século XXI é minoria a comunidade filosófica que preza o contato com outras ciências em prol da investigação plural de determinados tópicos; e menor ainda a que de fato a realiza em suas pesquisas. Minha intenção aqui não é produzir um ensaio desta discussão na história da filosofia – ou ainda uma história da decisão nas ciências –, mas procurar avançar no processo de montagem deste quebra-cabeça de informações acerca do comportamento humano no que tange a maneira que tomamos decisões. É procurar identificar e estabelecer fatores para uma base mínima abrangente de critérios para se realizar pesquisa no campo da ação em filosofia, que há muito mantém-se em sua poltrona, distante do progresso de outras áreas que justamente surgiram dela própria para analisar tais tópicos com maior atenção.

Estes são os principais motivos para a utilização da expressão “tomada de decisão” (*decision making* na literatura internacional), em contraponto a termos como ação, agência, racionalidade ou racionalidade prática. Ainda que sejam predominantes na bibliografia filosófica de epistemologia moral e ética, são termos que condizem e vão ao encontro das análises de outras áreas sobre este fenômeno tão caro à filosofia como um todo: o processo de decisão. “Tomada de decisão” utiliza-se amplamente em outras áreas, gerando resultados significativos em qualquer banco de dados das mesmas, sendo

intersecção comum a todas ao se tratar de temas relativos a este processo. Evidentemente, nem tudo é de relevância filosófica, e muito menos ao tratarmos especificamente de epistemologia moral. Mas mesmo sendo uma parcela diminuta do total, ainda assim é um torrencial de informações significativas para entendermos a plenitude dos processos envolvidos desde a percepção de estímulos até a realização da ação *per se*.

E dentro deste paradigma, se encontra também o estudo sobre “viés cognitivo”, área de pesquisa que foi estabelecida por Daniel Kahneman e Amos Tversky (DANIEL KAHNEMAN; SHANE FREDERICK, 2002; KAHNEMAN; TVERSKY, 1972) na psicologia cognitiva alinhada à economia comportamental e, mais recentemente, as neurociências, no contexto do estudo da tomada de decisão. Estes estudos surgiram a partir da revolução psicológica de Herbert Simon na economia (GIGERENZER; SELTEN, 2002; SIMON, 1957; WHEELER, 2020), que trouxe limitações à racionalidade plena da economia clássica na tomada de decisão de agentes (sejam humanos ou máquinas de computação), e que culminou na introdução da noção de “dissonância cognitiva”, introduzida por Leon Festinger, Henry Ricken e Stanley Schachter (2008) em 1956, inaugurando a pesquisa supracitada de vieses. Atualmente, o estudo destes vieses cognitivos abarca especificamente as maneiras em que diferentes influências externas, tendências ou inclinações de algum âmbito influenciam nossa percepção e juízo e como afetam nossas ações como um todo.

É importante esclarecer que a utilização de “cognitividade” aqui difere das variações utilizadas na tradição filosófica: exceto quando referenciado, as variações de “cognitivo”, “cognição” e afins farão referência à seu uso nas ciências cognitivas em geral, cujo termo torna-se representativo de processamento de informações. De maneira nenhuma me atrelo necessariamente à posições cognitivistas, por exemplo no viés epistemológico, quando me refiro à cognitividade ou à ciências cognitivas em geral. Cognição, neste contexto, nada mais sinaliza além da apreensão idiossincrática de informações e seu processamento psicológico; e embora possamos deduzir algumas ramificações que tais tipos de estudos podem trazer para a filosofia posteriormente, sua presença no texto não atrela seu uso de modo algum ao cognitivismo ou não-cognitivismo epistemológico e afins. Não é o intuito deste trabalho categorizar ou afirmar que determinada teoria filosófica é

definitivamente superior ao tratarmos da bibliografia acerca do tópico; mas sim, ilustrar o debate e estabelecer critérios mínimos para pesquisas na área devem partir a fim de construir teorias filosóficas de relevância científica, e quiçá mesmo teorias científicas de relevância filosóficas.

No que especificamente tange concepções de racionalidade, Wheeler (2020, p. 72–74) caracteriza sete definições principais deste conceito que, segundo ele, pode ser considerado um "conceito-mala", ou seja, que deve ser sempre desempacotado antes de ser tratado: (1) como *coerência* entre ação e objetivo em um sentido bayesiano; (2) como pressuposição de *intencionalidade comunicativa*, ou seja, a pressuposição de uma teoria da mente<sup>3</sup> de outras pessoas capaz de interpretar causalmente o mundo à sua volta; (3) a razão *instrumental* humeana, de uma avaliação do comprometimento individual como crença gerada por evidências do mundo; (4) a *racionalidade substantiva* de Weber, inspirada por Kant, que pressupõe propósito além da capacidade instrumental; (5) a concepção *pragmatista* como em Peirce e Dewey, que desenvolve a perspectiva instrumental humeana para um foco no ato de dúvida e mudança de crença ao invés de mera reflexão sobre uma crença já existente; (6) uma concepção *etológica*<sup>4</sup> envolvendo a administração do comportamento baseado no processamento de informações complexas, bastante inspirada nos pragmatistas e Hume; e, por fim, (7) a noção de "ausência de defeito", ou de "irracionalidade" no contexto do processo deliberativo, semelhante à concepção bayesiana de coerência pois, se há coerência, há a ausência de defeito, e portanto há racionalidade. Assim como o próprio autor reconhece que estas definições, por mais variadas, não abarcam a totalidade de diferentes delineações do que configura a "racionalidade" como característica de um agente, também não pressuponho exaurir ou mesmo eliminar aqui todas as definições da mesma.<sup>5</sup>

---

<sup>3</sup> Ao contrário do que o nome sugere em um contexto filosófico, "teoria da mente" (no original "Theory of Mind" ou "ToM") em termos psicológicos ou neurológicos refere-se à capacidade de atribuir intencionalidade e capacidade reflexiva a outrem.

<sup>4</sup> Área da biologia que trata do estudo comportamental comparativo animal. É importante não esquecermos que seres humanos também são animais.

<sup>5</sup> Nota-se também, como o autor também aponta, que este problema definicional afeta qualquer formulação positiva de racionalidade, dada a ampla variabilidade de concepções frequentemente mutuamente excludente. Todavia, fugiria bastante ao escopo deste trabalho revolucionar toda a terminologia de todas as áreas de pesquisa – filosóficas ou não – que envolvem o problema da razão. Para uma ilustração do que seria algo semelhante a uma tentativa assim no contexto da psicologia cognitiva, v. o fim de cada capítulo em KAHNEMAN, 2013.

Desta maneira, quando utilizar estes outros termos neste trabalho, a não ser quando especificando diferentes significados específicos por autores ou linhas de pensamento, estarei me referindo a esta concepção de racionalidade baseado sumariamente em (6), ou seja, uma razão instrumental situada em um contexto pragmático aplicável em um viés predominantemente científico como delineado por Damásio, de

reason as the ability to think and make inferences in an orderly, logical manner; and rationality as the quality of thought and behavior that comes from adapting reason to a personal and social context. I do not use reasoning and decision making interchangeably since not all reasoning processes are followed by a decision. (DAMASIO, 1994, p. 269)

Da mesma maneira, utilizaremos esta perspectiva ao tratarmos de emoções e sentimentos. Frequentemente utilizados intercambiavelmente, sem distinção, tomaremos “emoções por uma coleção de mudanças ocorrendo em ambos, no cérebro<sup>6</sup> e no corpo, geralmente incitados por um conteúdo mental particular” (DAMASIO, 1994, p. 270)<sup>7</sup>, enquanto sentimento seria “the perception of those changes” (1994, p. 270)<sup>8</sup>. Ou seja, o sentimento é a percepção consciente das emoções, que por si pode ser acessado pela memória e configura “a disposition to experience emotions under certain circumstances” (SILVEIRA, 2016, p. 191) – posição defendida por Jesse Prinz (2007a). As emoções ocorrem de maneira pública, sendo visíveis por outros através de respostas fisiológicas como a ruborização e expressões do rosto, vocalização de sons, ereção folicular e outros fatores observáveis, enquanto o sentimento ocorre privadamente no uso da memória. (DAMASIO, 2003, p. 29)

Este paradigma metodológico que tomarei para este trabalho é o da filosofia experimental, veia do naturalismo filosófico, perspectiva de recentes décadas que procura aproximar o empreendimento filosófico do viés científico distanciando-se ontologicamente e epistemologicamente de elementos não-

---

<sup>6</sup> Utilizaremos predominantemente o termo “cérebro” em contraste a “encéfalo”, “sistema nervosa central” etc. a fins de clareza, uma vez que se trata de um texto fora das áreas usualmente relacionadas com o estudo de neuroanatomia.

<sup>7</sup> No original: “*emotion for a collection of changes occurring in both brain and body, usually prompted by a particular mental content*”.

<sup>8</sup> Damásio se refere ao termo “*feeling*” ao invés de “*sentiment*”, porém denota o mesmo sentido; tanto que Francisco de Lima e Nythamar de Oliveira o traduz como “sentimento”. (DE LIMA; DE OLIVEIRA, 2019, p. 7–8)

naturais para a análise do comportamento humano (ALFANO; LOEB, 2017; KNOBE; NICHOLS, 2008; STURGEON, 2007; IGANSI, 2014a, 2014b). Em minha pesquisa de mestrado (IGANSI, 2014a) procurei analisar a validade e importância metaética do naturalismo moral pós-humeano através da aplicação da falácia naturalista conforme os escritos póstumos de Moore e a síntese de Dall'Agnol em termos lógicos contemporâneos (IGANSI, 2014c). Desta análise concluí que o viés deste empreendimento naturalista acerca do fenômeno moral possui como foco central o entendimento factual dos fatores naturais envolvidos no comportamento humano – ou seja, é uma teoria de cunho descritivo, e, portanto, não pretende criar normas. Isto gera um problema em que, assim, não é útil *prima facie* para guias práticas de ação na esfera moral no que tange os tópicos que abrange.

Por outro lado, análises de teorias não-naturalistas como as de Moore e Rawls, ou como Jesse Prinz faz de Aristóteles, Bentham, Platão, Mill dentre outros (PRINZ, 2007a, cap. 4, 2009; cf. também SINNOTT-ARMSTRONG; MILLER, 2017) e Joshua Greene de Kant (GREENE, 2007, 2017) apontam o problema inverso: pretendem elas esta força normativa, útil para as diversas áreas que necessitam de um parâmetro prático imediato, todavia carecem de um diálogo informado com a vasta e sempre crescente biblioteca de evidências empíricas acerca das diversas áreas de agência moral humana que pretendem reger. Uma vez que toda teoria normativa pretende atuar sobre indivíduos reais em um contexto sócio-histórico-psicológico real, a falha em dialogar com estas informações – o ato de limitar-se à uma “filosofia de poltrona”, como associado pelos experimentalistas (KAHANE, 2013; KNOBE; NICHOLS, 2008) – as rende no melhor dos casos apenas como hipóteses advindas de observações idiossincráticas. Teorias filosóficas que tangem temáticas empiricamente testáveis acerca de nossa natureza comportamental – de pressupostos epistêmico-ontológicos à ação em si – sem o devido crivo experimental são, no máximo, hipóteses desinformadas; e no mínimo, irresponsáveis. E o contrário também é significativo: teorias científicas sobre a ação sem devida análise filosófica nada mais são que dados enviesados por paradigmas filosóficos sem justificativa devida, afinal “não há tal coisa como ciência ausente de filosofia;

existe apenas ciência cuja bagagem filosófica é levada a bordo sem exame“(DENNETT, 1995, p. 21, tradução nossa)<sup>9</sup>.

De maneira mais precisa, dentro da tentativa de compreensão da forma como a ação se dá sob um viés epistemológico, o objetivo desta tese é entender o quanto do processo deliberativo moral é atribuível à uma capacidade racional objetiva. Além disso, também o quanto deste processo é sujeito a fatores externos alheios, secundários ou inconscientes através de um diálogo interdisciplinar com outras áreas do conhecimento que tratam de temas que se sobrepõe com áreas filosóficas quanto a este tema. A partir desta proposta, partem as seguintes questões: (i) quais fatores externos, ou “não morais”, influenciam a concepção do juízo moral, e como identificá-los e contorná-los, se possível? (ii) O que caracteriza a deliberação moral em contraponto a uma deliberação não-moral? (iii) é possível depender da razão para a elaboração de juízos morais?

De maneira análoga à Herbert Simon (SIMON, 1957, p. 99–100), o objetivo aqui não é “refutar” teorias racionalistas ou reduzir o todo da epistemologia moral à psicologia, economia e neurociências, mas oferecer um porto seguro no meio do caminho entre ambos os extremos a fim de melhores discussões sobre temas tão caros à pesquisa de todas estas áreas – ou seja, facilitar que esta nova camada de informações e metodologia acrescente à nossa estrutura filosófica, não que a substitua.

Iniciaremos esta tarefa no capítulo 2 a partir de uma análise de concepções de razão na tradição filosófica de maneira breve, focando por fim na interpretação de Hume que será a base do viés filosófico que abraçaremos neste trabalho. O objetivo desta primeira parte é analisar como esta relação dicotômica entre razão e emoção perdura e cede à primeira predominantemente, mesmo em casos icônicos do contrário como nas primeiras formulações de hedonismo. Se estabelece então uma posição padrão generalizada acerca do papel da racionalidade na deliberação de acordo com a tradição filosófica, que servirá como pano de fundo para compararmos com a perspectiva contemporânea cientificamente informada que desenvolveremos no resto do trabalho.

---

<sup>9</sup> Original: “*There is no such thing as philosophy-free science; there is only science whose philosophical baggage is taken on board without examination*”.

No capítulo 3 trabalhamos o naturalismo filosófico, historicamente a consequência teórica atualizada para o século XX da posição humeana do capítulo anterior. O naturalismo é o paradigma teórico geral no qual este trabalho se situa, mas mais especificamente se dá na filosofia experimental, que traz o naturalismo filosófico para tratar diretamente com outras ciências não estritamente filosóficas a fim de um aprofundamento descritivo maior de fenômenos envolvidos em problemas filosóficos.

Assim, encerraremos o terceiro capítulo com uma explicação básica desta vertente para então seguirmos para estas outras ciências importantes para o problema em questão – o papel da razão na ação – no capítulo 4. Apresentaremos nele os conceitos interdisciplinares que serão chave para investigar o problema da relação entre razão e emoções no contexto da deliberação ético-epistemológica humana. O termo principal, como vimos acima, será a tomada de decisão, amplamente utilizado na psicologia, economia e outras ciências comportamentais como “termo guarda-chuva”, genérico para a interconexão entre as diferentes áreas do conhecimento sobre este fenômeno. Iniciaremos analisando brevemente o histórico de teorias da ação fora do contexto filosófico a partir de áreas como a economia comportamental com e psicologia a partir de Herbert Simon e a crescente compreensão nessas áreas de fatores que pareciam “restringir” a racionalidade em certos testes comportamentais. Adentraremos na psicologia, sociologia e neurociências através da análise do viés cognitivo, o supracitado termo-chave associado à variedade de fatores externos à nossas faculdades cognitivas que as influenciam com ou (predominantemente) sem nosso conhecimento consciente.

Este capítulo é importante pois, tratando-se de um trabalho advindo de e direcionado para o público filosófico, dará as ferramentas básicas para compreendermos como interpretar dados destas áreas para a investigação dentro da filosofia. Prosseguiremos então para o capítulo 5, onde investigaremos tais dados a fim de estabelecer o estado-da-arte de como estas outras ciências compreendem o que entenderíamos como razão no processo de decisão. Versaremos brevemente por uma miríade de diferentes evidências acerca de nossas faculdades agenciais. Desde patologias cognitivas como psicopatia e a síndrome do cérebro dividido até como preconceitos como racismo e machismo influenciam nossa percepção e decisão, passando até mesmo por elementos

alheios dificilmente antes tratados em teorias filosóficas como a influência do nojo em nosso ajuizamento.

Prosseguiremos para o capítulo 6, então, para aplicarmos o método da filosofia experimental sobre estes dados e dialogar com outras evidências dela própria a fim de compreendermos suas implicações para o estudo da ação na filosofia. Aqui, estes dados dialogarão com estudos específicos e suas evidências especificamente da área da filosofia experimental ou de debates inseridos no contexto filosófico. Dentre estes, veremos alguns acerca do problema do caráter na ética das virtudes e mesmo de nossas faculdades cognitivas dentro do escopo das grandes escolas éticas em geral com neuroimageamento, etologia e outras ferramentas. Por fim, abordaremos o intuicionismo social enquanto a primeira grande escola da filosofia experimental, que trouxe muitas das evidências presentes neste trabalho.

Assim, encerraremos as discussões no capítulo 7, onde faremos as considerações finais desta trajetória filosófico-científica acerca da razão na tomada de decisão para propôr uma estrutura mínima para o debate acerca do agente e da ação na filosofia minimamente adequadamente informado com o montante de pesquisas de outras ciências. Apontaremos as limitações do empreendimento interdisciplinar na filosofia e como este tópico ainda mantém-se aberto para debate – como qualquer problema no viés científico. Verificaremos que, de fato, a importância da análise dos fatores externos à razão no processo de ajuizamento moral é crucial para a ética; todavia, nosso poder de racionalidade não é completamente descartável uma vez que, com devida informação e consciência destes fatores, é capaz de nos auxiliar a compensá-los e nos permitir agir de maneira mais adequada. Analizaremos também brevemente um pouco sobre pesquisas recentes acerca de teorias da razão empiricamente informadas e como podemos proceder o debate a partir disso.

Por fim, no capítulo 8, realizaremos as considerações finais em um panorama do que foi visto pelo texto e da importância da filosofia experimental e da análise interdisciplinar de problemas filosóficos para a pesquisa de seus problemas, além de apontar quais as limitações desta pesquisa e onde podemos desenvolver ainda mais estudos sobre o problema.

## 2. RAZÃO NA TRADIÇÃO FILOSÓFICA

As regras da lógica, os axiomas de probabilidade, os princípios da teoria da utilidade – os humanos deprezam a todas, e fazem tal por via de regra. Mas somos irracionais ao fazê-lo? Isto depende no que significa ser racional.

(WHEELER, 2020, p. 71, tradução nossa)

Toda a teoria filosófica que trata da deliberação parte de determinados pressupostos acerca das capacidades epistêmicas do agente. Por mais que a terminologia de “racionalismo” ou “razão” não seja utilizada ou diretamente elicitada, há um uso comparativamente comum em tipo aos racionalismos mais puros – embora em gradação menor. Neste capítulo procurarei não apenas identificar como a perspectiva que utilizamos aqui de “racionalidade” está predominantemente presente em teorias filosóficas como também estabelecer um panorama geral do papel deste conceito através da tradição filosófica.

O objetivo não é de maneira alguma exaurir todas as definições de racionalidade possíveis, e da mesma maneira também não criar uma definição por demais abrangente cuja crítica não se aplique a nenhuma teoria ou escola específica na prática. O empreendimento em vista neste capítulo será de delinear brevemente o problema da racionalidade na história da filosofia, focando em alguns posicionamentos mais representativos a fim de delinear hipóteses gerais que abarquem aspectos principais de teorias com um foco maior na razão sobre outros fatores no processo deliberativo.

### 2.1 Dos pré-socráticos à modernidade

“A mente do racionalismo ocidental é a filha leiga da alma salvável do cristianismo.”

Paulo Leminski

Iniciaremos esta seção tratando de algumas principais concepções de razão e racionalidade na filosofia antiga, dos pré-socráticos a Aristóteles e seu desenvolvimento no medievo. Posteriormente, analisaremos alguns autores empiristas e racionalistas na modernidade, que nos trazem definições e concepções mais relevantes para o debate contemporâneo. Por fim, trataremos especificamente de David Hume e Immanuel Kant, por ambos serem referências

cruciais para o desenvolvimento de múltiplas concepções contemporâneas de ação, razão e racionalidade.

Na antiguidade, diz-se de uma lenda antiga à época de Pitágoras e seu culto protorracionalista, cerca de um século antes de Sócrates, que Hipaso de Metaponto foi assassinado (pelos deuses em certas versões, pelos outros fiéis em outras) por causa de um cálculo matemático (SMITH, 2019, p. 1–2). A Escola Pitagórica abraçava os números, a matemática e a geometria como chaves para o entendimento do cosmos. Assim, os números “naturais” (como os inteiros positivos eram denominados na época) possuíam uma relação perfeita e racional entre si, e através deles poderíamos entender a realidade e tratar de problemas filosóficos importantes. Assim, reza a lenda, quando se descobriu que a relação da diagonal do quadrado com os seus lados era incomensurável, ou seja, que ao contrário de um número inteiro, levava a uma quantidade infindável de casas decimais – sendo assim *irracional* – Hipaso fez questão de divulgar esta informação. Os pitagóricos, inconformados com a possibilidade de outros descobrirem que o fundamento maior de sua escola baseava-se em caos, na irracionalidade, ordenaram a morte de Hipaso por afogamento. Todavia, como as melhores lendas, “provavelmente nunca aconteceu, mas como toda boa lenda não precisa ter acontecido de fato para comunicar sua lição profunda”<sup>10</sup> (SMITH, 2019, p. 2). A ficção novamente nos traz uma imagem descritiva significativa deste posicionamento.

Ainda que a emoção e outros fatores sejam considerados na deliberação, o viés racionalista sempre tendeu a dominar a perspectiva epistemológica e ética na filosofia – da prudência aristotélica até o fato da razão de Kant, dos pré-socráticos até os consequencialistas contemporâneos. Mesmo hedonistas e emotivistas como os epicuristas e alguns consequencialistas, conhecidos por sua priorização do prazer e da felicidade, os viam apenas como fins de um processo de deliberação essencialmente racional. Epicuro, por exemplo, pai do hedonismo filosófico, exalta a prudência e sabedoria na identificação do prazer em prol da felicidade.

---

<sup>10</sup> No original: “[i]t almost certainly never happened, but like any good legend it does not need to have happened in fact in order to convey its profound lesson”.

Quando então dizemos que o fim último é o prazer, não nos referimos aos prazeres dos intemperantes ou aos que consistem no gozo dos sentidos, [...] mas ao prazer que é ausência de sofrimentos físicos e de perturbações da alma. Não são, pois, bebidas nem banquetes contínuos, nem a posse de mulheres e rapazes, nem o sabor dos peixes ou das outras iguarias de uma mesa farta que tornam doce uma vida, mas *um exame cuidadoso* que investigue as causas de toda escolha e de toda rejeição [...]. (EPICURO, 1999, p. 42–45, grifo noss)

Apesar do fim máximo ser a felicidade, é apenas através da análise atenciosa de nossas paixões que somos capazes de identificar os prazeres adequados a nos levarem à melhor vida.

O epicurismo surgiu em contraste às teorias de Platão, talvez o primeiro “racionalista” em um sentido forte do termo. Platão defendia que apenas através da intuição divina pela razão pura (especulativa ou teórica, em oposição à razão prática) seríamos capazes de realmente entender o mundo. As sensações corpóreas do mundo físico são negadas, e este é entendido como mera cópia frugal das Formas ideais que habitam o mundo sobrenatural das ideias. Ações eticamente boas, para Platão, também dependeriam da compreensão racional da forma pura de Bom – “a existência das Formas, e particularmente a Forma do Bom, explica a existência do valor ético e da obrigação”<sup>11</sup> (PUTNAM, 2005, p. 17).

Aristóteles, por outro lado, frisava que praticar a virtude era um hábito em conformidade com a “função própria” (*ergon*) de nossas faculdades enquanto animais sociais através de

um tipo de razão que direciona e estrutura desejos e emoções [...] (EN I.7.1098a3-5 cf. I.13.1102b11-1103a10.) [...] Logo, primeiro, um ser humano floresce apenas se emociona-se e age com razão e de acordo com virtudes. (CHEW, 2009)

Este viés, assim como o do estoicismo, manteve-se bastante forte também na era medieval graças aos esforços árabes de estudo da filosofia grega. Ambos, colocando a racionalidade e a virtude acima das paixões, serviam bem ao paradigma cristão predominante na época. Através da leitura que Tomás de Aquino fizera de Aristóteles que esta perspectiva naturalista primitiva da filosofia antiga culminou nas leituras contratualistas modernas (Hobbes, Rousseau, etc.)

---

<sup>11</sup> No original: “*the existence of the Forms, and particularly the Form of the Good, explains the existence of ethical value and obligation*”.

e, por conseguinte, no jusnaturalismo. Isto se deu através da identificação de que então teríamos uma “racionalidade natural”<sup>12</sup> que por sua vez identificaria estes princípios “naturais” em um intuicionismo bastante próximo de Platão. O fato da humanidade ser composta criaturas racionais a submeteria a “uma lei que eles, diferentemente de criaturas não-rationais, poderiam seguir ou ignorar”<sup>13</sup> (DARWALL, 1995, p. 5, tradução nossa). A concepção racional de vias platônicas, portanto, predominou até a modernidade – variando apenas a perfeição das Formas puras que precisamos alcançar para concepções divinas.

Na modernidade, grandes eventos como a Revolução Copernicana, o Iluminismo, a Revolução Industrial e a Revolução Científica foram palco para o desenvolvimento de teorias filosóficas com maior foco no entendimento científico da época. De Francis Bacon a René Descartes temos muito da estrutura na qual muitas teorias contemporâneas se desenvolveram. Bacon fora o pai do movimento empirista moderno, propondo uma versão do método científico baseado no ceticismo e no raciocínio indutivo e estimulando uma interconectividade entre filosofia, ciência e política. Segundo Darwall, este movimento pretende tratar de problemas filosóficos com “uma epistemologia empirista e uma metafísica naturalista” (DARWALL, 1995, p. 14, tradução nossa), o que Bacon procurou representar em seus escritos. Todavia, isto fora sumariamente em viés religioso, com a intenção última de restaurar a humanidade para antes de sua “queda do paraíso” através da ciência e da filosofia como transformadoras de mundos – não apenas como ferramentas. Assim, para Bacon, há uma valoração ética em agir conforme a lógica e a razão: “[w]e need to apply self-discipline and rational assessment, as well as restraining our passions, in order to lead an active moral life in society (KLEIN; GIGLIONI, 2020, p. 39).

Seguindo no período, Baruch Spinoza e René Descartes, por outro lado, foram alguns dos maiores racionalistas da modernidade. Ambos compartilhavam com Bacon a admiração pelo divino. No entanto, mesmo sob a mesma

---

<sup>12</sup> Como veremos na próxima seção, apesar do nome é importante não confundir esta perspectiva com o naturalismo filosófico contemporâneo: o jusnaturalismo baseava-se em um racionalismo ontológico idealista de forte normatividade, enquanto mesmo as teorias mais racionalistas do naturalismo contemporâneo pretendem no máximo à descritividade numa esfera aquém de entidades sobrenaturais.

<sup>13</sup> No original: “a law that they, unlike nonrational creatures, can follow or flout”.

designação geral de pensamento, ambos também possuíam visões diametralmente opostas.

Spinoza era um monista cujas ideias de consciência e natureza foram relevantes séculos depois – embora indiretamente e de outras maneiras – para William James, Sigmund Freud e mesmo Charles Darwin (DAMASIO, 2003, p. 12–13). Spinoza desenvolveu sua teoria baseado numa substância única divina que permeia tudo que existe. Assim, nós, o seu deus cristão, e toda a realidade em nossa volta seriam vários aspectos diferentes dessa mesma substância – análoga a algumas concepções pré-socráticas. Corpo e mente seriam, de tal maneira, apenas diferentes manifestações de uma mesma coisa, agindo concomitantemente para o auto-melhoramento em prol da felicidade. A razão é, como predominantemente na tradição filosófica, uma dádiva divina a qual utilizamos para tais fins – o que marca sua alcunha racionalista.

Todavia, ao contrário desta tradição, Spinoza defende que a razão por si só, isolada de outros fatores, não é capaz de controlar os afetos, ou paixões: apenas paixões estimuladas pela razão são capazes de controlar outras paixões negativas. Evidentemente, porém, paixões negativas são assim, por definição, irracionais. Apesar de ainda manter-se no padrão tradicional de associar a razão ao bem moral e apenas aspectos negativos à irracionalidade, Spinoza traz uma diferente perspectiva das relações entre racionalidade e emoções. Não apenas dá devido peso a fatores além de nosso próprio poder racional, como também preza a análise destes tópicos em um viés científico (DAMASIO, 2003, p. 280). Ironicamente, desta forma, assemelha-se mais com os empiristas do que com os racionalistas aos quais ele está associado.

Descartes, como múltiplos de seus contemporâneos modernos, pretende investigar um método para aquisição de conhecimento de maneira científica. Para isto, busca as bases mais sólidas, inquestionáveis, nas quais pretende fundamentar todo o conhecimento humano (crenças verdadeiras justificadas) em axiomas basilares que garantem a veracidade de proposições mais complexas sobre o mundo – uma perspectiva *mentalista* que foi herdada pelos empiristas e positivistas posteriores (KIM, 1988, p. 384). Uma vez que as sensações fisiológicas são passíveis de questionamento pelo fato de podermos nos enganar quanto a nossa percepção de cores, sensações térmicas, sons etc., a única raiz

indubitável é nossa própria capacidade de duvidar – ou seja, nossa própria consciência.

Assim, nossa mente, ou alma, “*res cogitans*” e corpo, “*res extensa*” seriam duas substâncias diferentes, interconectadas pela glândula pineal em nosso cérebro. Esta perspectiva dualista é central para o racionalismo, pois embora muitas concepções ignorem a divisão de substâncias, ainda há uma separação em um sentido forte entre as faculdades racionais e as emoções em geral. Assim, “raciocínio e juízo moral, e o sofrimento que advém da dor física ou abalamento emocional podem existir separadamente do corpo.”<sup>14</sup> (DAMASIO, 1994, p. 250).

Antiteticamente a Descartes, David Hume talvez seja o autor mais relevante na história da filosofia para o paradigma que trataremos neste trabalho. Apesar de compartilhar da mesma escola de pensamento e sua afinidade com a metodologia científica, Hume separava-se diametralmente de Bacon – e da majoritária tradição filosófica – ao identificar um maior peso às emoções (paixões) do que à razão na deliberação. Acima dos próprios hedonistas o autor reconhece a primazia das emoções sobre a agência humana, o que é singular dado seu contexto histórico da glorificação da faculdade da razão.

Hume desenvolveu sua teoria acerca do comportamento humano em um contexto bastante interessante. Seguindo a revolução copernicana no século 16, a revolução científica – movimento do qual Hume participara – atingiu seu ápice. Não por acaso, era um movimento que abraçava a metodologia empirista como foco de ensino e pesquisa em todas as áreas cabíveis. Este aspecto experimental prático exalta-se já na leitura do título de sua primeira grande obra, o *Tratado da Natureza Humana* (HUME, 1739, 2009), cujo subtítulo lê-se “Como uma Tentativa de Introduzir o Método Experimental de Raciocínio à Temas Morais.”<sup>15</sup> Ou seja, antes de desenvolver sua teoria ética, Hume tem a intenção de investigar a humanidade e sua moralidade através do método empírico, invertendo o procedimento tradicional filosófico de uma investigação racional pura para uma análise factual dos processos envolvidos. Ele

---

<sup>14</sup> No original: “reasoning, and moral judgment, and the suffering that comes from physical pain or emotional upheaval might exist separately from the body”

<sup>15</sup> No original: “*Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*”. Tradução própria.

observa que, a partir de uma observação meticulosa e com o auxílio do método experimental inaugurado por Bacon e aplicado de forma precisa por Newton na física, era possível utilizar o mesmo método para o entendimento da natureza humana. (HILLER, 2019, p. 63)

Sua proposta, portanto, é a de encontrar as bases em que o agente moral realiza as distinções morais a partir do método empirista.

O autor contrasta-se com a tradição, principalmente a racionalista, uma vez que identifica que são as paixões, ou os sentimentos, que subjazem as escolhas morais, e não a razão – embora esta última possa auxiliar de forma secundária. Isto se dá pois

apenas o sentimento tem capacidade para motivar o ser humano e somente ele é capaz de exercer influência suficiente em seu comportamento a ponto de fazer com que se tenha mais apreço pela virtude do que pelo vício. (SILVEIRA, 2010, p. 23)

Ou seja, vício e virtude são causados respectivamente por sentimentos relacionados ao agrado ou repúdio, e nisto, “a função da razão é prestar auxílio para que desse sentimento possa prover um apropriado discernimento de seu objeto” (SILVEIRA, 2010, p. 24). A moralidade é então estruturada com base na reação emocional do agente, sendo esta fonte motivadora para a ação. A razão, assim, é, em si, “inerte” quanto à motivação do agente moral. É necessário algum estímulo passional – ou seja, sensorial, *empírico*, para que possamos decidir sobre um curso de ação. Em outras palavras,

embora a razão, quando plenamente assistida e melhorada, seja suficiente para nos instruir nas tendências perniciosas ou úteis de qualidades e ações, não é sozinha suficiente para produzir quaisquer culpa ou aprovação. [...] É necessário um sentimento que aqui se demonstre para que se dê preferência às tendências úteis acima das perniciosas.<sup>16</sup> (HUME, 1777, p. 286, M App 1.3, tradução nossa)

Embora a razão possa orientar as emoções de forma a, e.g., descobrir novas informações que podem influenciar nosso raciocínio, a razão “por si só não é capaz de prevenir a ação ou resistir a paixão ao controlar a vontade” (COHON,

---

<sup>16</sup> No original: “*though reason, when fully assisted and improved, be sufficient to instruct us in the pernicious or useful tendency of qualities and actions; it is not alone sufficient to produce any moral blame or approbation. [...] It is requisite a sentiment should here display itself, in order to give a preference to the useful above the pernicious tendencies.*”

2010, tradução nossa) A razão na concepção humeana é, de tal maneira, instrumental, pois possui um papel secundário, administrativo acerca dos ímpetus gerados pela vontade. Assim,

[...] utilizando os instrumentos da lógica, podemos verificar a validade das seleções as quais nossas preferências auxiliaram em fazer [e] podemos ir além delas utilizando as estratégias de dedução e indução em proposições de linguagem prontamente disponíveis.<sup>17</sup> (DAMASIO, 1994, p. 200–201, tradução nossa)

Mas além da modulação do paradigma da ação do agente no âmbito moral, Hume possui outra diferença com a tradição filosófica que o precedeu até então: apesar de identificar a fonte do valor moral na psicologia do agente, não ousa prescrever normatividade a partir de tal. Aqui entra o problema de *ser e dever ser*,<sup>18</sup> a famosa *Lei de Hume*, conforme o autor e suas interpretações mais recentes, onde identifica-se que de uma descrição dos fenômenos morais não se seguem necessariamente normas morais. Dedicarei maior atenção a esta questão específica posteriormente; sumariamente, nos serve entender que esta modulação se traduz em trazer o nosso comportamento como fonte e campo de pesquisa da moralidade, e não o contrário como se buscava no racionalismo e outras perspectivas tradicionais prévias.

É aqui que a metáfora do anatomista e do pintor entra em cena. Ao concluir esta sua primeira obra, o autor compara a profissão de pintor, representando o moralista que trabalha com moralidade prática e imputa normatividade à sua teoria, com a profissão do anatomista, que por outro lado representa aquele que procura descrever o fenômeno moral da forma mais completa e acurada possível. O pintor é capaz de pintar formas mais “[...] atraentes para o olho e para a imaginação”<sup>19</sup> (HUME, 1739, p. 621, 3.3.6.6, tradução nossa) do que o anatomista o faria se colocado na mesma posição. Afinal, a perspectiva do anatomista é fria, até mesmo repulsiva dada que é advinda suas disseções. Todavia, para Hume, o pintor necessita de orientação do anatomista para elaborar suas formas no quadro, a tal ponto em que, sem tal,

---

<sup>17</sup> No original: “using the instruments of logic, we can check on the validity of the selections our preferences have helped make [and] can go beyond them using the strategies of deduction and induction in readily available language propositions.”

<sup>18</sup> Também conhecido como “Guilhotina de Hume” e “Is/Ought Gap” (aproximadamente, “lacuna entre ser/dever ser”).

<sup>19</sup> No original: “engaging to the eye and imagination”.

sua excelência seria impraticável (id.) De tal forma, com esta analogia, Hume pretende ilustrar que mesmo

as especulações mais abstratas acerca da natureza humana, independentemente do quão frias e entediante, tornam-se subservientes à moralidade prática; e podem render esta última ciência mais correta em seus preceitos, e mais persuasiva em suas exortações.<sup>20</sup> (HUME, 1739, p. 621, 3.3.6.6, tradução nossa)

Ao focar no papel análogo ao do anatomista no que tange temáticas da moralidade, Hume então subverte o paradigma tradicional da ética de sua capacidade de prescrever normas. Sua teoria é um empreendimento descritivo: por mais que o autor venha a reconhecer o princípio de utilidade como o critério subjacente da distinção moral, o autor não postula qualquer teoria normativa com base em suas observações. A teoria sobre a norma é descritiva, e não prescritiva.<sup>21</sup>

Posteriormente, no ápice do *Aufklärung* alemão, Immanuel Kant desenvolve sua teoria como uma resposta a estes racionalismos e empirismos anteriores. Em sua primeira obra, a *Crítica da Razão Pura*, ele reconhece os limites da razão em que há conceitos puros (*a priori*) os quais são incognoscíveis – como Deus, a imortalidade da alma e a liberdade –, portanto renegando um posicionamento metafísico forte de uma razão infalível. No entanto, apesar de reconhecer a importância do âmbito sensível (*a posteriori*), seu papel é secundário ao ato de intuição ao entendimento (viés da razão teórica), que é capaz de categorizá-lo.

Em sua segunda obra, Kant trata do uso da razão no âmbito prático, em seu viés de apetição, através do estudo da vontade, que é a inclinação de agir segundo representações. Nela, afirma “ter descoberto o princípio supremo da

---

<sup>20</sup> No original: “*the most abstract speculations concerning human nature, however cold and unentertaining, become subservient to practical morality; and may render this latter science more correct in its precepts, and more persuasive in its exhortations.*”

<sup>21</sup> Gostaria de esclarecer que esta leitura do pensamento humeano não é indisputada. Leituras normativas de sua leitura do utilitarismo moral (como “pintor”) são tão comuns quanto sua leitura descritiva da esfera moral como um todo (como “anatomista”). Todavia, considerando o paradigma central do naturalismo moral contemporâneo, não me adentrarei nesta discussão uma vez que o viés do anatomista é a leitura padrão desta corrente filosófica, e trabalhar sua contraparte fugiria do escopo deste trabalho.

razão prática”<sup>22</sup> (WILLIAMS, 2018, seç. 2, tradução nossa), nominalmente, o Imperativo Categórico (IC). Ao formulá-lo, Kant desenvolve que

a razão pura, prática de si mesma, é aqui imediatamente geradora de leis. A vontade é pensada como independente de condições empíricas e portanto, como pura vontade, tão determinada pela mera forma da lei, e este fundamento determinante é tido como a condição suprema de todas as máximas.<sup>23</sup> (KANT, 1997, p. 28, 5:31, tradução nossa)

O IC é, portanto, o princípio de todo o dever, o crivo máximo da moralidade de todas as ações. Ele existe em contraste ao Imperativo Hipotético (IH), que é um princípio instrumental da razão para determinados fins, ou seja, teleológico. O IH postula que para conseguir resultado X, devo fazer Y: há um direcionamento do uso da razão para fins específicos, instrumentalizando os meios. Já o IC surge em um contexto onde Kant desenvolve prolegômenos de uma teoria da ação na análise da razão pela vontade no viés do senso comum. Ora, uma vez que somos seres dotados de uma capacidade racional, segundo o autor, “acima dos animais,” (KANT, 1997, p. 129 5:162) temos a capacidade de representar situações em um contexto universalizável – remanescente da primeira formulação do IC como a *regra de ouro*. Assim, formula-se: “age de tal maneira que a máxima da tua vontade possa simultaneamente manter-se como o mesmo princípio no prover de uma lei universal”<sup>24</sup> (KANT, 1997, p. 28, tradução nossa). Desta maneira,

agir por respeito à lei moral, na perspectiva de Kant, é ser movido a agir por um reconhecimento que a lei moral é um padrão supremamente autoritativo que nos vincula e experienciar um tipo de sentimento, que é parecido com deslumbramento e medo, quando reconhecemos a lei moral como a fonte dos requerimentos morais.<sup>25</sup> (JOHNSON; CURETON, 2018, seç. 3, tradução nossa)

<sup>22</sup> No original: “to have discovered the supreme principle of practical reason”.

<sup>23</sup> No original: “pure reason, practical of itself, is here immediately lawgiving. The will is thought as independent of empirical conditions and hence, as a pure will, as determined by the mere form of law, and this determining ground is regarded as the supreme condition of all maxims.”

<sup>24</sup> No original: “act that the maxim of your will could always hold at the same time as a principle in a giving of universal law”.

<sup>25</sup> No original: “[t]o act out of respect for the moral law, in Kant’s view, is to be moved to act by a recognition that the moral law is a supremely authoritative standard that binds us and to experience a kind of feeling, which is akin to awe and fear, when we acknowledge the moral law as the source of moral requirements.”

Ou seja, em Kant, a lei moral é um resultado *a priori* da intuição de observações *a posteriori* dado o fato<sup>26</sup> de nossa racionalidade. Agir conforme a lei moral é agir de tal maneira de ter o próximo como um fim em si mesmo, independente das consequências da ação. Ou seja,

a razão pura é prática, ou seja, [...] a razão pura pode ser um fundamento determinante da vontade humana, logo que a consciência da lei moral não é apenas uma ideia hipotética ou problemática, mas, de fato, tem realidade na medida em que todo ser racional humano reconhece originariamente [...] a sua legitimidade e “produz”, com base nela, efeitos no mundo, quer dizer, ações morais. (CHAGAS, 2011, p. 141)

Temos nesta formulação um forte racionalismo que postula um realismo dos fatos morais, que, porém, dependem epistemicamente de uma intuição especial para serem reconhecidos pela razão. A deontologia kantiana apresenta o reconhecimento da normatividade através da identificação intuitiva de uma fundação factual para o juízo moral, ainda que a normatividade fuja do aspecto causal físico (natural *per se*) uma vez que necessita deste aspecto racional da intuição. Assim, uma vez que ações morais dependem do uso adequado da razão, “todas as ações imorais são irracionais pois elas violam o [Imperativo Categórico]”<sup>27</sup> (JOHNSON; CURETON, 2018, p. 1, tradução nossa) –, conforme Korsgaard define,

a abordagem kantiana para a filosofia moral é de tentar mostrar que a ética é baseada na razão prática: ou seja, nossos juízos éticos podem ser explicados em termos de padrões racionais que aplicam-se diretamente à conduta ou deliberação.<sup>28</sup> (1986, p. 5, tradução nossa)

Todavia, ao contrário do que alguns críticos tendem a apostar, o papel da racionalidade em Kant não é absoluto – embora, indubitavelmente, seja o foco e cerne primário de toda sua teoria. A formulação da racionalidade de Kant nos vê como seres limitados, incapazes, como dito anteriormente, de acessar certos conceitos puros *a priori* e dotados de apetições. Ou seja, a formulação do

<sup>26</sup> Há diferentes interpretações acerca da grafia original como *fakto* ou *factum* e das implicações que estes diferentes usos poderiam ter na teoria kantiana, porém não me adentrarei neste debate por fugir do escopo do atual trabalho.

<sup>27</sup> No original: “all immoral actions are irrational because they violate the [Categorical Imperative]”.

<sup>28</sup> No original: “[t]he Kantian approach to moral philosophy is to try to show that ethics is based on practical reason: that is, that our ethical judgments can be explained in terms of rational standards that apply directly to conduct or to deliberation.”

princípio da moralidade em Kant não é para anjos ou criaturas perfeitas<sup>29</sup> – “é um imperativo para seres finitos como nós, que possuem necessidades e inclinações e não são perfeitamente racionais”<sup>30</sup> (WILLIAMS, 2018, seq. 2).

Ainda assim, mesmo procurando afastar-se de um platonismo puro, Kant está muito mais próximo de Platão e dos racionalistas do que o contrário. Apesar de rechaçar um idealismo puro platônico retirando de sua normatividade a onipotência da esfera apriorística (análoga ao mundo das Formas platônicas ou as fundações dogmáticas cartesianas) em prol de uma racionalidade prática em diálogo com o mundo factual, ainda há uma sobrepujação da razão sobre outros fatores. Se a ação não é primariamente pautada de maneira racional, é necessariamente imoral. Este tipo de afirmação sem dúvida surpreenderia David Hume, a quem Kant atribui seu despertar do “sono dogmático” (KANT, 2004, p. 10). De fato, tratando-se da felicidade e das apetições “inferiores” como desejos ou paixões, o autor as elimina do cálculo moral dado seu viés inerentemente relativo: pessoas mudam de “gosto” durante sua vida, e o que provê prazer em determinado momento pode não o fazer no futuro, e vice-versa. O afastamento deste viés idiossincrático em prol da impessoalidade universalizável da razão pura através da apetição “superior”, ou a vontade do uso da razão para a ação, é o que caracteriza a moralidade para Kant, portanto (ASSUMPÇÃO, 2017, p. 108–109). Este é um ponto interessante, que pautará a introdução do nosso último capítulo.

Nesta seção, vimos que

na tradução filosófica, “razão” é frequentemente utilizada para referir-se aos aspectos da mente ativos em oposição aos passivos ou receptivos. “Razão” neste sentido é contrastado com percepção, sensação e emoção, que são pensados como formas de passividade, ou ao menos como envolvendo passividade.<sup>31</sup> (KORSGAARD, 2018, p. 297)

---

<sup>29</sup> Agradeço às aulas da profa. Dra. Flávia Carvalho pelo importante auxílio em compreender esta faceta de Kant.

<sup>30</sup> No original: “[it] is an imperative for finite beings like us, who have needs and inclinations and are not perfectly rational.”

<sup>31</sup> No original: “[i]n the philosophical tradition, “reason” is often taken to refer to the active as opposed to the passive or receptive aspects of the mind. “Reason” in this sense is contrasted with perception, sensation, and emotion, which are thought of as forms of passivity, or at least as involving passivity.”

Em prol de uma exaltação exclusiva das capacidades cognitivas, outros diversos fatores envolvidos na tomada de decisão ficaram predominantemente de lado na história da filosofia. Quaisquer outros fatores que pudessem influenciar no processo deliberativo possuem papel secundário; neste sentido, o viés

racionalista [...] desprezou o papel do corpo, da matéria e das paixões para aquisição de uma vida ética virtuosa no campo da filosofia prática ou mesmo para um conhecimento indubitável no campo epistemológico. (DE LIMA; DE OLIVEIRA, 2019, p. 5)<sup>32</sup>

Além disso, há diversos outros fatores que não tocamos acerca do que entra na definição de “agente racional” para estes autores: Aristóteles exclui escravos e mulheres, por exemplo, (KARBOWSKI, 2014) assim como Kant em seus diversos textos de antropologia que acabam por delimitar apenas o homem branco europeu como capaz da atividade racional adequada – material que supera em quantidade seus textos filosóficos. Com eles em perspectiva, “é claro que Kant não considera todos os humanos iguais. Mas o que exatamente torna não-brancos inferiores? A resposta parece ser racionalidade, ou melhor ainda, racionalidade incompleta”<sup>33</sup> (ABUNDEZ-GUERRA, 2018, p. 4, tradução nossa).

Não é por acaso que racismo, sexismo e noções de racionalidade “pura” – no sentido de neutra, isenta de influências no processo decisivo – coexistem com tamanha frequência. Trataremos do problema da ausência de possibilidade de uma neutralidade tanto da própria noção de racionalidade através de evidências científicas assim como da própria metodologia científica nos capítulos 5 e 6. Tal problema se tornará evidente principalmente ao tratarmos de viés cognitivo e percepção racial e de gênero, como também de problemas na prática científica e de definições populacionais e epistêmicas de sujeito.

Na seção seguinte, investigaremos os últimos defensores do racionalismo na contemporaneidade a fim de caracterizar de maneira sumária frente a este contexto histórico os mais recentes defensores de teorias racionalistas e como este problema é tratado hoje em dia, especialmente frente ao tópico interentemente empírico que é o do desacordo moral.

---

<sup>32</sup> Aqui os autores se referem estritamente à tradição racionalista na filosofia; todavia acredito que é uma avaliação que adequa-se ao panorama geral de como este tópico fora tratado em sua história em geral.

<sup>33</sup> No original: “*it is clear that Kant does not consider all humans equal. But what exactly makes nonwhites inferior? The answer seems to be rationality, or better yet, incomplete rationality*”

## 2.2 Racionalidade contemporânea e desacordo moral

Philosophy can at most tell us what it would be like to be rational.

– Christine Korsgaard (2018, p. 25)

Muito se passou entre Descartes e os modernos até os tempos recentes. Do racionalismo como uma posição inteiramente atrelada à um posicionamento metafísico idealista para concepções deflacionadas de uma racionalidade lógica e/ou científica, de fato o único elemento comum talvez seja a primazia de nossas faculdades de juízo acima de influências emocionais ou externas em geral. Como qualquer escola de filosofia, não há como estabelecer uma definição que não seja incompleta ou excludente – não ao menos sem excluir diversas perspectivas ou representar erroneamente por simplificação as que poderiam ser incluídas. Afinal, inúmeros são os autores que o adotam, e da mesma forma inúmeras são suas definições. Todavia, nosso objetivo não é exaurir todas as definições de racionalismo e muito menos refutá-lo como um todo. Nosso objetivo é entender o fenômeno da tomada de decisão em diálogo com as outras ciências, de maneira que se torne claro que as evidências atuais apontam para uma razão deflacionada que não existe sem a influência majoritária de outros fatores alheios a ela. Mas a metodologia científica preza a constante autocrítica e revisão frente a novos dados, e, portanto, seria antitético à proposta deste trabalho definir um “vencedor” claro – podemos, no máximo, apontar qual teoria dá melhor conta das evidências.

Pretendo, portanto, desenvolver uma definição minimamente aceitável de qual seria o papel da “razão” e da “racionalidade” na tomada de decisão para os racionalistas contemporâneos. Para tal fim, irei tratar brevemente de algumas teorias desde o neopragmatismo de Putnam e o liberalismo neokantiano de Rawls até de realistas naturalistas – já que como vimos na seção anterior nem todos se subscrevem a uma perspectiva fortemente humeana da primazia das paixões sobre a razão. Para delinear as suas posições, iremos analisar de maneira breve como tratam o problema do desacordo moral, que historicamente “provê uma ilustração gritante da maneira como a ciência em geral pode e não pode apropriadamente ser trazida para tratar de questões sobre metaética”<sup>34</sup>

---

<sup>34</sup> No original: “provides a striking illustration of the way science in general can and cannot properly be brought to bear on a question about metaethics”.

(ALFANO; LOEB, 2017, p. 31), e aproxima o debate acerca da racionalidade do juízo moral das mais variadas escolas de pensamento filosófico. Nos é interessante aqui não apenas pela discussão sobre nossas faculdades racionais, mas também pois ao tratar temas filosóficos em uma esfera puramente empírica – do desacordo moral entre diferentes agentes em diferentes contextos e culturas –, verificamos a importância do paradigma experimental para os mesmos, que é o foco deste trabalho.

Mesmo ao tratarmos do racionalismo de séculos atrás, o viés experimental mantêm-se verdadeiro pois “de acordo com o racionalismo, moralidade é baseada na razão ou racionalidade ao invés de nas emoções ou idiosincrasias culturais”<sup>35</sup> (NICHOLS, 2002, p. 285); e assim, independentemente das perspectivas dos autores, são colocações empiricamente testáveis. Em outras palavras, este tópico

possui elementos empíricos substantivos: pressupõe observações antropológicas sobre como desacordos morais procedem – ou falham em proceder. [E] avaliar as observações antropológicas requer uma investigação cuidadosa sobre a diversidade avaliativa do mundo real – em oposição a filosoficamente imaginada.<sup>36</sup> (DORIS; PLAKIAS, 2008, p. 303)

Realistas morais frequentemente realizam estes tipos de afirmações sem devida referência à dados que este problema empírico requer, além de tenderem a defender alguma formulação racionalista em algum grau. Portanto, parte desta seção irá se dedicar a ilustrar as formulações destes autores quanto ao problema do desacordo moral, mas iremos tratar deste último e como oferece um desafio real ao realismo moral ao tratar das especificidades empíricas deste tema e suas implicações para uma concepção experimentalista da moral no capítulo 4. Trataremos primeiro de abordagens realistas quanto ao juízo moral na perspectiva de Doris e Plakias (2008), e posteriormente de outras perspectivas antirrealistas como Putnam e Rawls.

---

<sup>35</sup> No original: “[a]ccording to rationalism, morality is based on reason or rationality rather than the emotions or cultural idiosyncrasies [...]”

<sup>36</sup> No original: “has substantial empirical elements: It presupposes anthropological observations about how moral disagreements proceed—or fail to proceed. [And] evaluating the anthropological observations requires careful inquiry into real-world—as opposed to philosophically imagined—evaluative diversity.”

É contraprodutivo, evidentemente, tentarmos nos engajar com escritos de séculos atrás vista a quantidade de debates, críticas e adaptações tais teorias tiveram por todo o tempo em que foram relevantes, e portanto tomarei autores mais contemporâneos a fim de sintetizar estes posicionamentos gerais em um contexto atual. Portanto, trataremos aqui de posições de autores naturalistas como Richard Boyd e Bloomfield e de não-naturalistas como Michael Smith e Shafer-Landau, a fim de entendermos qual o posicionamento “padrão” dos realistas frente ao desacordo moral. Segundo Doris e Plakias (2008) há dois tipos de respostas-padrão a este problema: dos *convergentistas*, que identificam o problema do desacordo moral como um problema superficial superável, e os *divergentistas*<sup>37</sup>, que descartam o desacordo moral como um pseudo-problema dentro de seu realismo. Independente de suas diferenças, ambos posicionamentos em geral possuem uma fraqueza: se evidências explícitas de desacordo moral profundo – ou seja, não apenas que tal desacordo exista, mas que não seja meramente superficial – existirem, traria um grande problema para estas teorias<sup>38</sup>.

Primeiramente, os divergentistas, representados aqui por Bloomfield e Shafer-Landau, afirmam que o desacordo moral é, sumariamente, o desacordo de diferentes opiniões, e estas, por sua vez, não constituem a verdade moral. Segundo eles, não é a mera opinião de diversos agentes que irá definir o que é verdade ou não. Bloomfield associa seu realismo moral com um realismo sobre a saúde. Negar um realismo sobre a moralidade é negar a diferença entre a diferença entre vida e morte. Assim, o discurso moral se relacionaria ao discurso dos médicos em uma medicina baseada em evidências, onde o prognóstico para uma boa saúde pode sempre ser concordado mediante análise e tempo. Podem haver, então, diferentes concepções do que é saudável, mas é uma questão de tempo até haver concordância entre os profissionais.

Esta perspectiva é problemática. Embora isto possa ser verdade em casos simples, como gripes, infecções e outros casos comuns, esta questão taxonômica torna-se terreno controverso quando analisamos o que pode ou não ser considerado doença – especialmente como definimos doenças na área de

---

<sup>37</sup> Tradução própria dos originais *convergentism* e *divergentism*, respectivamente.

<sup>38</sup> Desacordo moral *fundamental*, tese defendida principalmente por Loeb e Tolhurst (DORIS; PLAKIAS, 2008, p. 305)

saúde mental, onde historicamente houve a institucionalização forçada desproporcionalmente de minorias a fim de controle social – e nos diversos casos de opiniões divergentes entre profissionais sobre algum prognóstico. Aqui, no máximo, vemos apoio a um “realismo de retalhos”<sup>39</sup>: como na moralidade, “algumas áreas sobre o discurso da saúde demandam, e outras resistem, um tratamento realista”<sup>40</sup> (DORIS; PLAKIAS, 2008, p. 309, tradução nossa).

Schafer-Landau, por sua vez, rejeita a semelhança do discurso moral com o da ciência ou da medicina e propõe um realismo retrógrado em vias racionalistas semelhantes às do jusnaturalismo. Segundo ele, através de um intuicionismo racional percebemos que o desacordo moral é impossível, e, da mesma forma que Bloomfield, afirma que com o tempo o fato que verdades morais o são de forma a priori torna-se óbvio. O desacordo moral é, para Schafer-Landau, um erro momentâneo por falhas perceptivas ou influências de crenças falsas, ambos solucionáveis com o bom uso da razão. Embora o autor negue um convergentismo, Doris e Plakias afirmam que seria sua única saída, pois seu

ceticismo sobre a possibilidade de desacordo fundamental parece a nós um reconhecimento tácito de nosso ponto: se o desacordo fundamental é uma experiência central da experiência moral, isto problematiza o realismo. De fato, Shafer-Landau oferece conjecturas de convergência; como um realista, suspeitamos que ele precisa fazê-lo.<sup>41</sup> (DORIS; PLAKIAS, 2008, p. 310)

Quanto aos convergentistas, iniciemos com Boyd. O autor defende um realismo científico onde a evidência empírica serve como base não-redutível e heurística para uma teoria moral consequencialista, a partir da regulação constante de crenças enquanto um fenômeno social em um *equilíbrio reflexivo* coerentista entre a teoria e sua metodologia (cf. DALL’AGNOL, 2005, p. 181–190; LENMAN, 2013, p. 6). Dall’Agnol aponta que a perspectiva de Boyd concebe termos éticos como *designadores rígidos* (no sentido de Kripke) de que há coisas relativas às necessidades naturais básicas que constituem uma

---

<sup>39</sup> “*Patchy realism*” no original.

<sup>40</sup> No original: “some areas of health discourse demand, and others resist, a realist treatment”

<sup>41</sup> No original: “*skepticism about the possibility of fundamental disagreement seems to us a tacit acknowledgement of our point: If fundamental disagreement is a central feature of moral experience, this problematizes realism. Indeed, Shafer-Landau offers convergence conjectures; as a realist, we suspect he must.*”

definição natural do fato moral, ainda que não o identificando como apenas uma única propriedade natural.

Assim, um entendimento realista da moralidade implica em um entendimento realista da ciência em prol da concepção de um *domínio dos fatos* desassociado da cultura. Boyd reconhece, todavia, diferenças entre questões científicas e morais: tópicos da ciência possuem mais facilidade de concordância entre cientistas e são normalmente temporários, enquanto na moralidade temos problemas virtualmente permanentes e cuja concordância não é fácil – se mesmo possível. Ainda assim, deveríamos trabalhar em prol de alcançar este domínio dos fatos para além das variações interculturais – que ele denomina de *distorções*. O realismo de Michael Smith, embora não-naturalista, segue a mesma proposta de Boyd com um misto de Shafer-Landau: dada a evidência empírica de que podemos entrar em acordo em diversos tópicos, temos motivos o suficiente que alcançaremos a convergência de opiniões dada uma situação ideal de racionalidade plena.

Assim, ao tratar sobre o descordo moral, Boyd, Brink, Sturgeon e Smith afirmam que este o é “racionalmente resolúveis quando disputantes ocupam condições otimizadas”<sup>42</sup> (DORIS; PLAKIAS, 2008, p. 304, tradução nossa). Outros autores que se encaixam nesta definição são Richard Hare, Hilary Putnam e John Rawls. Hare, embora acredite nunca ter alcançado sua aspiração última de defender a racionalidade da moral (HARE, 2002, p. 269), criticava os descritivistas analíticos pois acreditava que o reconhecimento de fatos morais diferentes levaria invariavelmente ao relativismo moral – algo que segundo ele racionalistas seriam incapazes de aceitar. Mesmo sendo um racionalista “brando”, em que adota argumentos emotivistas e principialistas (“imperativistas”, como ele se refere) (HARE, 2002, p. 288–289), seu objetivo era no foco procedural para uma decisão racional na moral, sem comprometer-se necessariamente com fatos morais, pois de acordo com ele

a objetividade de enunciados morais não consiste neles serem factuais ou descritivos. Consiste ao invés neles serem tais que todos os pensadores racionais irão assentir a eles; e pensadores racionais

---

<sup>42</sup> No original: “*rationaly resolvable when disputants occupy optimal conditions*”.

podem assentir a prescrições assim como a enunciados factuais.<sup>43</sup> (HARE, 2002, p. 304, tradução nossa)

Já Putnam, embora não seja um racionalista no sentido forte, assemelha-se neste quesito em sua concepção da ética como “florescimento cognitivo humano”, onde a racionalidade de comunidades epistêmicas permitiria conciliar amplo desacordo moral. Isto se dá pois há uma valoração inerente às nossas avaliações factuais acerca dos requisitos destas comunidades para o entendimento do “mundo real”, que “depende de nossos valores (e, novamente, vice-versa)”<sup>44</sup> (PUTNAM, 1981, p. 135, tradução nossa) – coerentismo herdado de Quine, um dos mais exponentes naturalistas no contexto epistemológico, e desenvolvido no viés da moralidade. Esta concepção estaria pautada em sua perspectiva da “morte da ontologia”, em que inspira-se na mereologia<sup>45</sup> e nos jogos-de-linguagem wittgensteinianos a fim de estabelecer uma teoria moral “maleável” o suficiente para abarcar posicionamentos éticos diametralmente opostos em oposição à uma teoria “rígida”, incapaz de dar conta do fenômeno dinâmico do desacordo moral (PUTNAM, 2005, p. 28); ou seja, não implicaria em um realismo metafísico. Não por acaso, é uma posição fortemente influenciada por Aristóteles e Kant (2005, p. 23–26). Assim, para Putnam, ao estarmos

desprovidos da antiga ideia realista de verdade como ‘correspondência’ e da ideia positivista de justificação como fixada por critérios ‘públicos’, ficamos com a necessidade de ver nossa busca por concepções de racionalidade como uma atividade humana intencional, que, como qualquer atividade que surge sobre o hábito e o mero seguir de inclinação ou obsessão, é guiada por nossa ideia de bom.<sup>46</sup> (PUTNAM, 1981, p. 136–137)

<sup>43</sup> No original: “*the objectivity of moral statements does not consist in their being factual or descriptive. It consists rather in their being such that all rational thinkers will assent to them; and rational thinkers can assent to prescriptions as well as to factual statements.*”

<sup>44</sup> No original: “*depends upon our values (and, again, vice versa)*”.

<sup>45</sup> Modelo lógico criado por Stanisław Lezniewski, influenciado por Husserl, que se utiliza da teoria dos conjuntos para denotar objetos: a união de quaisquer duas entidades (“coisas” quaisquer) denotariam um objeto. Um sapato e a Torre Eiffel poderiam ser um objeto: ambos compartilham lados esquerdos, parte superior e inferior, etc. Mas apesar da inspiração por Husserl, Lezniewski afasta-se de uma necessidade substantiva para a existência de objetos – de fato, a existência dos objetos “em si” é irrelevante para seu status ontológico: o único critério que importa para tal é a convenção escolhida para a análise de entidades. Uma “coisa” pode ser definida como a soma de quaisquer outras duas “coisas”. (cf. PUTNAM, 2005, p. 36)

<sup>46</sup> No original: “[b]ereft of the old realist idea of truth as ‘correspondence’ and of the positivist idea of justification as fixed by public ‘criteria’, we are left with the necessity of seeing our search for better conceptions of rationality as an intentional human activity, which, like every activity that rises above habit and the mere following of inclination or obsession, is guided by our idea of the good.”

É interessante que, para estabelecer sua teoria moral frente ao problema do desacordo ético, Putnam se utiliza do famoso experimento de pensamento do “cérebro em uma cuba”, que seria a radicalização do solipsismo existencialista característico da filosofia primeira (maneira que ele e Quine referiam-se às tradições filosóficas fundacionalistas, como o próprio racionalismo cartesiano) em que se acreditaria que um “cientista maligno” (semelhante ao “gênio maligno” de Descartes) teria nos sujeitado a uma operação e nosso cérebro posto em uma cuba com os devidos cuidados para seu funcionamento contínuo. Através de um supercomputador avançadíssimo, controlar-se-ia nossa percepção em uma realidade virtual a ponto de acreditarmos estarmos em um corpo em uma realidade (PUTNAM, 1981, p. 6-7).

Sua refutação inicial a este posicionamento foge do escopo do presente trabalho, porém, a aplicação do mesmo na questão do desacordo moral torna-se vital para nossa compreensão de sua proposta de superação (PUTNAM, 1981, p. 130-135). A diferença com sua formulação inicial consiste em que, ao invés de apenas uma pessoa acreditar nesta possibilidade como ocorrendo com ela naquele dado momento, é a população inteira da Austrália sob um “guru” que supostamente possui “acesso” ao conhecimento das intenções do cientista malévolo. Todavia, apesar disto, não possuímos nenhum desacordo científico nem ético com eles, dotados de uma concepção de mundo praticamente igual a nossa – exceto pelo fato que veem tal como uma alucinação gerada pelo supercomputador. Mesmo com todas as semelhanças e acordos possíveis, há um desacordo gritante em que enxergamos certa *loucura* (PUTNAM, 1981, p. 132) na perspectiva destes australianos. Loucura de fato, pois suas mentes teriam algum tipo de doença, algum transtorno para poderem pensar de tal maneira.

E como pensaríamos em tentar dialogar com eles acerca desta noção errônea? Primeiramente, há a possibilidade de tentar nos aproximar com uma ideia de coerência, mas que eles julgam não ser importante. Depois, podemos oferecer o conceito da Navalha de Ockham, evitando a multiplicação desnecessária de entidades abstratas – porém, apesar de sua simplicidade funcional, dificilmente seria aceita por um povo que se sente feliz em criar cada vez mais objetos fora da cuba sem propósito algum para nossas experiências.

Em terceiro, podemos tentar convencê-los do aspecto holístico do corpus íntegro, holístico da ciência, ao invés de aspectos separados.

É interessante notar aqui que Putnam trata com certa frequência perspectivas dissidentes às suas como doenças mentais, tanto metaforicamente como vimos ao tratar de populações hipotéticas que defenderiam o paradigma do “cérebro na cuba”, ou também um “super-benthamismo” (PUTNAM, 1981, p. 139-141) e explicitamente ao fim da última conferência da primeira parte de *Ethics without ontology* (PUTNAM, 2005, p. 82-85) ao se referir às teorias de Paul Churchland, Bernard Williams e Simon Blackburn, como nota também Mark Timmons (TIMMONS, 1991, p. 371). Se esta é sua opinião pessoal ou está tentando demonstrar ao leitor a inseparabilidade entre fato e valor em sua asserção é uma questão em aberto, mas novamente há uma correlação evidente na tradição contemporânea entre “sanidade” e “razão” como mencionamos anteriormente.

Se algum destes passos poderia funcionar, continuará um mistério, mas o fato é que todas as alternativas que ofereceríamos consistem em como já partimos desta noção de florescimento cognitivo humano, ou seja, que nossa perspectiva de ciência é um sistema de vários valores que como já vimos não depende de um realismo. A ausência de conexão com uma ontologia e o subjugar dos enunciados éticos aos mesmos escrutínios de enunciados científicos permitem uma objetividade dos juízos éticos, mesmo que não se adeque a uma imagem de “fatos naturais” (PUTNAM, 2005, p. 77-78).

O fato da perspectiva de Putnam aceitar posicionamentos destoantes, e, por muitas vezes, incompatíveis de juízos éticos não implica que todos tais juízos possuam o mesmo valor de verdade, e também não engendra uma subjetividade pura. Como vimos no exemplo do culto do cérebro na cuba, existem padrões, critérios para a aceitabilidade racional aplicados pelas comunidades epistêmicas, e o ato de procurar melhores concepções de racionalidade pressupõe uma concepção de “bom” – já que critérios cognitivos arbitrários iriam render os fatos como inúteis. Assim, o domínio de conceitos e descritores para caracterizar pessoas e o cenário condizem com tais critérios de aceitabilidade racional, sujeitos à adequação em equilíbrio reflexivo em prol de descritores “bons” ou “ruins” (PUTNAM, 1981, p. 135-139).

A importância da racionalidade é inerente à abordagem de Putnam, pois evita um relativismo completo dos juízos éticos uma vez que se cede uma maior objetividade de veracidade para aquelas concepções éticas que melhor dão conta do mundo em prol do florescimento. Da mesma forma, evita um absolutismo anti-histórico pois prevê a evolução conceitual e valorativa através da reavaliação e percepção de metas, como concorda Bernard Williams (NAVIA, 2016, p. 138–140). Os contínuos e inevitáveis padrões de aceitabilidade racional das comunidades epistêmicas pressupõe o distanciamento de um autoritarismo, pelas noções de razoabilidade e florescimento (PUTNAM, 1981, p. 147-149), assim como os casos extremos do nazismo, uma vez que a valoratividade cognitiva permite que rechacemos ideias factualmente errôneas e moralmente danosas – “anti-eudaimônicas” (NAVIA, 2016, p. 140-141).

Apesar de seu posicionamento naturalista forte, a perspectiva ética de Putnam remete mais ao jusnaturalismo, que como vimos defende um intuicionismo moral fortemente associado às faculdades racionais, um racionalismo extremo análogo à fé cristã, tema bastante presente também nos trabalhos de seus autores. Putnam já defende versões de intuicionismo nos exemplos do cérebro na cuba e principalmente no argumento da “terra gêmea”, em que procura estabelecer juízos de identidade sintéticos (COPP, 2000), e assim também o é na ética, uma vez que ao sermos dotados de razão somos capazes de intuir o que é bom e mau a fim de nosso próprio florescimento individual – terminologia também comum na área. Este ponto não é por acaso: “*eudaimonia*”, o termo que Aristóteles delineia como o objetivo da ética, é tradicionalmente traduzida como “felicidade”, quando não apenas mantida no original. Os tais jusnaturalistas – que também frequentemente remetem a Aristóteles em suas perspectivas éticas –, por sua vez, tendem a traduzir como “florescimento” (“*flourishing*”, em autores da língua inglesa). Assim, apesar de cruzar longas distâncias para uma ética sem metafísica, Putnam inadvertidamente cai num intuicionismo jusnaturalista cuja única diferença jaz no aspecto social da racionalidade.

John Rawls, por outro lado, defende que é possível conciliar uma ideia de desacordo moral fundamental através da racionalidade na esfera pública, ou o que ele chama de “razoabilidade”. Segundo ele, juízos morais fazem parte da esfera subjetiva, uma vez que cada indivíduo possui uma concepção do que é

“bom”. Rawls então parte para uma separação substantiva entre a esfera da ética, ou do que é subjetivamente bom ou ruim, da política, daquilo que é publicamente – e portanto, objetivamente – justo ou injusto (RAWLS, 2002, p. 50–52).

O embasamento ético de sua teoria estaria em um coerentismo (“equilíbrio reflexivo”) entre os juízos ponderados dos indivíduos (IGANSI, 2014d, p. 8–9), ou seja, aqueles juízos que se tem maior “certeza” pessoalmente do que outros “mais incertos” e que portanto não hesitaria-se em torná-los públicos (RAWLS, 2002, p. 49–50), e os valores políticos alcançados com seu processo contratualista, conciliando assim uma teoria política com o desacordo moral que ele denomina de “pluralismo razoável”. E como é de praxe neste tipo de teoria, preza-se por uma isenção do agente de tudo aquilo que se é considerado pessoal tendo como fim um julgamento imparcial, impessoal. Isto se dá pelo uso do processo da posição original sob o véu da ignorância por indivíduos racionais e razoáveis (capazes e dispostos a interagir com a esfera pública), uma espécie de experimento mental contratualista hipotético o qual então é capaz de gerar os princípios de justiça mais razoáveis o possível (RAWLS, 2002, cap. 3, 4).

A “posição original” é análoga a um “estado de natureza”, onde o indivíduo localiza-se anteriormente às contingências (sorte natural, circunstâncias sociais e outros acidentes imprevisíveis) das quais a sua vida na sociedade pode gerar. Nela, usa-se o “véu da ignorância”, simbolizando que o indivíduo não saberá qual a sorte na qual estará inserido quando “entrar” na sociedade: qual etnia, religião, classe econômica etc. Assim, julga-se quais valores institucionais devem ser estimulados a fim de gerarmos princípios abrangentes e não-axiomáticos para um governo devidamente democrático.

No entanto, assim como há um deflacionamento do que é “justiça” ao reduzi-la a uma “equidade” (*fairness*, no original), similar à uma teoria dos jogos, onde subtrai-se a perspectiva metafísica platônica para o *fair*, expressão semanticamente cognata a “jogo limpo” na língua inglesa, amplamente utilizada em jogos e outros empreendimentos do cotidiano, há também o da razão, ao ser reduzida à razoabilidade. Rawls utiliza a concepção padrão de racionalidade como presente nas teorias sociais, ou seja,

é entendida na maneira familiar da economia. Assim, as partes são racionais em que podem classificar seus fins últimos consistentemente; deliberam guiadas por tais princípios de tal maneira a: adotar os meios mais efetivos para seus fins; selecionar a alternativa mais provável para avançar tais fins; organizar atividades para que, *ceteris paribus*, mais ao invés de menos daqueles fins sejam cumpridos.<sup>47</sup> (RAWLS, 2001, p. 87, tradução nossa)

Inclusive, o autor remete a Herbert Simon em uma nota de rodapé (RAWLS, 2005a, p. 143), que, conforme veremos no próximo capítulo, critica veementemente os modelos clássicos de racionalidade que pressupõe agentes no completo controle de suas faculdades deliberativas sem levar em consideração todas as possíveis limitações contextuais que estas podem sofrer. Assim, ao definir sua concepção de pessoa, Rawls afirma que

todo indivíduo é racional, razoável, igual e livre. Isto significa, respectivamente, que ele possui uma concepção de bem própria, a capacidade de adequar-se à publicidade de sua sociedade, as mesmas capacidades intelectuais e morais de todos os outros indivíduos e os mesmos direitos e deveres, e a liberdade de escolher e seguir o que quiser. (IGANSI, 2014d, p. 10)

Desta maneira, o indivíduo político necessita possuir pleno domínio de dois “poderes morais” para participarem da posição original: a racionalidade que, para Rawls, seria intrinsecamente associada à capacidade de identificar o que é moralmente bom em um nível individual, e a razoabilidade, que seria a capacidade de pensar o que é justo em um âmbito societal, público. Todavia, o autor intencionalmente deixa aberta a definição de ambos os poderes, apesar destas e outras breves tentativas definicionais, a fim de que ambos os termos subentendam-se através do uso (RAWLS, 2001, p. 82).

Rawls frequentemente se utiliza de terminologia como “senso de justiça”, e refere-se nominalmente à psicologia moral para justificar sua teoria, todavia como Quine nunca trabalha nem cita nenhuma pesquisa, dados ou material em geral relacionado a estes tópicos. E de maneira semelhante à Putnam, associa a ausência de razão à indivíduos com transtornos mentais, que não devem ser inseridos no processo supracitado para a escolha de princípios que vão reger a

---

<sup>47</sup> No original: “*is understood in the way familiar from economics. Thus the parties are rational in that they can rank their final ends consistently; they deliberate guided by such principles as: to adopt the most effective means to one’s ends; to select the alternative most likely to advance those ends; to schedule activities so that, ceteris paribus, more rather than less of those ends can be fulfilled.*”

sociedade num âmbito geral – ou seja, que inclui eles mesmos (RAWLS, 2005b, p. 18–22, 183–186; PUTNAM et al., 2019, seç. 4, 4.1). Assim, deixa para o nível legislativo posterior para dar conta de como manter seu bem-estar social, perspectiva amplamente criticada por sua falta de conhecimento de tópicos de saúde mental e psicologia de fato (PUTNAM et al., 2019, seç. 4.1).

Rawls tenta amenizar seu posicionamento posteriormente ao considerar uma representatividade futura, ou seja, de imaginarmos a possibilidade de eventualmente estarmos subjugados à uma situação de debilidade física ou psicológica para podermos abarcar mesmo aqueles que são incapazes de participar da posição original (RAWLS, 2001, p. 86–87). Todavia, ainda cabe o problema também reconhecido pelo feminismo de pessoas que não fazem parte de minorias subrepresentadas serem responsáveis por gerar as leis que irão regê-las (PUTNAM et al., 2019, seç. 4.1).

A razoabilidade contratualista rawlseana procura então tentar fugir de uma discussão do desacordo moral a partir de uma suposta “universalizabilidade localizada” de princípios políticos, pressupondo que nossas capacidades racionais seriam capazes de uma sublimação última à imparcialidade e compreensão plena das vivências de todos os grupos abarcados dentro de uma democracia plural. Como veremos indiretamente posteriormente, apesar de ser uma tentativa admirável de dissociar um governo democrático de alguma perspectiva única e excludente, cada parte destes pressupostos acaba por cair por terra ao termos uma compreensão mais aprofundada de nossas capacidades cognitivas e entendemos nossa relação com o discurso deontológico.

Concluo esta seção concordando então com Christine Korsgaard, embora por outros motivos, que “a extensão pela qual as pessoas são de fato movidas por considerações racionais [...] é além da competência da filosofia”<sup>48</sup> (KORSGAARD, 1986, p. 25). Tendo uma limitação prática na metodologia filosófica de ausência de diálogo com dados empíricos acerca do comportamento humano, dedicarei o próximo capítulo para estabelecer um entendimento básico do processo de tomada de decisão em âmbitos relevantes para tal discussão. Mas apesar dos limites tradicionais da filosofia, estes tópicos tratam de temas intrinsecamente filosóficos, indissociáveis de qualquer tentativa

---

<sup>48</sup> No original: “[t]he extent to which people are actually moved by rational considerations [...] is beyond the purview of philosophy”

de análise da agência humana. A partir disto, voltaremos no capítulo posterior a tratar, como Korsgaard propôs, das implicações destas evidências para o discurso filosófico acerca da racionalidade.

### 3. METODOLOGIAS CIENTÍFICAS NA FILOSOFIA

All behavior is shaped by the interaction of genes with  
the environment.

(KANDEL; SCHWARTZ; JESSELL, 1996, p. 555)

Tratamos, no capítulo anterior, de um panorama da racionalidade na tradição filosófica e como é entendida em termos contemporâneos, uma vez que perspectivas que prezam nossas faculdades racionais como confiáveis e capazes de funcionar de maneira otimizada além ou apesar de quaisquer influências emotivas ou externas em geral ainda se mantêm relevantes contemporaneamente. Neste capítulo, trataremos das bases filosóficas que sustentam este trabalho desde um panorama do naturalismo filosófico como paradigma de pesquisa até a prática de metodologia empírica da filosofia experimental, concepção-cerne da proposta do atual trabalho.

Primeiramente, iremos oferecer um breve panorama metaético da tradição naturalista na filosofia, com algumas definições gerais de autores e problemas que tratam. Seguiremos enfim para esclarecer o âmbito da filosofia experimental, que irá reger as discussões posteriores deste trabalho e justificar a utilização direta de ferramentas conceituais e evidências empíricas de outras áreas do conhecimento além da filosofia. A introdução do naturalismo na pesquisa da moral será meramente instrumental para apresentar a estrutura conceitual do resto do trabalho, uma vez que não abraçaremos nenhuma teoria em si, focando nas posições gerais sobre a possibilidade de interação entre filosofia e ciência e da concepção ontológica do mundo necessárias para o uso da metodologia experimental em temas filosóficos. Estas informações, portanto, servirão principalmente para entendermos o contexto da filosofia experimental, que se trata predominantemente de uma metodologia de pesquisa em filosofia sem conexão necessária com nenhuma teoria específica de naturalismo à parte destes paradigmas gerais.

#### 3.1 Naturalismo filosófico<sup>49</sup>

---

<sup>49</sup> Boa parte desta seção já tratei de maneira mais desenvolvida em outros trabalhos (IGANSI, 2014a, 2014b, 2015, 2021). Aqui segue uma versão resumida e adaptada para o paradigma em questão, além de expansão de detalhes específicos que deixei de lado nestes outros trabalhos. Assim, para se aprofundar em como exponho estes autores ou o naturalismo filosófico em geral, cf. estes trabalhos.

O movimento naturalista na filosofia conforme contemporaneamente entendido teve seu início na metade do século XX. Como dito acerca do racionalismo, é impraticável um cerceamento adequado de uma definição infalível de naturalismo. Todavia, já tratei bastante quanto a este problema definicional em trabalhos anteriores (IGANSI, 2014b, 2015), então sintetizo que o viés predominante de naturalismo que tomarei é de uma tradição pós-humana de descrição do fenômeno moral ausente de quaisquer propriedades sobrenaturais. Ou seja, é

um paradigma que procura eliminar qualquer aspecto sobrenatural do discurso filosófico acerca do mundo, e mais especificamente, da moral: o que é bom, justo, certo, etc. nada mais são do que fenômenos naturais específicos de acordo com a fisiologia e a natureza evolutiva humana e o ambiente físico e cultural em que se está situado – obviamente em um contexto empírico. Assim, através das múltiplas veias das ciências empíricas, procura-se um entendimento mais informado, baseado em evidências, acerca destes fenômenos usualmente insulados na esfera de investigação filosófica. (IGANSI, 2014b)

Há nesta perspectiva uma posição realista constituída de (a), uma defesa ontológica dos fatos morais como fatos naturais e (b), epistemológica, em que podemos conhecê-los como qualquer outro fato natural – a *definição padrão* de naturalismo moral conforme Sturgeon (2007, p. 92–93) e Copp (2003, p. 179) sugerem. Reitero que a abrangência desta definição não é sem controvérsia; todavia, tal definição é suficiente para deixar claro o ponto de partida neste trabalho.

As raízes do movimento evidentemente jazem no nascimento da filosofia analítica e principalmente nos debates do Círculo de Viena no início do século XX, os momentos mais férteis de debates entre ciência e filosofia que formou as primeiras estruturas para interações entre temas filosóficos e metodologias científicas da era contemporânea. Todavia, o naturalismo surge como uma resposta à falha do projeto analítico de uma redução do mundo à lógica e linguagem, e portanto destoa-se no que busca uma análise descritiva e cientificamente holística<sup>50</sup> dos fenômenos os quais procura tratar. É importante ressaltar que na literatura há uma diferença entre naturalismo ontológico, que

---

<sup>50</sup> No sentido que uma maneira integrada dos sistemas científicos e filosóficos tendem a superar o impacto de ambas as áreas separadamente.

trata todas as entidades possíveis do mundo como entidades naturais analisáveis pelo método científico das ciências naturais, e o metodológico, que vê a prática científica indissociável da prática filosófica, diferenciando-se apenas nos objetos de pesquisa (cf. PAPINEAU, 2021). Neste trabalho, adotaremos ambas as perspectivas uma vez que adotaremos uma compreensão naturalista da ontologia possível do mundo e nos utilizaremos de uma metodologia que procura alinhar a pesquisa de âmbito científico com a análise de problemas filosóficos, a saber, a da filosofia experimental, que analisaremos posteriormente nesta seção.

Embora o viés naturalista possa ser rastreado até Aristóteles e outros da antiguidade que procuravam focar em métodos empíricos, ou ao menos dentro de um modelo científico da época, para investigar o mundo e problemas filosóficos, da maneira como o compreendemos hoje surge aproximadamente com o pragmatismo no fim do século XIX. Charles Peirce e William James destoavam-se filosoficamente de seus pares por defenderem uma perspectiva empiricista extrema, procurando excluir quaisquer fragmentos de metafísica a fim de definir o mundo exclusivamente em fatos analisáveis pelas ciências naturais. Peirce, químico de formação e polímata de prática, defendia uma concepção monística ao estilo de Espinoza, todavia oposta ao seu racionalismo metafísico em que tudo na existência é fisicamente natural e pode ser analisado pelo método científico. Palavras e ideias seriam símbolos cujo significado se daria pela prática, e a realidade poderia ser analisada por três perspectivas abrangentes de uma compreensão evolutiva: um idealismo hegeliano pragmático – ou seja, sem depender de um espírito metafísico – para a evolução cultural, o evolucionismo geológico de Lyell para estruturas físicas e a teoria da seleção natural de Darwin como explanação para o desenvolvimento biológico (BURCH; PARKER, 2022).

James também se inseria nesta perspectiva de empirismo extremo, porém com forte foco na psicologia em detrimento à visão mais logicamente-orientada de Peirce – o que causou este último a alterar a maneira que se referia à sua versão de pragmatismo de “pragmaticismo”. Sua perspectiva focava em entender conceitos epistemológicos como crença e conceitos a partir de sua efetividade prática, ao invés depender de análises abstratas como predominantemente na psicologia de seu tempo.

Mas o primeiro a utilizar a nomenclatura de fato fora John Dewey, que tomou como foco a vida prática de indivíduos para a análise filosófica. Partindo da posição de James com este foco em um contexto fortemente empirista, Dewey desenvolve uma teoria epistemológica coerentista onde a experiência do indivíduo informa e regula suas ações e visão de mundo, ao passo que estas da mesma maneira acabam influenciando suas experiências de maneira progressivamente cíclica. A perspectiva Hobbesiana clássica do indivíduo racionalista e egoísta educado à socialização pelo Estado é quebrada para dar lugar à uma interpretação naturalizada onde o conhecimento, a ética, estética e aprendizado como um todo surgem de maneira social, situando o indivíduo humano novamente como um animal social dentro de um viés filosófico cientificamente informado (HILDEBRAND, 2021).

Isto é importante para nós pois estabelece uma característica definitiva do naturalismo, ainda que não predominante: não apenas o preceito básico de trabalhar teorias filosóficas dentro de um arcabouço adequadamente científico, mas também, por conseguinte entender que nossas capacidades e comportamentos estão necessariamente inseridos dentro de um contexto social, cultural e político. Embora nosso foco nestre trabalho seja razoavelmente específico, de entender o processo de tomada de decisão na moralidade, é invariável que hajam elementos sociais que irão pesar sobre como estes se desenvolvem e são influenciados por determinados fatores externos.

Outro autor importante para entendermos as bases deste trabalho fora um dos primeiros proponentes de maior relevância do naturalismo como posicionamento filosófico fora na metade do século XX. Willard V. O. Quine em sua “Epistemologia Naturalizada” (1969) procurava quebrar com um dos paradigmas predominantes em prespectivas racionalistas, o fundacionalismo – que ele chamava de “filosofia primeira”. Podemos rastrear esta perspectiva especificamente ao *cogito* cartesiano (KIM, 1988, p. 381–383), e como Paul A. Gregory (2008) coloca, é uma *perspectiva euclidiana* de justificar conclusões e premissas a partir de axiomas fundacionais de maneira cíclica. Quine procura superar esta relação justificacional (lógica), baseada em uma epistemologia normativa, com uma relação causal-nomológica, baseada em uma psicologia cognitiva – ou seja, com um viés descritivo do processo de aquisição de crenças (KIM, 1988, p. 388).

Isto leva à uma inversão da perspectiva clássica de conhecimento: uma vez que nossas representações (*output*) são *torrencialmente* mais complexas que nossa experiência do mundo (*input*), que é limitada, (QUINE, 1969, p. 82–83; KIM, 1988, p. 389–390) temos uma relação horizontal entre nossa teoria de mundo *a priori* (lógico-matemática), as teorias científicas, as leis da natureza e nossa experiência empírica. Há, portanto, um equilíbrio reflexivo entre teoria e experiência, uma via de duas mãos entre elas onde ambos se adequam conforme o necessário. Portanto

há então uma contenção recíproca, embora contenção em diferentes sentidos: epistemologia na ciência natural e ciência natural na epistemologia. [...] Estamos atrás de um entendimento da ciência como uma instituição ou processo no mundo, e não temos a intenção que este entendimento seja melhor que a ciência da qual é seu objeto.<sup>51</sup> (QUINE, 1969, p. 83–84, tradução nossa)

Assim, nossa interação com o mundo não parte de uma *tabula rasa* que constrói uma concepção de realidade a partir de diversas experiências específicas; mas sim, nossa cultura, comunidade epistêmica e constituição fisiológico formam já um arcabouço cognitivo o qual nos permite experienciar o mundo em um espectro amplo porém limitado, o que por sua vez nos convida a ajustar nossa teoria de acordo com estes estímulos, e assim *ad infinitum* – algo não muito distante da formulação pragmática de Dewey. Estamos inseridos em um contexto específico desde o nascimento e não há como tratar dele fora dele mesmo. Em um viés semelhante à Wittgenstein, Quine aponta que

a tarefa do filósofo difere da dos outros, então, em detalhe; mas não em tal maneira drástica como aqueles que supõe imaginar para o filósofo um ponto de vantagem fora do esquema conceitual que ele pretende encarregar-se. Não há tal exílio cósmico.<sup>52</sup> (QUINE, 2013, p. 254, tradução nossa)

Em contraste com Quine, que apesar de prezar pela investigação empírica jamais trabalhou diretamente com dados ou investigação deste teor, J.

---

<sup>51</sup> No original: “[t]here is thus reciprocal containment, though containment in different senses: epistemology in natural science and natural science in epistemology. [...] We are after an understanding of science as an institution or process in the world, and we do not intend that understanding to be any better than the science which is its object.”

<sup>52</sup> No original: “[t]he philosopher’s task differs from the others’, then, in detail; but in no such drastic way as those suppose who imagine for the philosopher a vantage point outside the conceptual scheme that he takes in charge. There is no such cosmic exile.”

L. Mackie foi o primeiro a se utilizar de alguma evidência empírica para trabalhar um realismo frente ao fato do desacordo moral, problema cuja relevância vimos brevemente no capítulo anterior. Seu naturalismo trata da estranheza socialmente aceita de que a prescritividade de juízos morais está no mundo natural. Ou seja, as vias pelas quais alcançamos o critério social de motivação e normatividade se dá por uma intuição “especial” que revelaria o fator elusivo “moral” à propriedades naturais, e assim, temos seu argumento da estranheza<sup>53</sup> em que

conceitos morais pretendem designar propriedades que são ultimamente estranhas. Temos razão alguma para pensar que tais propriedades existem. Assim, conceitos morais são na verdade vácuos. Juízos éticos são falsos ou sem sentido pois atribuem propriedades que são mais reais do que fadas ou flogisto.<sup>54</sup> (PRINZ, 2007, p. 88, tradução nossa)

O conceito de moralidade é então como um constructo misterioso o qual atribuímos ao mundo natural, e estranhamente atribuímos verdade ou falsidade como se os critérios fossem tangíveis como cores ou moléculas. É uma crença que de fato formamos, mas que, todavia, sempre acaba sendo falsa pois estas propriedades misteriosas simplesmente não existem; assim a nomenclatura da teoria de Mackie como “teoria do erro”.

Simon Blackburn concorda em parte, pois atribui a estas crenças um “semi-realismo”, ou seja, de fato estas crenças tratam-se de coisas reais: das expressões de agentes acerca de seu julgamento do mundo e da objetividade moral que tais pessoas pressupõe do resto do mundo. Todavia, desta maneira, as crenças são puramente subjetivas, resultando parcialmente em um realismo moral que, todavia, não pressupõe uma objetividade do juízo moral; portanto a nomenclatura de “*semi-realismo*”, em contraste à ambos não-cognitivistas puros e realistas morais. Blackburn também critica amplamente Mackie em que sua teoria se refere apenas a como usamos a linguagem para falar de moralidade sem abarcar a dinâmica em que a língua se adapta e se reinventa – o que também também implica em uma crítica de como sua “teoria do erro” funcionaria

---

<sup>53</sup> No original: “*Argument for queerness*”.

<sup>54</sup> No original: “*moral concepts purport to designate properties that are utterly queer. We have no reason to think that such properties exist. Therefore, moral concepts are really vacuous. Ethical judgments are false or meaningless because they ascribe properties that are no more real than fairies or phlogiston.*”

acerca de assuntos não-morais, uma vez que não há uma linguagem imutável e completa de todas as coisas presentes no mundo factual.

Como última definição de um naturalismo moral “padrão” propriamente dito, gostaria de considerar a posição de Copp. O autor parte da definição-padrão que vimos acima para defender que “uma propriedade é natural apenas se e somente se quaisquer proposições sintéticas sobre sua instanciação que podem ser conhecidas, podem apenas ser conhecidas empiricamente”<sup>55</sup> (COPP, 2007, p. 39) e assim, “o naturalismo seria definido em termos do que podemos ou não conceber a partir da experiência” (BARBOSA, 2015, p. 44). Há um realismo em que

determinadas propriedades morais, tais como bondade, justiça e virtude, são propriedades naturais, por isso o naturalismo ético, embora seja uma espécie de realismo moral, dispõe de uma condição específica de admitir propriedades naturais e se distingue de outras formas de realismo. (BARBOSA, 2015, p. 41)

Assim, diferentemente de Mackie e Blackburn, Copp não pretende comprometer-se com um naturalismo irrestrito, apenas com um realismo das propriedades morais como propriedades naturais (COPP, 2007, p. 34). Estas propriedades podem ser conhecidas de forma empírica: de maneira análoga a Nicholas Sturgeon, a ética não é como a ciência, mas o conhecimento ético pode ser sujeito à observação e indução empírica. Assim, há uma caracterização epistemológica das propriedades morais em que elas são necessárias para podermos fazer uma abordagem científica completa do mundo natural como estudado pelas ciências. Todavia, nem todo conhecimento empírico é científico (já que o naturalismo ético, segundo ele, é diferente de um cientificismo), ou seja, não precisa adequar-se a teorias científicas, uma vez que podemos ter conhecimento baseado em nossas sensações próprias.

A diferença então com o não-naturalismo jaz na questão do acesso epistêmico ao conhecimento empírico de uma proposição sintética, que configura uma propriedade natural. O foco no sintético aqui, para Copp, não nega a existência de verdades conceituais. A questão aqui é de como se dá nossa

---

<sup>55</sup> No original: “*a property is natural if and only if any synthetic proposition about its instantiation that can be known, could only be known empirically*”

imersão no mundo natural através de nosso acesso epistêmico. Desta maneira, Copp delinea sua concepção naturalista predominante:

A ideia subjacente é que verdades morais refletem fatos empíricos sobre a natureza humana, as necessidades de sociedades e afins. Estes fatos variam de mundo possível para mundo possível, então as verdades morais podem também variar de mundo possível para mundo possível. É por isto que precisamos ter experiência de nosso mundo para ter conhecimento moral. Passamos a ter conceitos morais, ambos ‘grossos’ e ‘finos’<sup>56</sup> e, além disso, passamos a ter uma ‘teoria’ normativa substantiva sobre quais coisas são certas e erradas, boas e ruins, virtuosas e viciosas.<sup>57</sup> (2007, p. 45)

De acordo com esta posição de Copp, segundo Barbosa,

nós conseguiríamos mapear o mundo natural ou descobrir os mapeamentos disponíveis deste mundo, os quais nos permitiriam identificar a presença de propriedades morais no mundo através de uma investigação empírica. (2015, p. 44)

Assim, dado que (1) temos valores culturais padrões e (2) somos capazes de refletir acerca de nossos valores, não há razão para o naturalista negar proposições morais fracamente *a priori*, uma vez que podem ser alteradas frente a novas evidências empíricas.

Todavia, não é apenas no contexto de discurso puramente filosófico – ou *de poltrona* – que se encontra o naturalismo. Outros autores deram início à empreendimentos que facilitaram o desenvolvimento de um relacionamento crescentemente mais próximo entre filosofia e as ciências naturais onde o naturalismo começa a afastar-se do viés puramente teórico para abarcar um arcabouço e perspectiva cada vez mais científica.

Na década de 70, Paul Churchland, que fora orientado por Wilfrid Sellars, outro naturalista que compartilhava uma perspectiva semelhante à de Quine,

---

<sup>56</sup> Terminologia advinda de Bernard Williams, que concebia conceitos morais em duas categorias principais: os “finos”, ou “tênuos” (“*thin*”), aqueles como “bom”, “virtuoso” e afins que são abrangentes e abstratos demais, não possuindo uma associação empírica adequada para representá-los; e “grossos”, “espessos” (“*thick*”) como coragem, mentira ou brutalidade, os quais possuem elementos práticos em seu entendimento que permitem exemplificação empírica.

<sup>57</sup> No original: “*The underlying idea is that moral truths reflect empirical facts about human nature, the needs of societies, and the like. These facts vary from possible world to possible world, and so the moral truths also might vary from possible world to possible world. This is why we need to have experience of our world in order to have moral knowledge. We come to have moral concepts, both ‘thick’ and ‘thin,’ and, beyond that, we come to have a substantive normative ‘theory’ as to which things are right and wrong, good and bad, virtuous and vicious.*”

aprofundou-se nas neurociências a fim de aplicar o conhecimento cada vez mais aprofundado da área sobre nossa consciência, cognição e comportamento em geral na filosofia. Ele e sua esposa, Patricia Churchland, foram responsáveis pelo desenvolvimento da *neurofilosofia*, uma perspectiva materialista eliminacionista extrema em que conceitos de “psicologia popular” (“*folk psychology*”), ou seja, ideias e linguajar advindo do senso comum ou de uma psicologia ou qualquer disciplina não informada pelas neurociências dariam lugar ao linguajar científico conforme avançamos no entendimento do sistema nervoso, genética e biológico em geral. Ao tratar de ética, por exemplo, em seu livro *Braintrust* (2008), Patricia Churchland resume o paradigma metodológico da neurofilosofia:

desenvolvimentos recentes nas ciências biológicas nos permitem ver através do emaranhado, começar a discernir caminhos revelados por novos dados. O fenômeno de valores morais, até então tão enigmáticos, o é menos agora. Não completamente claro, apenas menos enigmático. Ao valer-se de novos dados convergentes da neurociência, biologia evolutiva, psicologia experimental e genética, e dada uma estrutura filosófica conciliadora com tais dados, podemos agora abordar significativamente a questão de onde surgem os valores.<sup>58</sup> (CHURCHLAND, 2008, p. 3, tradução nossa)

Apesar da neurofilosofia não participar ativamente da elaboração de pesquisas práticas como posteriormente ocorre na filosofia experimental, seu impacto é inegável em seu desenvolvimento, além de debates atualmente relevantes na filosofia da mente, cognitiva e computacional.

Por fim, em uma linha semelhante aos Churchlands, Daniel Dennett, previamente aluno de Quine e orientando de Gilbert Ryle, foca na aplicação de conceitos neodarwinistas e uma compreensão do estado-da-arte científica para tratar de problemas filosóficos. Evidentemente, como já citado, esta não é uma perspectiva nova: a tentativa de endender a esfera destes problemas dentro deste tipo de compreensão científica informada com o montante da biologia e outras ciências da época – deixando de lado as referências óbvias à Aristóteles

---

<sup>58</sup> No original: [...] *recent developments in the biological sciences allow us to see through the tangle, to begin to discern pathways revealed by new data. The phenomenon of moral values, hitherto so puzzling, is now less so. Not entirely clear, just less puzzling. By drawing on converging new data from neuroscience, evolutionary biology, experimental psychology, and genetics, and given a philosophical framework consilient with those data, we can now meaningfully approach the question of where values come from.*”

e os “físicos da natureza” –, em especial na teoria de Darwin, pode ser facilmente rastreada aos pragmatistas ou mesmo aos modernos tardios como Richard Spencer<sup>59</sup>, conhecido contemporaneamente principalmente pela crítica exagerada de G. E. Moore à sua teoria no início do século XX (MOORE, 1993, p. 97–110).

Spencer fora talvez o primeiro a tentar desenvolver uma teoria filosófica com embasamento científico – “filosofia sintética”, como chamava (WEINSTEIN, 2019, seq. 1) – focado em Darwin, todavia baseou sua teoria num intuicionismo à maneira jusnaturalista associando os traços selecionados de que favorecem a adaptação com uma noção de “progresso”. Moore identificou semi-corretamente esta associação como a “falácia naturalista”<sup>60</sup>, ou confundindo algo intrinsecamente “bom” em teor moral com um fato da natureza. A identificação é parcial pois a metaética de Spencer distancia-se de um darwinismo social propriamente dito, terminologia utilizada para caracterizar teorias que se utilizam de uma interpretação limitada de Darwin a fim de justificar perspectivas elitistas cometendo tal falácia. Em seu trabalho ético, o autor foca num utilitarismo “racionalista” que, em contraste ao de cunho “empirista” de Bentham não pressupõe que é possível realizar escolhas de maneira ausente de intuições.

Promover a utilidade nunca é simplesmente uma questão de escolher opções, especialmente quando muito está em jogo, calculando e criticamente comparando utilidades. Ao invés, a emergência do raciocínio utilitário prático começa sempre que nossas intuições morais colapsam. A ciência moral testa e refina nossas intuições morais, que frequentemente provam-se “necessariamente vagas” e contraditórias.<sup>61</sup> (WEINSTEIN, 2019, seq. 4)

<sup>59</sup> Trato mais sobre Spencer em outros trabalhos (2014c, 2014b), todavia estas minhas críticas estão marcadas pelo exagero contemporâneo suscitado por Moore que citaremos a seguir.

<sup>60</sup> A falácia naturalista é um dos conceitos mais controversos e igualmente mal entendido na história contemporânea da filosofia. Já versei aprofundadamente sobre o tópico e os equívocos suscitados pela apresentação desorganizada de Moore em 2014c e 2014a. Em resumo, uma interpretação plena deste conceito conforme apresentado por Darlei Dall’Agnol (DALL’AGNOL, 2005, cap. 3) ao adaptar o inédito prefácio à segunda edição do *Principia Ethica* (1993), publicada postumamente apenas 90 anos depois do original em que Moore finalmente esclarece sua argumentação, envolve dois erros lógicos: erro categorial, como confundir um objeto com sua própria classe ou ele mesmo, ou erro de identificação equivocada, rastreável ao argumento da questão em aberto de identificar uma entidade holística apenas em uma de suas partes. Além disso, abrange também quaisquer argumentos realizados que cometam algum destes erros em seu processo inferencial. Ou seja, não é uma falácia *per se*, e nem trata sobre naturalismo como no entendimento aqui presente, mas sim sobre objetos naturais na lógica.

<sup>61</sup> No original: “*Promoting utility is never simply a matter of choosing options, especially when much is at stake, by calculating and critically comparing utilities. Rather, the emergence of utilitarian practical reasoning begins wherever our moral intuitions breakdown. Moral science tests and refines our moral intuitions, which often prove “necessarily vague” and contradictory.*”

Ironicamente, apesar da nomenclatura de sua veia de utilitarismo, a posição de Spencer se assemelha à imagem geral naturalista que pressupõe fatores externos à escolha além de puras deduções objetivamente racionais.

Apesar de Dennett não ter uma teoria ética propriamente dita, a maneira como procura tratar de problemas filosóficos em um viés evolucionista, de certa forma paralela à de Spencer, consiste com o empreendimento naturalista com foco empírico que será a base teórica para a aplicação da filosofia experimental como veremos a seguir. Nos aprofundaremos propriamente em sua escola de interpretação da teoria darwiniana, o neodarwinismo, e sobre as interações sobre biologia e comportamento no capítulo 4, onde trataremos da base conceitual das áreas científicas não especificamente filosóficas que constituirão parte do cerne dos capítulos posteriores deste trabalho.

O trabalho em que ele apresenta de maneira mais desenvolvida seu paradigma geral chama-se “A Ideia Perigosa de Darwin” (DENNETT, 1995)<sup>62</sup>. Nele, apresenta uma interpretação da teoria da seleção natural de Darwin – a qual analisaremos com maior profundidade no próximo capítulo ao apresentarmos a importância da mesma não apenas para o presente trabalho, mas como para a filosofia em geral – e quais as implicações dela para diversas outras áreas do conhecimento como linguagem, epistemologia, matemática e ética. Embora seja mais um trabalho de divulgação científica do que um material fortemente embasado em todas as áreas sobre as quais pretende versar sobre, cumpre um papel importante de aproximação paradigmática de todas estas áreas com uma compreensão naturalista do universo.

Em suma, Dennett identifica nesta obra que Darwin teve a percepção de que o sistema que rege toda a vida como conhecemos é baseado em um algoritmo automático. Anteriormente, perspectivas teológicas teleológicas permeavam a biologia como parte de uma obra divina do cristianismo, onde toda a vida seria uma criação planejada de uma figura sobrenatural. Mas as observações e análises anatômicas comparativas de Darwin, baseadas em sua viagem intercontinental no navio HMS Beagle, somadas à análise de fósseis, etologia e torrenciais outras evidências permitiram que ele percebesse este

---

<sup>62</sup> “*Darwin's Dangerous Idea*” no original.

padrão de lenta mudança em espécies. Seus genótipos, fenótipos e comportamentos estariam intrinsecamente atrelados a este processo constante de busca ao que Dennett atribui como “estratégias evolutivamente estáveis” (DENNETT, 1995, p. 254, tradução nossa) para a sobrevivência frente às dificuldades do contexto ambiental em que os espécimes estão inseridos.

Assim, um processo mecânico, cego, ausente de intencionalidade seria capaz de auto-regulação em prol da adaptação daquela espécie no meio em que vive, gerando cada vez mais sistemas mais complexos – não necessariamente perfeitos ou “melhores”, apenas que favoreceram ou simplesmente ocorreram por acaso e não atrapalharam a sobrevivência. A “ideia perigosa” então é que para se ter um bom *design* não é necessário um *designer*, ou alguma pressuposição anterior intencional (divina ou não) que realiza o planejamento de seres vivos com algum fim em mente.

As interconexões entre filosofia e ciência são intrínsecas em tal perspectiva segundo ele, pois “a Revolução Darwiniana é ambas, uma revolução científica e uma revolução filosófica, e nenhuma revolução poderia ter ocorrido sem a outra” (DENNETT, 1995, p. 21, tradução nossa). Em outros trabalhos como “*Elbow Room*” (1984) e “*Freedom Evolves*” (2004), que respectivamente podem ser traduzidos como “espaço de manobra” e “liberdade evolui”, o autor trata especificamente do problema do livre-arbítrio no debate filosófico com o paradigma apresentado acima.

O primeiro, que também pode ser traduzido literalmente como “espaço para os cotovelos”, faz jus à questão de que nosso feitiço biológico por um lado nos limita por não termos uma “liberdade plena” para fazermos qualquer coisa: da mesma forma que não podemos voar sozinhos por não nascermos com asas, nossas decisões e pensamentos em geral são bastante limitados de acordo com nosso feitiço neurológico e nossas experiências subjetivas. Todavia, isto não engendra um determinismo absoluto; pelo contrário. Tal feitiço na verdade nos possibilita pensar conceitos como estes, e refletir acerca de nossas capacidades psicológicas e tomar decisões mais bem-informadas. Afinal, como diria Carl Sagan, “somos uma maneira do Cosmos conhecer a si mesmo” (“Cosmos”, 1980), fazendo referência ao fato de que nossos átomos advêm de fissões nucleares do centro de estrelas muito, muito distantes e paulatinamente uniram-se e foram selecionados até formar nossa espécie, que é por sua vez capaz de

entender este processo. Assim, ao invés do conceito de “liberdade” ser entendido como algo absoluto, deve ser entendido como este “espaço de manobra” de nossos “cotovelos” frente à todas nossas limitações – embora o autor lembre que o que ocorre é o contrário, nossas capacidades fisiológicas nos *capacitam* a pensar além delas, ao invés de nos impedir.

Em seu segundo livro, trabalha não apenas como esta compreensão de liberdade é capaz de ampliar-se conforme as capacidades cognitivas em espécies ficam mais complexas – cachorros possuem maior liberdade que amebas, assim como nós a temos mais que cachorros – (DENNETT, 2004, p. 162–166), como compreender como ela funciona nos capacita a exercitá-la de maneira melhorada (DENNETT, 2004, cap. 9 e 10). Inclusive muito do que veremos posteriormente sobre tomada de decisão ele trata nestes trabalhos, mas por tratarem-se de trabalhos predominantemente de divulgação científica não possuem profundidade e peso nos tópicos em questão como necessitamos aqui.

Na verdade, há um grande fator limitador no naturalismo filosófico que o impede de ser exclusivamente a estrutura deste trabalho. Raros são os autores supracitados que de fato engajam com evidências empíricas diretamente a fim de guiar suas teorias; talvez apenas os Churchlands e Dennett, mas ainda assim de uma maneira bastante secundária. Como já dito, o objetivo deste trabalho é analisar e entender as implicações de evidências empíricas e análises de outras áreas relevantes acerca de nossa agência, e para isto devemos dar um passo adiante para uma linha filosófica que cumpre exatamente este papel: a filosofia experimental.

### 3.2 Filosofia experimental

*A noção que o filósofo moral pode estudar os conceitos da moralidade meramente refletindo, numa poltrona no estilo de Oxford, no que ele ou ela e aqueles à sua volta dizem é improdutivo.*

– Alasdair MacIntyre, “Depois da Virtude”

Tendo em vista este crescente diálogo entre o naturalismo filosófico e as ciências aplicadas com as quais procura interagir, não surpreende que uma metodologia híbrida de filosofia aplicada e teórica em diálogo interdisciplinar

direto tenha surgido a fim de avançar este debate para além da poltrona filosófica. Ao contrário de autores como Quine e outros que prezavam o papel da ciência no empreendimento filosófico em um naturalismo metodológico, porém raramente citavam ou dialogavam com pesquisas de fato empíricas por acadêmicos de áreas afins ao tema em questão, os experimentalistas procuram tratar diretamente com evidências e outros pesquisadores de diversas áreas, além de realizar ativamente experimentos práticos como parte de seu processo de pesquisa com frequência. Apesar de podermos rastrear esforços de alinhar teorias filosóficas com práticas científicas à Aristóteles e mesmo alguns outros da antiguidade, a disciplina experimental *per se* na filosofia surge apenas mais de dois mil anos depois no século XXI, quando autores não apenas dialogam como participam com outras áreas aplicadas do conhecimento, ou como Jesse Prinz coloca, “arregaçando suas mangas e sujando suas mãos”<sup>63</sup> (PRINZ, 2007b, p. 189).

Há duas maneiras de se proceder quando se utiliza da filosofia experimental para tratar de problemas filosóficos, ou dois projetos: o negativo e o positivo. O projeto negativo engendra tratar de problemas ou autores já existentes e analisa-los sob o crivo das evidências empíricas. Quando Aristóteles fala que um agente virtuoso possui um caráter firme que o guia a agir de acordo com a atitude mais virtuosa, ele está fazendo uma série de afirmações empiricamente analisáveis acerca de nossa psicologia e capacidades cognitivas em geral. Assim, podemos comparar suas afirmações sobre a existência de caráter, a possibilidade de virtudes psicológicas e inúmeras outras com dados já existentes na literatura científica sobre o assunto – e de fato vamos fazê-lo no capítulo 6, uma vez que envolve um diálogo direto com o problema da tomada de decisão sob um viés filosófico – ou mesmo realizarmos experimentos a fim de obtermos novas informações com ferramentas tecnológicas e metodológicas que foram surgir mais de dois mil anos após a morte do autor em questão. Além disso, presume-se uma concepção universal e objetiva do que caracteriza uma virtude ou vício, de como se age etc. sem levar em consideração como diferentes povos ou grupos podem interpretar estas coisas – uma posição associada com a tradição filosófica predominante a partir do diálogo de Teeteto, por Platão, onde

---

<sup>63</sup> No original: “*rolling up their sleeves and getting their hands dirty*”.

Sócrates busca inquirir acerca de uma concepção única, absoluta do que é conhecimento (KNOBE; NICHOLS, 2017, seq. 2). Além destes exemplos, vamos utilizar o projeto negativo da filosofia experimental também para o utilitarismo e a deontologia, especificamente acerca de Kant, a fim de aprofundar o debate filosófico-experimental sobre nossas ações.

Por outro lado, o projeto positivo visa engajar com problemas filosóficos em geral não com o objetivo de criticar posições já estabelecidas, mas para ir além delas a partir de sua metodologia diferenciada (ALFANO; LOEB, 2017, seq. 2.2). Embora invariavelmente teorias que já versaram sobre o tópico em questão acabem surgindo em determinados momentos, elas não são o foco da investigação. Este é o caso do presente trabalho: tratar do problema da agência a partir dos debates recentes nas ciências empíricas. Um passo além, mais característico ainda do projeto positivo em filosofia experimental, seria efetivamente planejar ou reproduzir uma ou mais pesquisas empíricas acerca do tópico a fim de ampliar ou estender a gama de dados e entendimento acerca do assunto; infelizmente o autor não foi capaz de tal durante o período da pesquisa por uma série de fatores, resumindo-se à revisão bibliográfica como fonte de material e debates.

Embora não iremos nos comprometer com nenhuma teoria experimentalista específica, e sim com sua metodologia predominantemente de maneira positiva, cabe apresentar algumas concepções já estabelecidas por outros autores que dialogarão com o material a fim de entendermos como é possível, então, relacionar de maneira produtiva pesquisa empírica com teorias filosófica<sup>64</sup>. Apresentarei brevemente a concepção de Jesse Prinz, que discute com uma ampla gama de autores naturalistas e não-naturalistas e é um dos filósofos experimentais mais proeminentes internacionalmente em termos de produção filosófica e pesquisa empírica. Contrastarei-o com Mackie a fim de ilustrar algumas diferenças procedimentais acerca do tema do desacordo moral, posição que trabalhamos anteriormente ao apresentar o naturalismo, com foco no debate acerca da cognitividade do juízo moral, tópico ainda mais central a este trabalho.

---

<sup>64</sup> Apresento a posição do autor mais aprofundadamente com algumas discussões em 2014a.

Prinz defende uma posição quase diametralmente oposta à de Mackie, que vimos anteriormente, desenvolvendo o que chama de “sentimentalismo construtivista”. O autor, como os naturalistas supracitados, preocupa-se com as bases naturais da moralidade, ou uma “genealogia da moral” de cunho naturalista. Apesar do nome de sua obra principal sobre moralidade ser *A Construção Emocional da Moral* (2007)<sup>65</sup>, Prinz procura defender um cognitivismo construtivista de cunho emotivista empiricamente informado – todavia, sua conclusão está quase mais próxima de um não-cognitivista do que parece. Isto se dá, pois segundo o autor, a moralidade depende fortemente da cultura para sua normatividade, mas vai além no que tange à apofanticidade do juízo moral. A estrutura para realizá-lo não é inata, todavia depende de mecanismos fisiológicos e, por conseguinte, processos psicológicos que são majoritariamente comuns à espécie humana como um todo. A avaliação moral, portanto, surge em comunidades cujas capacidades cognitivas e comportamentais associadas à moralidade são compartilhadas, e de uma maneira análoga à crítica de Wittgenstein à linguagem privada, não existe juízo moral realizado de maneira individual. Assim,

a moralidade se dá apenas em comunidades morais, [e] uma vez inserido, toda a asserção moral que tal agente fizer estará embasado em suas faculdades fisiológicas e em um contexto de verificação pública na qual ele estará inserido. É evidente portanto que, formulado desta maneira, invariavelmente todo o juízo moral seja verdadeiro. (IGANSI, 2014a, p. 65)

Assim, o ponto fulcral de diferença com Mackie jaz em um termo: utilidade. Afinal, ambos concordam que há uma ausência de necessidade de conexão ontológica entre as propriedades de objetos no mundo e conceitos morais. Mackie, todavia, utiliza-se deste argumento para defender a total vacuidade – e, por conseguinte, falsidade – do juízo moral, enquanto Prinz realiza um paralelo entre ele e o juízo de cor como na perspectiva de Locke: como qualidades secundárias dos objetos, ou seja, que possuem realidade física, porém dependem dos sentidos subjetivos dos agentes. Juízos morais e cores possuem propriedades empiricamente analisáveis, como a maneira como nosso corpo

---

<sup>65</sup> “*The Emotional Construction of Morals*” no original.

funciona ao interpretar frequências de luz ou reagindo a uma determinada ação moralmente relevante, por exemplo.

Mas estes elementos empíricos não correspondem estritamente a uma cor ou a alguma valoração moral específica: diferentes iluminações, condições fisiológicas do olho do observador e diversos outros elementos influenciam na cor que vemos, assim como criação, humor no momento, quem está envolvido no evento moral e diversos outros elementos influenciam no ajuizamento moral. Assim, conceitos morais e conceitos de cor não dependem necessariamente da mente, e não há garantia que sejam alcançados em critérios perfeitos. Todavia, o linguajar acerca de cores é predominantemente útil no dia-a-dia para uma miríade de situações, e da mesma maneira o é a linguagem moral. Assim, “[e]xitem fatos morais; mas eles são apenas fatos puramente subjetivos” (PRINZ, 2007, p. 138, tradução nossa). O juízo moral não é objetivo nem existe independentemente de agentes morais, mas é ontologicamente real enquanto uma propriedade secundária, é cognoscível e *útil* – e, portanto, não é vácuo.

Para finalizar, não apenas suas perspectivas destoam-se, como sua metodologia. Mackie trata apenas de maneira breve de dados empíricos da antropologia ao trazer o problema do relativismo que serve como ponto de partida para seu antirrealismo (ALFANO; LOEB, 2017, p. 32–33). Prinz, por outro lado, utiliza-se de montantes de evidências empíricas em cada parte de cada um dos seus trabalhos – isto quando não é ele mesmo parte da equipe de pesquisadores adquirindo estes dados práticos. Embora filosoficamente ambos possuam valor discutivelmente amplo e diferenciado em suas respectivas áreas, o que nos é mais interessante aqui é que Prinz procura contrastar e dialogar suas posições e argumentos diretamente com outras ciências. Desta maneira, invariavelmente seu trabalho dá melhor conta de dados acerca do juízo moral do que Mackie – como a própria crítica de Blackburn que vimos anteriormente também sugere.

Apesar de ser nossa principal perspectiva de trabalho, esta seção dedicada à filosofia experimental mantêm-se curta pois iremos nos aprofundar de maneira prática em sua metodologia propriamente dita ao aplica-la no capítulo 6, onde debateremos sobre as implicações filosóficas de diversos dados científicos sobre o processo de tomada de decisão uma vez que trataremos de autores e discussões já dentro da literatura até então – inclusive muito do

material que levou Prinz às conclusões supracitadas. A importância deste viés em nossa pesquisa filosófica é que “filósofos geralmente tentaram estabelecer a conexão entre emoção e juízo moral em reflexões de poltrona” (PRINZ, 2006, p. 31), tratando de tópicos diretamente empíricos sem nem ao menos se aproximar de um livro de psicologia ou biologia. Aqui, como é o propósito central deste trabalho, procuraremos uma perspectiva integrativa interdisciplinar da investigação do juízo moral e do papel das emoções e da razão em sua elaboração.

O intuito deste capítulo fora elaborar as estruturas filosóficas com quais as teorias, conceitos e evidências de ambos os próximos capítulos irão dialogar. Observamos o escopo do naturalismo enquanto escola filosófica e como destoa-se da tradição em geral à parte de Hume e poucas outras exceções, a fim de desenvolver-se especificamente na filosofia experimental, que procura elaborar esta ponte direta entre um viés científico na filosofia e as investigações e dados relevantes para o debate de problemas filosóficos. O próximo capítulo nos auxiliará a ilustrar que tipo de visão científica é esta da qual nosso naturalismo partirá com base na biologia evolutiva, economia comportamental, psicologia cognitiva e neurociências. O posterior a este, o capítulo 5, trará então dados destas ciências os quais tratam dos problemas de racionalidade e agência dentro desta metodologia empírica. No capítulo 6, como dito acima, seguiremos para debater estes dados com estes problemas em cunho filosófico aplicando diretamente a metodologia da filosofia experimental sobre os dados vistos no capítulo anterior e em outros dela própria, para então analisarmos o estado-da-arte do tópico central para este trabalho: é a racionalidade racional?

#### 4. A ESTRUTURA CIENTÍFICA DA TOMADA DE DECISÃO

*O homem em sua arrogância considera-se um grande trabalho merecedor da interposição por uma divindade. Mais humildes e eu acreditamos ser mais verdadeiro considerá-lo criado a partir dos animais.*  
– Charles Darwin<sup>66</sup>

Tendo estabelecido a estrutura e paradigma filosóficos no qual este trabalho se apoia, nominalmente dentro de uma compreensão teórica naturalista alinhada à uma prática experimental, procuraremos entender neste capítulo como o que a filosofia entende por “racionalidade” relaciona-se com pesquisas em outras áreas do conhecimento focadas em testagem empírica e análise prática como a biologia evolutiva, economia comportamental, psicologia cognitiva e neurociências.

Como mencionado anteriormente, as problemáticas relacionadas à racionalidade e deliberação na filosofia tendem a cair sob a definição do termo abrangente “tomada de decisão” nas outras ciências. Da economia às neurociências, os mais diversos tópicos acerca da esfera de características do fenômeno da escolha e ação nas mais variadas áreas congruem-se sob esta terminologia. Para tratarmos do processo de tomada de decisão, ou seja, investigarmos o processo deliberativo em harmonia com as disciplinas que não são estritamente filosóficas, é necessário antes nos familiarizarmos com sua terminologia e bases teóricas.

Portanto, iremos examinar brevemente o arcabouço técnico de quatro principais grandes áreas de pesquisa fora da filosofia que se sobrepõem significativamente no tópico da deliberação: (i) da biologia evolutiva, que nos irá prover o paradigma científico abrangente do agente filosófico enquanto animal biológico inserido num contexto natural de seleção genética e fenotípica que diretamente molda sua capacidade cognitiva e todos os outros aspectos envolvidos no decidir; (ii) da economia comportamental, área da economia responsável por analisar a racionalidade de agentes na tomada de decisão; (iii) da psicologia cognitiva, que procura tratar de tópicos caros à epistemologia de maneira estritamente científica e experimental; e, por fim, (iv) das neurociências, que englobam tópicos da psicologia cognitiva e do funcionamento geral do

---

<sup>66</sup> Presente em seus manuscritos 21 anos antes da publicação da “Origem das Espécies”, (P. H. BARRETT, 1974, p. C197)

sistema nervoso e sua relação com comportamento, percepção e nossa fisiologia em geral.

Esta divisão não apenas é didática como também cronológica: o entendimento evolutivo da biologia permitiu uma compreensão “animalista” da humanidade – ou seja, tratando pessoas não como seres especiais, mas como inseridas em um contexto análogo ao de outros animais, tendo desenvolvido apenas características diferentes, mas igualmente naturais –, proporcionando à economia a compreensão de nosso processo racional como imperfeito, dentro das limitações do mundo real. Isto estimulou estudos computacionais e cognitivos que permitiram o rápido crescimento da psicologia cognitiva, que por sua vez desenvolveu as ferramentas metodológicas para tratar da mente como um problema científico adequadamente, propiciando a ascensão das neurociências como área especializada na análise dos processos subjacentes de nossas faculdades cognitivas.

Um dos principais conceitos que permeia todas estas linhas de pesquisa é o de viés cognitivo, frequentemente representado por uma de suas instâncias mais comuns (e origem da área de pesquisa) que é o da dissonância cognitiva. Viéses cognitivos, em suma, tratam exatamente destes elementos alheios à deliberação em questão que acabam influenciando, podendo ser uma extensa variedade de fatores desde estar com fome ou sentir nojo até o posicionamento no campo de visão e inclusive preconceitos. Assim, trataremos com maior atenção em uma seção própria englobando elementos de todas as ciências em questão para posteriormente tratar de diversos exemplos empíricos onde ele é observável.

Assim como na filosofia, até o século XX a predominância de modelos teóricos sem devido embasamento empírico também predominava em todas estas áreas, todavia por motivos diferentes: a prática científica na época ainda estava limitada metodologicamente e tecnologicamente. Neste capítulo, vamos tratar brevemente de apresentar estas áreas em seus conceitos, terminologia e práticas a fim de familiarizarmo-nos e sermos capazes de entender a plenitude dos experimentos e análises posteriores, e a relevância que têm para o estudo da tomada de decisão. Por fim, apresentaremos também o conceito de “processo duplo”, que abrange todas estas áreas e trata especificamente do processo deliberativo e dos vários elementos e aspectos que o influenciam, noção esta

que rendeu o prêmio Nobel em Economia dos seus proponentes e foi altamente influente nas recentes décadas (KAHNEMAN, 2013).

#### 4.1 Um paradigma evolutivo<sup>67</sup>

A “síntese moderna”, também conhecida como “neodarwinismo” como vimos anteriormente, é uma adaptação da teoria da seleção natural original de Charles Darwin informada por uma perspectiva pós-Mendeliana de genética, onde toda a concepção darwinista de seleção de traços fisiológicos e comportamentais, adaptabilidade e hereditariedade é algoritmicamente compreendida e formalizada<sup>68</sup> de acordo com as novas informações biológicas do último século. Assim, um tópico como o da racionalidade é tratado de acordo com seu contexto biológico e evolutivo, como um sistema cognitivo fruto de milênios de seleção natural e composto de outros diversos sistemas mais simples – ou “guindastes”<sup>69</sup> – que se desenvolveram de determinada maneira (aleatoriamente ou artificialmente) e por acaso foram bem-sucedidas. Estes sistemas variam da percepção até a capacidade da linguagem e o desenvolvimento da cultura, que se superam e se acumulam entre si gerando sistemas mais complexos. Em suma,

Junto ao caminho de amebas e cucos para nós, houve uma transição majoritária com forças rivais ao nascimento endossimbiótico dos eucariontes: a evolução da linguagem e cultura, um do dos grandes guindastes da evolução.<sup>70</sup> (DENNETT, 2009, p. 10063, tradução nossa)

É importante, claro, uma elucidação da definição de evolução. Ao contrário do estigma tradicional do termo, “evolução” não designa melhoramento, mas sim meramente uma alteração. Um traço pode ser selecionado, ou “evoluir”,

---

<sup>67</sup> Sou imensamente grato aos importantíssimos debates com o prof. dr. Matheus Silveira, me. Rafael Hiller, dra. Heloísa Allgayer e à excelentíssima prof. dra. Anna Carolina Regner por me ajudarem a navegar Darwin corretamente.

<sup>68</sup> Embora Darwin seja possivelmente um dos autores mais importantes na história da humanidade por permitir a compreensão do desenvolvimento da vida em termos científicos, sua teoria original sofre em termos de formalidade e metodologia científica. Recomendo a leitura de “Significado e função do princípio de seleção natural na Origem das Espécies” de Heloísa Allgayer (2013) não somente pelo excelente trabalho organizacional da obra de Darwin como também para sua compreensão.

<sup>69</sup> “Cranes” no original.

<sup>70</sup> No original: “*along the path from amoebas and cuckoos to us, there was a major transition with powers to rival the endosymbiotic birth of the eukaryotes: the evolution of language and culture, one of the great cranes of evolution*”.

tanto por facilitar a sobrevivência de uma espécie em um determinado contexto – polegar opositor, faro sensível, urina com odor forte contra predadores –, como também simplesmente pode se desenvolver de maneira a não afetar negativamente ou apenas significativamente tal probabilidade de sobrevivência, como nossos dentes sisos, vesícula biliar, baço ou apêndice<sup>71</sup>.

Darwin desenvolve este entendimento do funcionamento geral da vida no contexto revolucionário antropocêntrico pós-iluminista, onde já era reconhecida a perspectiva “desmistificada” da humanidade e da razão em Hume e outros como vimos anteriormente, somada à leitura da teoria de Thomas Malthus onde a progressão aritmética dos recursos e geométrica das populações gera uma desproporção entre ambos, resultando na luta para apoderar-se de recursos e espaço. Isto permitiu que Darwin reconhecesse que há um princípio comum que rege tal diferença, um princípio mecânico na própria natureza responsável por “administrar” o desenvolvimento dos seres orgânicos, ou em outras palavras, moldam as características que os definem como espécie e influenciam em como interagem com o ambiente, de forma a torna-los mais aptos à sobrevivência.

Como muitos mais indivíduos de cada espécie nascem que podem possivelmente sobreviver; e como, conseqüentemente, há uma luta recorrente pela existência, segue-se que qualquer ser, se variar mesmo minimamente de alguma maneira lucrável para si, sob as complexas e ocasionalmente variáveis condições de vida, terá uma chance melhor de sobreviver, e assim ser selecionado naturalmente. Do princípio forte da herança, qualquer variedade selecionada tenderá a propagar sua nova e modificada forma.<sup>72</sup> (DARWIN, 1876, p. 3, tradução nossa)

Assim, a síntese moderna reinterpreta o processo evolutivo de seleção como um algoritmo responsável por reger a vida de maneira cega, não-intencional – a “perigosa ideia de Darwin” segundo Dennett (1995, p. 59), que vai além e propõe que o argumento do princípio de Darwin

---

<sup>71</sup> Ainda que muitos destes ofereçam chance de perigo à vida, o desenvolvimento evolutivo não se importa tanto com longevidade individual como com a capacidade de procriação e cuidado mínimo da prole a fim de garantir a sua sobrevivência.

<sup>72</sup> Do original: “As many more individuals of each species are born than can possibly survive; and as, consequently, there is a frequently recurring struggle for existence, it follows that any being, if it vary however slightly in any manner profitable to itself, under the complex and sometimes varying conditions of life, will have a better chance of surviving, and thus be naturally selected. From the strong principle of inheritance, any selected variety will tend to propagate its new and modified form”.

consiste de duas maneiras de demonstrações, a demonstração lógica que um certo tipo de processo iria necessariamente ter algum certo tipo de resultado, e a demonstração empírica que as condições necessárias para tal tipo de processo de fato foram encontradas na natureza.<sup>73</sup> (DENNETT, 1995, p. 48–49 tradução nossa)

Nominalmente, se tivermos (i) *variações*, ou seja, mutações de um indivíduo para o outro, de uma prole para a outra, (ii) *seleção*, entre as variações que são mais aptas a sobreviver num ambiente, e (iii) *hereditariedade*, onde a seleção de variações segue com a procriação, teremos inexoravelmente uma consequência lógica necessária que é a evolução pela seleção natural. Ou seja, este

[...] é o mecanismo pelo qual são preservadas e acumuladas variações quando úteis aos indivíduos que as possuem, favorecendo sua sobrevivência e de sua prole, que herdará as modificações exibidas pela espécie parental. (ALLGAYER, 2013, p. 14)

De maneira mais recente e holística com o todo da pesquisa interdisciplinar surgiu a síntese estendida, que insere no paradigma geral evolutivo os recentes dados sobre epigenética e outras áreas. Nela,

a noção da natureza enquanto um sistema perde o significado. Pelo contrário, ela passa a ser compreendidos como uma série de fenômenos probabilísticos e cegos, onde o direcionamento concebido como sobrevivência do mais apto perde sua finalidade e, conseqüentemente, seu aspecto teleológico, enquanto elemento que preserva o sistema chamado natureza. (SILVEIRA; ALLGAYER; HILLER, 2016)

Uma compreensão naturalista do agente filosófico pressupõe, portanto, que nossa biologia e ambiente interagem de maneira intrinsecamente relacionada e bilateral. Nosso genótipo e fenótipo, ou como nossa carga genética é selecionada, combinada e manifestada de maneira orgânica não é determinante de nosso comportamento ou potencialidades, mas capacitador: os sistemas que desenvolvemos ao longo de nossa história evolutiva formam redes de “guindastes” que permitem estruturas fisiológicas e cognitivas capazes de

---

<sup>73</sup> No original: “*consists of two sorts of demonstrations, the logical demonstration that a certain sort of process would necessarily have a certain sort of outcome, and the empirical demonstration that the requisite conditions for that sort of process had in fact been met in nature*”.

perceber, entender e processar o mundo de cada vez mais formas permitindo interagir com ele de uma maneira muito mais ampla e complexa.

Ressaltando esta interdependência entre natureza e ambiente para o desenvolvimento e seleção de características, há alguns elementos originais de Darwin que gostaria de considerar. Primeiramente temos o hábito, que diferentemente do uso comum do termo, ou até mesmo de autores como Hume, "dentro do contexto da biologia, refere-se às ações instintivas de animais e às tendências naturais das plantas" (SILVEIRA, 2010, p. 75, nota de rodapé) em uma relação direta com sua fisiologia. Através da alteração de hábitos do organismo é possível gerar variações fisiológicas que, por sua vez, podem ser herdadas fisiologicamente por sua prole. O tamanho das mamas de cabras e vacas em locais onde são constantemente ordenhadas é claramente maior em comparação às que não o são; um exemplo também da seleção artificial, onde há a influência direta de humanos para a seleção intencional de determinados traços em outras espécies (SILVEIRA, 2010, p. 75–76). Da mesma maneira, ambientes e traços que favorecem a alimentação, ou seja, disponibilidade de nutrientes e/ou características que permitam sua melhor captação também influenciam na sobrevivência e conseguinte seleção natural (SILVEIRA, 2010, p. 77–82).

Naturalmente, segue-se que não apenas elementos materiais influenciam na sobrevivência de indivíduos. Darwin acrescenta a seleção sexual como elemento importante dentro da seleção natural, onde há uma seleção interna a espécies acerca de indivíduos mais prováveis a procriarem em detrimento de outros. Características que influenciam este tipo de seleção podem ser físicas, como aparência, por exemplo "quando machos e fêmeas de qualquer animal possuem os mesmos hábitos gerais da vida, mas diferenciam-se em estrutura, cor ou ornamento, tais diferenças foram predominantemente causadas pela seleção sexual" (DARWIN, 1952, p. 90, tradução nossa)<sup>74</sup>.

Mas estas características também podem ser sociais, o que nos é o mais importante aqui. Como vimos, nossa estrutura fisiológica foi capaz de gerar "aparelhagens" que guiam nosso comportamento. Estas aparelhagens englobam

---

<sup>74</sup> No original: "*when the males and females of any animal have the same general habits of life, but differ in structure, colour, or ornament, such differences have been mainly caused by sexual selection*".

toda nossa percepção do ambiente através de nossos sentidos, nossos processos cognitivos que processam estes estímulos e também a expressão resultante deste processo de maneira comportamental. Indivíduos vivendo na floresta que vêem uma forma de alguma criatura na escuridão onde não há nada e acaba tendo medo e fugindo, por exemplo, possuem mais probabilidade de sobreviver do que aqueles que nunca fogem, pois eventualmente pode ser que tenha algo ali. De modo semelhante, as interações sexuais entre indivíduos estão intrinsecamente relacionadas com suas relações sociais, pois “os indivíduos que possuíam maior sociabilidade e tinham características que favoreciam o convívio social, uma vez que o sucesso da caça e da proteção dependem diretamente dessa convivência” (SILVEIRA, 2010, p. 83) – especialmente em animais sociais como humanos e lobos.

Assim, as expressões fisiológicas e estrutura psicológica desenvolvidas nas interações sociais de espécies homínidas a partir do período Pleistoceno (de 200.000~120.000 anos atrás até cerca de 8.000 anos atrás) – período em que nossa espécie surgiu – nas savanas gramíneas africanas com caça de tamanho razoavelmente superior ao nosso favoreceram na seleção natural aqueles que por algum motivo optaram por viver em grupo. Geneticamente, estes primeiros humanos compartilham entre 99,4% e 99,9% de seu DNA conosco – assim como eu e você (presumindo que você também seja membro dos *homo sapiens*), ou qualquer um de nós com quaisquer outros seres humanos no planeta terra (AUTON et al., 2015; KOWALSKI, 2007). Ou seja, todas nossas variações idiossincráticas de fisiologia e comportamento estão compreendidas em uma relação entre essa ínfima diferença genética e os diversos ambientes e vivências nos quais nos encontramos e desenvolvemos – mais uma demonstração de que não há uma divisão entre natureza e ambiente, mas um diálogo constante e íntimo que engloba todas as maneiras em que nossas inclinações, capacidades e limitações naturais interagem com a ambientação na qual estamos inseridos.

Torna-se claro aqui, então, como uma estrutura psicológica complexa se desenvolve a fim de dar conta de uma aprimoração constante de habilidades sociais para a crescente necessidade de convívio. A expressão das emoções então torna-se elemento-chave nestas interações, sendo foco de estudo posterior de Darwin em sua obra “A expressão da emoções nos homens e nos

animais” (1890), onde estuda as características fisiológicas das emoções etologicamente. Aqui, a palavra “expressão deve ser entendida como modificações ou movimentos em qualquer parte do corpo, como o balançar da cauda de um cão, o levantar de ombros de um homem ou a dilatação dos capilares da pele” (SILVEIRA, 2010, p. 90).

A tese principal é que há uma relação intrínseca entre emoções e manifestações comportamentais e fisiológicas: “[a] expressão do medo, por exemplo, caracteriza-se por gritos e um aumento da frequência cardíaca, que pode levar a um suadouro e ao enrubescimento da pele” (SILVEIRA, 2010, p. 85). Darwin dedica-se à análise do comportamento infantil, de pessoas com transtornos psicológicos, da apreensão de expressões faciais de outrem por fotografia, a análise etológica destas expressões em animais como cães e gatos e, o mais importante para nós, um estudo global para investigar a variação cultural das relação entre expressões e emoções (SILVEIRA, 2010, p. 87–90). Ao enviar 16 cartas ao redor do mundo para esta investigação, Darwin tem como objetivo observar a capacidade do reconhecimento facial e sua associação com emoções específicas a fim de estabelecer sua manifestação como inata ou dependente da cultura. As respostas foram praticamente unânimes: não apenas a execução como o reconhecimento fisiológico de emoções é universal, como é natural. Ou seja, “do mesmo modo que expressamos sentimentos de maneira instintiva e inata, nosso reconhecimento dos mesmos antecede a qualquer julgamento ou reflexão da razão” (SILVEIRA, 2010, p. 88). Emoções como tristeza, raiva, medo ou surpresa não apenas independem de normas sociais ou culturas, como também estão presentes em outras espécies (ROXO et al., 2011, p. 2431). Após extensivas conjecturas comparativas entre o comportamento e faculdades humanas e de outros animais, Darwin conclui que

a diferença da mente entre o homem e os animais superiores, por maior que seja, certamente é uma de grau e não de tipo. Temos visto que os sentidos e intuições, as várias emoções e faculdades tais como amor, memória, atenção curiosidade, imitação, razão etc., das quais o homem se vangloria, podem ser encontradas de maneira incipiente, ou mesmo em uma condição bem desenvolvida, nos animais inferiores.<sup>75</sup> (DARWIN, 1889, p. 126, tradução nossa)

---

<sup>75</sup> No original: “*the difference in mind between man and the higher animals, great as it is, certainly is one of degree and not of kind. We have seen that the senses and intuitions, the various emotions and faculties, such as love, memory, attention, curiosity, imitation, reason, &c., of which*

Em suma, temos, portanto, um comportamento selecionado naturalmente e sexualmente através de constante convívio social em grupos de crescente tamanho em uma escala de tempo de até centenas de milhares de anos. Tal convívio favoreceu o exercício de nossos mecanismos cognitivos e comportamentais para a capacidade de execução e identificação de expressões emocionais a fim de relações sociais cada vez mais complexas. Ou seja, poucos exemplos da conexão direta entre nossa natureza genética e fisiológica com nosso ambiente e criação são mais evidentes do que nossa evolução social. Iremos desenvolver alguns destes conceitos originalmente apresentados por Darwin acerca da caracterização, expressão e reconhecimento de emoções nos capítulos vindouros, todavia não fugirá muito deste escopo geral apresentado aqui.

Evidentemente, esta apresentação não é esgotadora do montante de informações e estrutura que a biologia evolutiva pode nos prover. Mas nos é suficiente a título de ilustração de uma das colunas científicas centrais para a compreensão do agente dentro de uma perspectiva naturalista: somos animais como quaisquer outros, com capacidades e limites de acordo com nosso histórico evolutivo e outros fatores muitas vezes aleatórios. De acordo com Leda Cosmides e John Tooby, uma compreensão biológica evolutiva de nosso equipamento cognitivo permite, portanto, uma intersecção com a economia e psicologia acerca do processo de tomada de decisão.

Em seu artigo “Melhor que racional: Psicologia Evolutiva e a Mão Invisível”<sup>76</sup> (1994), resumem esta relação da seguinte maneira: a seleção natural na biologia seria análoga a um processo de “mão invisível” na economia, e é unicamente responsável no processo evolutivo por desenvolver estruturas fisiológicas complexas em organismos como o olho dos vertebrados e a capacidade de tomar decisões. A aparelhagem cognitiva gerada por este processo, responsável por nosso comportamento, está imbuída, portanto, em teorias do comportamento econômico – assim, as características desta aparelhagem definem e constituem os princípios universais humanos que guiam

---

*man boasts, may be found in an incipient, or even sometimes in a well-developed condition, in the lower animals”.*

<sup>76</sup> No original: “*Better than Rational: Evolutionary Psychology and the Invisible Hand*”.

a tomada de decisão econômica (COSMIDES; TOOBY, 1994, p. 328). Iremos nos aprofundar nesta perspectiva na próxima seção por meio de quem primeiramente possibilitou-a dentro da própria economia, Herbert Simon e sua concepção de razão limitada; além de compreender sua relevância contextual na história da economia e as ferramentas importantes que trouxe para todas as outras ciências que tratamos neste trabalho quanto ao problema da tomada de decisão.

## 4.2 Razão Limitada e Economia Comportamental

The classical theory does not tolerate the incomparability of oranges and apples.  
(SIMON, 1957, p. 108)

Na economia, foi tradição predominante tratar o agente como um ser puramente racional tal qual na tradição filosófica, no que se denomina a “teoria da escolha racional” até o século XX. De acordo com Mlodinow, “os fundamentos da economia clássica se esteiam na ideia de que as pessoas [...] comportam-se de maneira racional, de acordo com o princípio orientador de seus próprios interesses” (MLODINOW, 2013, p. 28). Este modelo postula um indivíduo com uma racionalidade perfeita,

um agente hipotético que possui informação completa sobre as opções disponíveis para a escolha, perfeita antecipação das consequências de escolher tais opções, e os meios para resolver um problema de otimização (tipicamente de considerável complexidade) que identifica uma opção que maximiza a utilidade pessoal do agente.<sup>77</sup> (WHEELER, 2020, p. 2, tradução nossa)

Este indivíduo racional hipotético é chamado comumente de *homo economicus*, um indivíduo “que, no curso de ser ‘econômico’ também é ‘racional’”<sup>78</sup> (SIMON, 1957, p. 99, tradução nossa). Esta perspectiva de agente fora e ainda é utilizada amplamente na teoria dos jogos, e representa a manifestação da concepção de cognitividade humana que se possui nestas tradições, originalmente proposta

<sup>77</sup> No original: “a hypothetical agent who has complete information about the options available for choice, perfect foresight of the consequences from choosing those options, and the wherewithal to solve an optimization problem (typically of considerable complexity) that identifies an option which maximizes the agent’s personal utility”.

<sup>78</sup> No original: “who, in the course of being ‘economic’ is also ‘rational’”.

por John Stuart Mill na primeira metade do século XIX como parte de sua teoria concernente ao raciocínio para o cálculo de utilidade de sua teoria utilitarista.

Segundo Simon, nesta perspectiva o agente é capaz de calcular razões de maneira objetiva, analisando probabilidades analogamente a um processo inferencial lógico-matemático, onde o resultado de sua deliberação seria uma conclusão necessária de seu conhecimento pleno das premissas, estatística e das normas de processamento lógico – ou uma “omnisciência lógica”, tanto do agente como de alguma interface computacional envolvida (SIMON, 1957, p. 102–103; WHEELER, 2020, p. 4–6, 11–12). A racionalidade na economia clássica então, pode ser resumida como “a otimização (maximização ou minimalização) de alguma função”<sup>79</sup> (GIGERENZER; SELTEN, 2002, p. 3, tradução nossa).

Todavia, ao invés de focar nos processos de raciocínio e no aspectos comportamentais envolvidos neles para a predição da escolha, centralizava-se neste cálculo de maximizar a utilidade, uma pressuposição que postulava que o “comportamento poderia ser previsto sem preocupar-se com os processos decisivos do ator econômico”<sup>80</sup> (SIMON, 2000, p. 27, tradução nossa) e que então “sucessos significativos eram de maneira última alcançados ao formalizar a teoria do equilíbrio geral em um alto nível de sofisticação matemática”<sup>81</sup> (SIMON, 2000, p. 27, tradução nossa), tanto formalmente em linguagem lógica quanto figurativamente em linguagem não-matemática. Isto é o cerne do utilitarismo clássico: predominantemente, prazeres e dores poderiam ser quantificados objetivamente, e, portanto, o cálculo de utilidade para a melhor ação também poderia sê-lo. Posteriormente, este paradigma tornou-se também a base para o pensamento de teorias economistas. Como veremos a seguir, apenas mais recentemente que este paradigma se desenvolveu do foco no processo de raciocínio para a escolha em si, e, por conseguinte, no processo de tomada de decisão.

Podemos rastrear o empreendimento de reconhecer o processo deliberativo como localizado entre indivíduos auto-interessados sem um

---

<sup>79</sup> No original: “*the optimization (maximization or minimization) of some function*”.

<sup>80</sup> No original: “*behavior could be predicted without concern for the decision processes of the economic actor*”.

<sup>81</sup> No original: “*major successes were ultimately achieved in formalizing general equilibrium theory at a high level of mathematical sophistication*”.

atrelamento inerente à racionalidade última maior até Adam Smith em sua “Riqueza das Nações” (2007), escrita no século XVIII, onde ele desenvolve uma perspectiva humeana da ação baseada numa concepção empiricista de simpatia (ou empatia, em termos mais contemporâneos).

A crença que indivíduos, como agentes livres e racionais, são capazes de ir além da busca exclusiva de seu respectivo bem-estar foi fortemente asserida por autores clássicos como Adam Smith [quem] em particular, distinguia ações motivadas pela “prudência” (levando em consideração vantagens diretas e indiretas) daquelas influenciadas pela “simpatia”, “generosidade” e “espírito público”, para cada as quais ele deu papéis distintos e diferenciados.<sup>82</sup> (SEN, 1994, p. 386, tradução nossa)

Smith não fala em momento algum em racionalidade como parte de sua teoria, e o termo “racional” é apenas utilizado duas vezes para qualificar outros conceitos que não tangem o processo deliberativo (SIMON, 2000, p. 26–27; SMITH, 2007, p. 506, 513). Assim, ao se referir à agentes em ação, Smith usa uma terminologia subjetiva, imprecisa, denotando as particularidades individuais ao tomar escolhas:

seres humanos utilizam sua perspicácia para ganhar o que querem, mas Adam Smith não faz afirmações que podem otimizar ou maximizar, e ele reconhece a variedade dos desejos humanos sem agrupar todos estes desejos em uma função de utilidade comum.<sup>83</sup> (SIMON, 2000, p. 27, tradução nossa)

Este paradigma foi relevante para alguns de seus contemporâneos e sucessores, todavia caiu em desuso após a ascensão do idealismo alemão, retornando novamente à atenção dos economistas na década de 1950 com Herbert Simon e a introdução do conceito de “*bounded rationality*” (GIGERENZER; SELTEN, 2002, p. 1; SIMON, 1957; WHEELER, 2020), traduzido usualmente como “racionalidade limitada”, mas que literalmente traduz-se por “racionalidade atada”, em referência aos elementos que impedem

---

<sup>82</sup> No original: “*The belief that individuals, as free and rational agents, are capable of going beyond the exclusive pursuit of their respective wellbeing was strongly asserted by classical authors such as Adam Smith [who] in particular, distinguished actions motivated by "prudence" (taking into account indirect as well as direct advantages) from those influenced by "sympathy," "generosity," and "public spirit," to each of which he gave a distinct and differentiated role*”.

<sup>83</sup> No original: “*[h]uman beings use their wits to gain what they want, but Adam Smith makes no claims that they optimize or maximize, and he recognizes the variety of human desires without crowding all of these desires into a common utility function*”.

que a razão atue de maneira autônoma e independente ao realizar uma escolha. O conceito serve como “uma abreviação de seu dossiê contr economia neoclássica e seu chamado para substituir as presunções da racionalidade perfeita do *homo economicus* com uma concepção de racionalidade feita sob medida para agentes cognitivamente limitados”<sup>84</sup> (WHEELER, 2020, p. 1 grifo do autor, tradução nossa), dando importância também portanto à totalidade do contexto em que a escolha é realizada.

Esta perspectiva, então, retorna ao foco no comportamento relativo ao processo do raciocínio como um todo, e não apenas da escolha em si como tornou-se padrão. Assim, reconhece-se que o indivíduo, ao realizar sua escolha, não somente é um sujeito auto-interessado dotado de capacidade de escolha, como também que é um animal social cujos poderes deliberativos podem ser afetados por fatores externos aos diretamente envolvidos na decisão em si, ou seja: “o estado da informação pode muito bem ser observado como uma característica de quem toma a decisão como uma característica de seu ambiente”<sup>85</sup> (SIMON, 1957, p. 100).

Desta maneira, a tomada de decisão é como o ato de cortar utilizando

um par de tesouras, onde uma lâmina representa os “limites cognitivos” de humanos de verdade e a outra a “estrutura do ambiente”. Mentes com tempo, conhecimento e outros recursos limitados podem ainda assim serem bem-sucedidos ao explorar estruturas em seus ambientes.<sup>86</sup> (GIGERENZER; SELTEN, 2002, p. 4, tradução nossa)

Ou seja, é invariável que ambos estes fatores sejam devidamente analisados para podermos entender o processo deliberativo: “quando a racionalidade é associada com processos de raciocínio, e não apenas com seus produtos, os limites das habilidades dos *Homo sapiens* para raciocinar não podem ser ignorados”<sup>87</sup> (SIMON, 2000, p. 26, tradução nossa). Assim, ao invés de negar a

---

<sup>84</sup> No original: “a shorthand for his brief against neoclassical economics and his call to replace the perfect rationality assumptions of *homo economicus* with a conception of rationality tailored to cognitively limited agents”.

<sup>85</sup> No original: “the state of information may as well be regarded as a characteristic of the decision-maker as a characteristic of his environment.”

<sup>86</sup> No original: “a pair of scissors, where one blade is the “cognitive limitations” of actual humans and the other the “structure of the environment.” Minds with limited time, knowledge, and other resources can be nevertheless successful by exploiting structures in their environments”.

<sup>87</sup> No original: “[w]hen rationality is associated with reasoning processes, and not just with its products, limits on the abilities of *Homo sapiens* to reason cannot be ignored”.

racionalidade como um todo, Simon propõe aproximar os estudos da tomada de decisão das evidências psicológicas que tangem o tópico (1957, p. 99–100) – uma tarefa que expandiu-se à uma miríade de áreas diferentes do conhecimento como dito anteriormente, da biologia e inteligência artificial à antropologia e à própria filosofia (GIGERENZER; SELTEN, 2002, p. 1) – para demonstrar empiricamente que os pressupostos da economia clássica racionalista não possuem embasamento científico algum acerca das reais capacidades deliberativas dos indivíduos sobre os quais ela pretende trabalhar (SIMON, 1957, p. 104).

Em suma, “a racionalidade humana real pode ser no máximo uma aproximação extremamente grosseira e simplificada do tipo de racionalidade global que é implicada, por exemplo, por modelos da teoria dos jogos”<sup>88</sup> (SIMON, 1957, p. 101, tradução nossa). Isto nos remete ao paralelo realizado por Tooby e Cosmides na seção anterior quanto ao paradigma evolutivo do agente; fica claro, agora, que “a mão invisível da seleção natural criou a estrutura da mente humana, e a interação destas mentes é o que gera a mão invisível da economia”<sup>89</sup> (COSMIDES; TOOBY, 1994, p. 328, tradução nossa). Assim, Simon estabelece este novo paradigma em que “agentes econômicos são vistos como limitadamente racionais, ou seja, pensa-se que agem sensivelmente dentro do contexto da tomada de decisão e dentro dos limites de suas habilidades cognitivas e emocionais”<sup>90</sup> (DE BONDT, 2008, p. 68).

Simon introduz, então, novas variáveis ao cálculo deliberativo clássico baseado nestas possíveis limitações da racionalidade. À variável tradicional que delineia todas as alternativas possíveis que o organismo (agência computacional tanto individual quanto digital) tem disponível, acrescenta-se a variável de “alternativas *consideráveis*”, ou seja, aquelas que em sua limitação perceptiva considera em detrimentos de outras que possivelmente tenha ignorado (SIMON, 1957, p. 102). Ao conceito de “*pay-off*”<sup>91</sup>, análogo à concepção de “risco”

---

<sup>88</sup> No original: “*actual human rationality can at best be an extremely crude and simplified approximation to the kind of global rationality that is implied, for example, by game-theoretical models*”.

<sup>89</sup> No original: “*natural selection's invisible hand created the structure of the human mind, and the interaction of these minds is what generates the invisible hand of economics*”.

<sup>90</sup> No original: “[e]conomic agents are seen as boundedly rational, i.e., they are thought to act sensibly within the decision-making context and within the limits of their cognitive and emotional abilities”

<sup>91</sup> Pode ser traduzido como “recompensa”, “resultado”.

predominante na economia que aluz à probabilidade da consequência de uma decisão desviar do objetivo intencionado, ele acrescenta o elemento psicológico de que existem determinadas situações e sequências de eventos que podem alterar nossa percepção do que é risco e do que seria um objetivo desejável, dificultando um modelo racional que antecipasse tais contextos (SIMON, 1957, p. 104–106, 108–110).

Afinal, no modelo econômico clássico, pressupõe-se que o organismo tenha esta ciência completa de todas as alternativas e riscos envolvidos em cada escolha, mas o reconhecimento destas limitações sobre as alternativas consideráveis e da dificuldade no cálculo de recompensas, verifica-se que o mecanismo cognitivo por trás da escolha pode também ser desconhecido. Assim, fatores como a primeira alternativa viável que encontrarmos ter maior preferência sobre outras possibilidades igualmente viáveis, ou o nível da “aspiração” do indivíduo de escolher algum curso de ação de acordo com a facilidade ou dificuldade de encontrar diferentes alternativas, onde também sua persistência pode influenciar, acabam entrando no cálculo deliberativo de tal maneira que o sistema clássico não dá conta. Portanto, o modelo de Simon procura superar o conceito tradicional de “racionalidade global”, esta capacidade universal e constante que historicamente se atribuía a nossas capacidades cognitivas, para esta racionalidade limitada, localizada em indivíduos com diferentes critérios hierárquicos para a tomada de decisão (SIMON, 1957, p. 110–113).

O que Simon e seus herdeiros trouxeram à luz é que, ao contrário do que os economistas tradicionalmente associavam com no máximo pequenos desvios ou correções posteriores à uma teoria da decisão racional predominante, os aspectos psicológicos, biológicos e afins não podem ser tolhidos de uma análise completa relativa ao processo deliberativo, uma vez que compõe nele os principais fatores decisivos. Assim, a expectativa tradicional da maximização da utilidade esperada da racionalidade “pragmática” conforme Simon define crescentemente cede seu lugar ao montante de evidências trazidas pela psicologia cognitiva e outras ciências, favorecendo então o estudo da incerteza na economia sobre os cálculos supostamente objetivos da tradição (SIMON, 2000, p. 34–38).

Iremos adentrar mais profundamente em pesquisas e evidências que corroboram a proposta de Simon posteriormente, mas gostaria de citar duas aqui brevemente a fim de ilustrar algumas consequências de sua revolução na economia. Embora, como todas as outras que trataremos posteriormente, estas teorias adentrem bastante na área de psicologia cognitiva, ambas se localizam mais na área geral de economia, e especificamente nesta veia comportamental, do que as outras.

Primeiramente, sobre um fenômeno bastante interessante acerca da elaboração de preços para produtos que provavelmente todos nós já vimos e possivelmente no qual caímos também. É comum vermos preços como “R\$1,99” ao invés do valor inteiro de “R\$2,00”, e o motivo é a utilização do conceito de “preço psicológico”<sup>92</sup>. Embora muitos entendam e percebam o motivo de seu uso, as evidências apontam claramente que ainda existe um padrão forte de avaliação errônea para o número mais baixo, passando a ilusão para o agente comprador de que o custo do produto é menor do que de fato o é. Esta estratégia específica de diminuir um centavo de um valor para que o dígito da esquerda seja menor do que mantendo tal centavo chama-se “efeito de dígito à esquerda”<sup>93</sup> e, embora não seja universal, de fato mantém-se relevante em diversos estudos (BIZER; SCHINDLER, 2005; MANNING; SPROTT; ARTICLE., 2009).

Por exemplo, ao serem fornecidos com uma quantia determinada de dinheiro em dólar, participantes que viram produtos com o efeito de dígito à esquerda tinham a ilusão de que possuíam poder de compra muito maior do que os que foram expostos a quantias inteiras, como se de fato o um centavo a menos implicasse em um dólar inteiro a menos (BIZER; SCHINDLER, 2005, p. 776–782). Muitas hipóteses surgiram para tentar dar conta deste fenômeno tão comum, todavia em suma é suficiente para nós sua mera existência para estabelecer um primeiro caso onde a teoria da racionalidade limitada claramente se sobressai em comparação com a teoria clássica: claramente, esta última não daria conta de explicar uma mudança tão significativa na percepção de pessoas se elas analisassem objetivamente a situação, já que meros centavos não fariam nenhuma diferença prática.

---

<sup>92</sup> “*Psychological pricing*” no original.

<sup>93</sup> “*Left-digit effect*” no original.

Outro caso é o da teoria da arquitetura, conhecida na língua inglesa originalmente como teoria do “nudge”, ou do “empurrãozinho”, da “cutucada”, numa tradução aproximada. Proposta por Richard Thaler e Cass Sunstein em 2008, a teoria entra em determinados méritos acerca da deliberação o qual não tomaremos neste texto, por fugirem do escopo do trabalho; ainda assim, trazem bastantes dados interessantes sobre o processo de escolha. Como o nome sugere, a teoria prevê que mesmo as menores mudanças no ambiente, ou as menores influências no agente podem trazer significativas alterações em seu comportamento. Sua obra inicia com um exemplo simples acerca da reorganização da disposição de alimentos em diversos refeitórios de escolas em um distrito. Dependendo da maneira em que a comida é disposta, e. g. se é a primeira ou última que a pessoa vê, se está na altura dos olhos, acima ou abaixo – todos estes fatores influenciam de uma maneira ou de outra como os indivíduos escolheram determinada comida em detrimento de outra (THALER; SUNSTEIN, 2008, p. 1–4).

Assim, os autores identificam que o cotidiano de todos nos expõe intencionalmente ou não a estas “empurradinhas”, do acaso da disposição de alimentos em sua própria geladeira a vendedores tentando influenciar sua compra como vimos com o artifício do preço psicológico. Estes que organizam situações como essas são chamados, então, de “arquitetos da escolha”, embora não necessariamente realizem esta influência de maneira intencional, como o caso de quem dispõe a comida na cantina sem ordem definida – mas que acaba influenciando invariavelmente a escolha dos agentes.

Os autores versam sobre diversos tópicos desde finanças até saúde, meio ambiente e o próprio conceito de liberdade em si. Por exemplo, falam acerca da diminuição do uso de energia em 40% em residências que possuíam uma esfera que brilhava verde quando o uso era modesto e vermelho quando exagerado – com algumas inclusive tocando sons desagradáveis neste caso –, algo que mera demonstração na conta, avisos e-mails e outros métodos mais sutis não foram capazes de realizar. Assim, sugerem também estratégias governamentais para lutar contra a mudança climática em nível global com gastos menores, uma vez que alguns outros experimentos semelhantes em mais larga escala tenham sido bem-sucedidos (THALER; SUNSTEIN, 2008, p. 193–196) como por exemplo com doação de órgãos, assim como negativamente, como com sistemas de

planos de saúde dada sua complexidade (THALER; SUNSTEIN, 2008, cap. 10, 11).

Retornando à comparação com a teoria clássica da economia, os autores reforçam que “[e]m situações complexas, o mantra de Apenas Maximizar Escolhas não é o suficiente para criar boas políticas. Quanto mais opções existem, e mais complexa a situação, mais importante é se ter uma arquitetura de escolha iluminada”<sup>94</sup> (THALER; SUNSTEIN, 2008, p. 174, tradução nossa). Toda a base da teoria da arquitetura é indiretamente estabelecida na revolução da racionalidade limitada que Simon trouxe à economia comportamental, assim como seu constante diálogo com a psicologia cognitiva.

De maneira geral, nas próximas seções iremos analisar de modo similar como estudos sobre a deliberação avançaram na psicologia cognitiva e nas neurociências, qual aproximadamente é o estado-da-arte na pesquisa que interessa a este trabalho e também alguns conceitos-chave que serão fundamentais para a compreensão das pesquisas referenciadas posteriormente e para a discussão acerca de sua relevância filosófica. Afinal, assim como a economia comportamental surge em diálogo direto com um entendimento evolutivo do agente humano, invariavelmente as outras ciências que tratam de questões semelhantes envolvidas na tomada de decisão também começam a se interseccionar. A psicologia cognitiva tratará de maneira geral dos mecanismos cognitivos envolvidos neste processo, e as neurociências mais especificamente ainda através das análises fisiológicas do que é nossa “caixa preta” comportamental, o sistema nervoso (CAMERER; LOEWENSTEIN; PRELEC, 2005, p. 9). Iremos voltar especialmente à economia comportamental e estes preceitos da teoria da racionalidade limitada especialmente quando tratarmos de certos vieses na psicologia cognitiva como aversão à perda e efeito de enquadramento, dentre outros.

Um dos campos em que estas três últimas ciências dialogam mais veementemente é o da neuroeconomia, a qual pretende investigar estes conceitos que vimos nesta seção dentro de um escopo fisiológico-empírico. As neurociências são capazes de informar a economia comportamental de duas

---

<sup>94</sup> No original: “[i]n complex situations, the Just Maximize Choices mantra is not enough to create good policy. The more choices there are, and the more complex the situation, the more important it is to have enlightened choice architecture”

maneiras principais: incrementalmente e radicalmente. De maneira incremental ela funciona introduzindo ou alterando variáveis clássicas na tomada de decisão com base empírica, como no caso que o consumo de drogas inibe o prazer do consumo futuro de outros bens. De maneira radical, por outro lado, potencializa a teoria da racionalidade restringida de Simon deixando explícito que as teorias clássicas não dão conta dos processos automáticos e afetivos e possuem altos graus de racionalização *post hoc* destes mesmos a fim de manter uma consistência estrutural em suas teorias (CAMERER; LOEWENSTEIN; PRELEC, 2005, p. 10–11). Veremos mais sobre estes processos automáticos e afetivos além de uma apresentação geral das neurociências na seção 4.4, e mais dados e discussões das relações entre estas áreas nos próximos capítulos.

### **4.3 Psicologia Cognitiva e Vieses**

Ao tratarmos do entendimento de como o processo de tomada de decisão funciona e de que maneiras é afetado, torna-se imprescindível a compreensão de seu estudo na psicologia. Por isso, faremos, em um primeiro momento, um breve apanhado histórico a fim de localizar a psicologia cognitiva como linha principal para nossos estudos acerca das faculdades deliberativas, para, em seguida, analisar o desenvolvimento conceitual de “dissonância cognitiva” e o estudo geral de vieses cognitivos que surgiram posteriormente a estas discussões. Encerraremos com um diálogo entre o entendimento destes conceitos trabalhados até então com a teoria do processo duplo, paradigma de análise na psicologia concernente à influência destes vieses na cognitividade que surgiu em intersecções com a economia comportamental em sua versão posterior à revolução de Herbert Simon como vista na seção anterior. Isto nos dará as bases, então, para seguirmos no próximo capítulo com os desenvolvimentos das neurociências acerca destes tópicos mencionados, estruturando então a base científica de nossa investigação filosófica.

#### *4.3.1 Psicologia e cognição*

A psicologia nasceu de uma cisão com a tradição filosófica da filosofia da mente e epistemologia para um viés prático e científico da análise da mente e do comportamento humano; portanto, diferentemente de outras áreas do conhecimento, sempre andou de mãos dadas com a metodologia empírica de

pesquisa. Em seu berço, ao fim do século XIX, Wilhelm Wundt, fisiólogo, já tentava tratar dos problemas “mentais” – hoje chamados de “cognitivos” – sob um viés experimental, tornando-se o primeiro a se autodenominar “psicólogo” e estabelecendo a estrutura do que iria futuramente tornar-se a psicologia experimental<sup>95</sup>. A partir de um contexto filosófico em diálogo com William James, Wundt já tentava explicitar “a falta de fronteiras distintas entre as várias disciplinas que combinavam-se para o surgimento do desenvolvimento da Psicologia”<sup>96</sup> (GROSS, 2020, p. 39). Ainda assim, a ciência teve muito a se desenvolver antes de alcançar o crivo contemporâneo em sua aplicação, e da mesma maneira levou tempo até que grandes conjecturas especulativas na psicologia dessem lugar para a análise empírica adequada (SOLSO; MACLIN; MACLIN, 2014, p. 108).

Da mesma maneira que a psicologia como área separou-se da filosofia a fim de maior especialização sobre a psique, a psicologia cognitiva surge de uma quebra com o behaviorismo, que era na época a linha de pesquisa mais preocupada com o aspecto prático e testável de hipóteses comportamentais das três principais, como o próprio nome sugere. O behaviorismo, também conhecido como comportamentalismo ou psicologia comportamental pela tradução de seu termo original “*behavior*”, “comportamento” em inglês, mantinha que apenas uma relação prática ocorria no processo comportamental: o estímulo-resposta. Os teóricos abstraíam qualquer conteúdo mental de uma ação uma vez que o estado psicológico poderia ser analisado puramente através da observação de como um agente responderia a um estímulo específico; e assim conceitos como memória, pensamento e outras que pressupunham conteúdo interno foram abandonadas até a “revolução cognitiva” na década de 30. De fato, o próprio Skinner, que cunhou o termo a partir das pesquisas de John Watson e Ivan Pavlov, afirma que “[o] Behaviorismo não é a ciência do comportamento humano, mas, sim, a filosofia dessa ciência” (SKINNER, 2004, p. 7), denotando – ironicamente, em nosso contexto –, o foco mais abstrato desta escola. Mesmo Sigmund Freud, um dos fundadores da psicologia, já propunha uma dicotomia entre o inconsciente e

---

<sup>95</sup> Há controvérsias acerca da exatidão destas informações acerca de seu pioneirismo, muitas delas auto-afirmadas por Wundt, que não vem ao caso. Para mais informações, cf. GROSS, 2020, p.40.

<sup>96</sup> No original: “*the lack of distinct boundaries between the various disciplines that combined to bring about Psychology’s development*”.

consciente – algo que a princípio seria bastante relevante para a discussão deste trabalho –, todavia, como reconhecem os behavioristas, a teoria freudiana carecia de objetividade e crivo empírico adequado (SOLSO; MACLIN; MACLIN, 2014, p. 108). Assim, mesmo sendo confirmados posteriormente, como veremos, iremos focar somente nestas pesquisas dado nosso foco experimental.

Tal “revolução cognitiva” foi trazida por Edward Tolman em 1932 analisando o comportamento de ratos em labirintos engenhosamente gerados de maneira a alterar os caminhos mantendo o local geral relativo da comida, de tal forma que o paradigma estímulo-resposta cairia por terra caso soubessem encontrá-la considerando sua posição relativa ao invés dos movimentos específicos aprendidos previamente. Este “mapa cognitivo” que foi atribuído aos roedores tornou-se uma hipótese muito mais adequada para o experimento do que quaisquer tentativas behavioristas puderam dar conta. Isto continuou ocorrendo com vários experimentos do gênero nos próximos 20 anos, contando com a ascensão das ciências cognitivas como um todo na área da informática, antropologia, na própria filosofia com autores como Chomsky e Fodor e também na economia com os desenvolvimentos de Simon. Pouco a pouco, a imagem interdisciplinar do intento deste trabalho torna-se mais clara. Assim, os últimos pregos no caixão do behaviorismo foram martelados, e esta subárea da psicologia especializada na aquisição e processamento de informações foi estabelecida: a psicologia cognitiva (GROSS, 2020, p. 41–42; SOLSO; MACLIN; MACLIN, 2014, p. 21–26).

Não há, como nas três principais linhas prévias de pesquisa na psicologia – a saber, psicanálise, behaviorismo e *gestalt* –, teorias ou autores principais representativos na psicologia cognitiva. Mas o aspecto computacional, como vimos pela com Simon anteriormente, torna-se evidente uma vez que a fisicalidade de nossas capacidades cognitivas toma o centro das atenções nos debates acerca da atenção, memória e processamento em geral de estímulos e geração de comportamento baseado nestes processos (GROSS, 2020, p. 23–26). Todavia, não adentraremos nesse aspecto computacional nem de puro processamento de informações da psicologia cognitiva – duas de suas áreas principais – mas sim em seu aspecto neurocientífico, como veremos na seção correspondente, enquanto neurociência cognitiva. Isto pois tais áreas focam na investigação direta destes processos supracitados de maneira abstrata,

enquanto aqui a preocupação é centrada nas funções fisiológicas que subjazem tais processos, ou seja, o que “limita” a racionalidade conforme identificado por Simon. (SOLSO; MACLIN; MACLIN, 2014, p. 30–36)

O estudo da psicologia cognitiva – como de todas áreas tratadas aqui – é amplo demais para uma apresentação adequada. Todavia, apenas alguns conceitos centrais serão importantes para este trabalho. Como estabelecemos, o funcionamento cognitivo é um resultado de nossos processos fisiológicos em interação com o ambiente, e assim sujeito à inúmeras variáveis. A mesma pessoa pode perceber e julgar algo diferentemente de acordo com iluminação, grau de fome, estímulos sonoros, odores etc., assim como pessoas diferentes na mesma situação. Iremos tratar diretamente de evidências para isto e suas implicações filosóficas nos próximos capítulos.

Mas talvez o aspecto mais importante do que já estabelecemos até aqui é que o ato de pensar não implica em um acesso abstrato à uma dimensão pura de ideias, mas sim na geração de representações a partir de nossas limitadas capacidades perceptivas, cognitivas e comportamentais. Conceitos não são objetos metafísicos acessíveis pela razão, mas sim um acúmulo de diferentes estímulos armazenados na memória que compartilham algo em comum. O conceito de cadeira, por exemplo, pressupõe visões de diferentes cadeiras com nossos olhos em diferentes condições de luz, seu toque, a palavra que a representa e o som associado, os contextos nos quais interagimos com objetos que compartilham estas características em comum e como lidamos e observamos outros fazê-lo com os mesmos, etc. O conceito de *hippie*, por sua vez, pode envolver a memória de pessoas, vestimentas, símbolos, acontecimentos históricos, eventos, posicionamentos políticos e determinados comportamentos – tanto como observador como participante. Todas estas informações estão em algum grau armazenadas em nossa memória de longo prazo após estas experiências.

Assim, não apenas conceitos, mas a escolha como um todo depende integralmente de nossas disposições, interações e contextos cognitivos e afetivos envolvidos. Conceitos, então, seriam um dos “guintastes” como visto anteriormente a que Dennett se refere: há uma simplificação da multitude de diferentes aspectos cognitivos através de sua associação. Uma vez que um fator importante para a seleção natural é a administração de recursos a fim de

dispender menor carga energética<sup>97</sup> em menor tempo, tal otimização é funcional. Mas da mesma maneira, como veremos a seguir, estas generalizações pragmáticas não apenas podem como não é raro levarem à falha.

Iremos nos adentrar nos conceitos principais os quais a psicologia cognitiva trata, como percepção, memória e atenção, por exemplo, quando formos focar em seu aspecto neurológico posteriormente uma vez que também estabeleceremos as áreas do sistema nervoso mais associadas com seu funcionamento. Por enquanto, é suficiente entender que o pensamento e a ação são processos baseados num entendimento de razão constringida em um entendimento evolutivamente estável, ou seja, sujeito a alta variação de acordo com o contexto. Na seguinte seção trataremos de vieses cognitivos e como estes afetam todo o escopo relacionado a nossas tomadas de decisão.

#### *4.3.2 Dissonância e outros vieses cognitivos*

Uma vez estabelecida enquanto ciência prática focada em nossas faculdades cognitivas e em como estas interagem com o ambiente através de nosso comportamento, a psicologia cognitiva parte para entender o fenômeno deliberativo como um todo. Iniciaremos esta seção introduzindo o conceito de “dissonância cognitiva” como desenvolvida na década de 1950 na psicologia social, e posteriormente seu impacto na análise psicológica e os outros vieses descobertos em pesquisas mais recentes. Estes conectam-se diretamente com os limites de nossa racionalidade que Simon referiu-se, uma vez que se tratam de mecanismos automáticos para facilitar o processamento cognitivo que, todavia, acabam por comprometê-lo de um funcionamento “perfeito” como nas pressuposições tradicionais da economia, filosofia e outras ciências comportamentais.

Em primeiro lugar, um “viés” cognitivo refere-se a como estaríamos inclinados (enviesados) a agir de uma determinada maneira em contraste a outras em determinadas situações. Uma maneira simples de exemplificar é quando estamos procurando alguém conhecido numa multidão: ignoramos quase todas as pessoas que não se adequam à descrição da pessoa em questão a fim de otimizar nosso tempo de procura. Ou seja, intencionalmente nos

---

<sup>97</sup> Obviamente, dado o contexto, em termos puramente calóricos e não-metafísicos.

enviesamos, nos dispomos a deliberadamente limitar nossa percepção, no caso, a fim de atingir um objetivo de maneira eficiente. Todavia, podemos não perceber a pessoa mesmo vendo-a diretamente se por acaso não corresponder ao critério que escolhemos: pode ter modificado o cabelo, usar roupas diferentes do usual, ou a iluminação do local pode fazê-la parecer diferente, dentre outros inúmeros fatores. Este exemplo nos é útil não apenas para ilustrar como vieses funcionam, como também que não necessariamente são negativos. Novamente, a economia de energia é evolutivamente estimulada a fim de poupar recursos tomando atalhos cognitivos, o que facilita seu desenvolvimento; mas da mesma maneira, podem nos levar a resultados indesejados.

Uma das primeiras e mais importantes ferramentas conceituais para o desenvolvimento do estudo de vieses cognitivos foi a dissonância cognitiva. Tal fora introduzida primariamente, como mencionado no primeiro capítulo, por Leon Festinger em sua obra “Quando a Profecia Falha”<sup>98</sup> (2008; SHARMA, 2014, p. 837), que trata do acompanhamento de um grupo de cultistas teóricos da conspiração que acreditavam que alienígenas iriam libertá-los da terra no dia do juízo final – todavia, este dia já havia chegado e posteriormente alterado incontáveis vezes. Cada uma dessas vezes em que o fim do mundo não chegava, os líderes e seguidores criavam posteriormente uma explicação, ou uma “racionalização”, a fim de conciliar a sua fé com o fato que o mundo ainda existia. Assim, o conceito “é centrado na volta da ideia que se uma pessoa sabe várias coisas que não são psicologicamente consistentes umas com as outras, ele irá, em uma variedade de maneiras, tentar fazê-las mais consistentes”<sup>99</sup> (FESTINGER, 1962, p. 93).

Desta maneira,

[d]issonância e consonância são relações entre cognições – ou seja, entre opiniões, crenças, conhecimento do ambiente e conhecimento de suas próprias ações e sentimentos. Duas opiniões, ou crenças, ou itens de conhecimento são dissonantes entre si se não se encaixam juntas – ou seja, se são inconsistentes ou se, considerando apenas os dois

---

<sup>98</sup> No original: “When Prophecy Fails”.

<sup>99</sup> No original: “centers around the idea that if a person knows various things that are not psychologically consistent with one another, he will, in a variety of ways, try to make them more consistent”.

itens particulares, um não se segue do outro.<sup>100</sup> (FESTINGER et al., 2008, p. 33)

Como no caso dos adoradores de alienígenas, uma vez que sua crença na liturgia extraterrestre entrava em conflito com o fato evidente que o fim dos dias não havia chegado, ao invés de reconhecer que esta sua crença prévia estava equivocada, de fato ampliava mais ainda sua crença nos salvadores de outro mundo. Isto se dá pois torna-se mais confortável criar uma explicação que “justifique” os motivos de uma crença menor estar errada já dentro do conjunto de crenças prévias, ao invés de aceitar que ela faz parte de um sistema falho.

A explicação para este tipo de raciocínio recai sobre princípios biológicos básicos de economia de energia e aversão à dor: em termos de esforço mental e injúria social, é muito mais trabalhoso e mesmo danoso reconhecer tal sistema de crenças como um erro do que simplesmente corrigi-lo com alguma explicação confortável (FESTINGER et al., 2008, p. 33–37). E ainda mais, a cada evento falho, cresce em força a racionalização, uma vez que este comportamento já se tornou padrão em resposta a um evento em constante repetição. A dissonância então se dá nesta lacuna entre estas cognições, e resulta na maneira em que tentamos lidar com ela sem reformular completamente nossa visão de mundo.

E não é necessário ir para exemplos tão extremos como do apocalipse extraterrestre. A figura do fumante que sabe que o ato de fumar faz mal a seu corpo e ainda assim permite-se fazê-lo sem grandes restrições é comum, geralmente com racionalizações como que apaziguaria sua ansiedade ou substituiria outros vícios piores (cf. FESTINGER et al., 2008, p. 8–9). É preciso que algo fique claro: estas não são mentiras deliberadas. O agente, mesmo ciente do absurdo da contradição, acredita veementemente que sua racionalização é uma explicação razoável. Este aspecto é um dos pontos cruciais para esta tese. A auto-enganação é um mecanismo evolutivo não apenas simples como comum, e muito mais amplo que apenas um processo objetivo de racionalidade pura. Emoções, crenças, vícios químicos e a mera preferência por

---

<sup>100</sup> No original: “Dissonance and consonance are relations among cognitions – that is, among opinions, beliefs, knowledge of the environment, and knowledge of one’s own actions and feelings. Two opinions, or beliefs, or items of knowledge are dissonant with each other if they do not fit together – that is, if they are inconsistent, or if, considering only the particular two items, one does not follow from the other.”

menor esforço cognitivo, por exemplo, reforçam um panorama muito mais rico do processo deliberativo conforme temos visto.

A partir desta ampla gama de fatores extrínsecos ao elementos tradicionalmente primários em um processo de tomada de decisão nesta apresentação da dissonância cognitiva e também, anteriormente, da racionalidade limitada e um viés evolutivo como um todo, os psicólogos Amos Tversky, especialista em psicologia matemática e cognitiva, e Daniel Kahneman, da economia, desenvolveram ao início da década de 70 a concepção ampla de “viés cognitivo” (DANIEL KAHNEMAN; SHANE FREDERICK, 2002; KAHNEMAN; TVERSKY, 1972). Este conceito engloba uma miríade de áreas acerca de fatores que influenciam a percepção e o processamento de informações, sendo aplicável na economia, psicologia, marketing e diversas outras áreas.<sup>101</sup>

Seus primeiros estudos analisavam aspectos específicos como o fato de que “pessoas não seguem os princípios da teoria da probabilidade ao julgar a probabilidade de eventos incertos” (KAHNEMAN; TVERSKY, 1972, p. 431, tradução nossa)<sup>102</sup>, analisando mecanismos psicológicos de crença heurística, ou seja, de maneira análoga a este trabalho, procurar entender os processos que levam pessoas à rescindir de uma “racionalidade” probabilística em prol de intuições frequentemente errôneas. Os autores focam especificamente nos processos cognitivos envolvidos na formação de crença através da análise de probabilidades e predição de valores que levam à decisão (TVERSKY; KAHNEMAN, 2013, p. 419). Estes processos, como Simon pensara anteriormente, são heurísticos, mas não apenas em um sentido casuístico: dependendo de diversos fatores externos envolvidos no processo deliberativo, nós tendemos a nos apoiar em modelos “intuitivos”, em um sentido de mecanismos semi-automatizados como os modelos de psicologia evolutiva vieram a identificar.

Viéses cognitivos afetam diretamente nossa percepção da probabilidade de um determinado resultado. Um exemplo primário na economia comportamental é o da “falácia do apostador”, por exemplo, que envolve a

---

<sup>101</sup> Não à toa, fora uma pesquisa que rendeu o prêmio Nobel em Economia à Daniel Kahneman.

<sup>102</sup> No original: “*People do not follow the principles of probability theory in judging the likelihood of uncertain events*”.

intuição falsa de que as chances de acertar algo baseado na sorte aumentam a cada tentativa, quando na verdade as chances não se alteram. Ao se jogar uma moeda no ar, há invariavelmente 50% de chance dela cair em cara ou coroa, independentemente de quantas vezes a ação for repetida; a falácia está em se pensar que, por ter aterrissado com uma de suas faces para cima uma ou mais vezes, há uma chance maior de que caia na outra na próxima jogada. Afinal, tendo em mente que ela cai em 50% das vezes em algum dos dois lados, pode-se pensar que, em duas tentativas, é bastante provável que uma será cara e a outra coroa. Assim, se ela der cara na primeira vez, o apostador “intui” que há mais chances de dar coroa. Todavia, a moeda não tem “memória” dos últimos resultados, e os subsequentes não dependem dos anteriores de maneira alguma (TVERSKY; KAHNEMAN, 2013, p. 431).

Três principais processos cognitivos heurísticos são tratados inicialmente por Tversky e Kahneman em seus primeiros trabalhos: o da representatividade, o da disponibilidade e o da âncora de ajuste. O primeiro tange o erro probabilístico de associar a representatividade de um membro com sua categoria e deixando de lado as probabilidades em geral, optando pela opção com menor chance estatística se esta representar mais o membro como tal é percebido do que o grupo geral de que participa (KAHNEMAN; TVERSKY, 1972, p. 431–433).

O exemplo mais famoso é o de Linda, uma moça hipotética de 30 anos que seria formada em filosofia, extrovertida e preocupada com causas sociais. Ao serem apresentados à esta descrição geral, os participantes foram inquiridos para tentar adivinhar a profissão de Linda, e invariavelmente, entre as opções “Linda trabalha como caixa de banco” e “Linda trabalha como caixa de banco e é ativista do movimento feminista”,<sup>103</sup> a maioria escolhia a segunda opção. Todavia, basta uma breve reflexão para torne-se evidente que, apesar de condizer com o estereótipo proposto pelo exemplo, estatisticamente as chances de um membro participar de um grupo X (caixas de banco) serão invariavelmente menores do que de ela participar de um grupo X e de um grupo Y (ativista do movimento feminista) simultaneamente (TVERSKY; KAHNEMAN, 2013, cap. 15). Ou seja, mesmo quando as opções logicamente e estatisticamente mais prováveis estão presentes junto com outras ilógicas e consideravelmente menos

---

<sup>103</sup> Existiram diversas formulações deste exemplo, esta conclusão é uma síntese baseada também nos outros trabalhos dos autores a fim de objetividade.

prováveis, o fator associativo da representatividade pesa muito mais – 89%, em uma das formulações (DANIEL KAHNEMAN; SHANE FREDERICK, 2002, p. 62; KAHNEMAN, 2013, p. 158).

Outro exemplo similar é o de Tom W. (BARON, 2007, p. 147–149; KAHNEMAN; TVERSKY, 1972), um jovem hipotético bastante inteligente porém não muito criativo, que prefere as coisas de maneira clara e objetiva, possui humor bobo e tem preferência por ficção científica, não é muito carismático e é um pouco egocêntrico, todavia com alto senso moral. Os participantes então deveriam selecionar a qual estereótipo de estudante de pós-graduação o jovem Tom W. se assemelharia, de exatas a humanas, e a maioria gritante, como no exemplo anterior, optou por áreas como informática e matemática – que estatisticamente são muito menos comuns do que ciências sociais ou humanidades em geral, áreas universalmente mais populares.

Uma noção geral de ambas heurísticas é a utilização do esforço mental – de uma forma ou outra – para medir a probabilidade subjetiva. É certamente mais difícil imaginar um processo incerto gerando resultados não-representativos. Estes últimos resultados, evidentemente, produzem mais surpresa, que é uma forma de esforço mental. Similarmente, quanto menor a disponibilidade das instâncias de um evento, mais difícil é acessar e construí-lo, e julga-se menor a probabilidade daquele evento. (KAHNEMAN; TVERSKY, 1972, p. 452, tradução nossa)<sup>104</sup>

Novamente, o esforço cognitivo torna-se peça central no processo de tomada de decisão. Assim, de maneira simples, Tversky e Kahneman identificaram que a “cognição é realizada por um conjunto de procedimentos heurísticos simples, que são efetivos em muitas circunstâncias mas propensos a erros em outras” (HASELTON; NETTLE, 2006, p. 47, tradução nossa)<sup>105</sup>. Outra maneira de referir-se a estas “heurísticas” é como “atalhos”, ou mesmo “guindastes” na perspectiva evolutiva: são maneiras simplificadas, econômicas e práticas para nossa estrutura cognitiva processar estímulos e informações e

---

<sup>104</sup> No original: “One general notion shared by both heuristics is the use of mental effort – of one form or another – to gauge subjective probability. It is certainly harder to imagine an uncertain process yielding a nonrepresentative outcome than to imagine the same process yielding a highly representative outcome. The former outcome, of course, produces more surprise, which is a form of mental effort. Similarly, the less available the instances of an event, the harder it is to retrieve and construct them, and the lower the judged probability of that event.”

<sup>105</sup> No original: “cognition is performed by a set of simple heuristic procedures, which are effective in many circumstances but prone to error in others”.

exercer respostas. O reconhecimento amplo dos mecanismos identificados originalmente na dissonância cognitiva, então, a suplanta em uma categoria inteira de vieses cognitivos, onde há várias outras instâncias de atalhos que nossa mente aplica em diferentes situações. São, de certa maneira, bastante semelhantes a falácias lógicas, como veremos. Todavia, o foco não está nos critérios de uma inferência, mas nos fatores alheios que influenciam a tomada de decisão. A quantidade de tipos diferentes de vieses é extensiva demais para estudarmos plenamente neste trabalho (BARON, 2007, p. 54–58; GROSS, 2020, p. 333; SHERMER, 2011, cap. 12), porém vamos tratar de alguns agora e voltaremos também no próximo capítulo a fim de aprofundar-nos em algumas evidências sobre eles.

Um exemplo de Tversky e Kahneman é o viés, ou efeito, de ancoragem (TVERSKY; KAHNEMAN, 2013, p. 1128–1130). Um dos primeiros propostos, consiste na geração de uma expectativa determinada baseada em uma experiência prévia, todavia sem correlação lógica alguma. Nesta apresentação, os autores questionam participantes acerca da quantidade de países africanos que participam das Nações Unidas. Ao serem questionados, rolava-se uma “roda da fortuna” com números aleatórios, e pedia-se para estimar se o número era maior ou menor que se encontrava no resultado da roda. Participantes que viram números maiores sugeriram valores mais altos e vice-versa: o primeiro número servia, então, como uma “âncora” na qual as pessoas baseavam seu juízo. Este exemplo, como muitos outros confirmando a existência deste efeito, confirma a importância da economia de energia ao mantermo-nos próximos à nossas crenças prévias, ainda que tenham sido induzidas momentos antes (BARON, 2007, p. 380).

Um dos mais famosos, especialmente na era da pós-verdade gerada pela internet e as redes sociais, é o *viés de confirmação*, que Michael Shermer refere-se como “a mãe de todos os vieses cognitivos” (SHERMER, 2011, p. 280), referindo-se ao fato que de alguma maneira ou de outra os outros parecem diferentes formulações deste. Ele ocorre quando a pessoa possui alguma crença e tende a ignorar evidências contrárias a ela, focando apenas naquelas que a confirmam. Se um indivíduo possui uma crença X e é confrontado com uma perspectiva contrária, tende a favorecer a importância de material a favor e de dispensar mesmo uma maioria de evidências contrárias (GROSS, 2020, p. 333).

Isto foi comum recentemente com o crescimento das teorias da conspiração como o movimento antivacina, que se baseia em um artigo científico retraído apesar do montante esmagador de evidências contrárias, e também – como imagina-se – ocorre comumente com posicionamento político, religioso e esportivo, terrenos férteis para uma miríade de vieses (GROSS, 2020, p. 333–334). Há, inclusive, uma associação deste viés com o suposto *backfire effect*, ou “efeito do tiro pela culatra” em bom português, que sugere que existe um reforço na crença de indivíduos ao serem expostos a argumentos e evidências contrários.

Tratamos nesta subseção do aspecto psicológico dos limites do funcionamento de nossa estrutura para a tomada de decisão através dos vieses cognitivos. Em suma, eles

quase sempre distorcem percepções para encaixarem-se em conceitos pré-concebidos. Crenças configuram percepções. Não importa qual sistema de crença está em questão – religioso, político, econômico ou social – estes vieses cognitivos modelam como interpretamos informações que passam pelos nossos sentidos e as moldam para adequarem-se à maneira que queremos que o mundo seja, e não necessariamente como de fato é.<sup>106</sup> (SHERMER, 2011, p. 279, tradução nossa)

Seu entendimento contemporâneo em conjunto com as outras ciências previamente apresentadas delinea cada vez mais claramente a imagem de que a objetividade de nossas faculdades agenciais é sujeita a uma série de restrições que podem levar ao erro, uma vez que focam na economia de energia. E esta economia é útil, de fato, uma vez que se fôssemos analisar com profundidade todas as variáveis possíveis em todas as decisões desde o momento que abrimos os olhos no início do dia, estaríamos exaustos e possivelmente não teríamos ido muito longe de nossa cama até a hora de voltar para ela. Como muito na biologia, o que muitas vezes consideramos “limitações” na verdade são fruto de otimização para tarefas as quais não mais realizamos, centenas de milhares de anos após nossos genótipos e fenótipos estabelecerem-se como tais. Pássaros são “limitados” a não serem capazes de dirigir um carro, assim como nós somos “limitados” em não podermos voar; mas em papéis invertidos, tais

---

<sup>106</sup> No original “they almost always distort percepts to fit preconceived concepts. Beliefs configure perceptions. No matter what belief system is in place—religious, political, economic, or social—these cognitive biases shape how we interpret information that comes through our senses and mold it to fit the way we want the world to be and not necessarily how it really is”.

limitações são consideradas capacidades. Vamos tratar um pouco mais deste aspecto a seguir.

Na próxima subseção nos adentraremos na teoria que servirá de ponte direta entre a psicologia cognitiva e as neurociências comportamentais: a teoria do processo duplo. Ela parte dos mesmos expoentes da heurística cognitiva, Tversky e Kahneman, a fim de entender de maneira mais profunda os mecanismos neurofisiológicos por detrás dos processos cognitivos envolvidos na tomada de decisão e processos mentais em geral. Iniciaremos sua explanação aqui pois trata-se de entender primeiramente quais são e como funcionam estes mecanismos cognitivos, para no próximo capítulo aprofundarmos-nos em sua configuração fisiológica.

#### *4.3.3 Dois sistemas, dois processos*

O estabelecimento na seção anterior de vieses que diretamente afetam nossa cognitividade, especialmente nas pesquisas de Tversky e Kahneman, levou ambos a identificar padrões notáveis ao analisarem seu funcionamento na deliberação e percepção em geral. Nesta seção, desenvolveremos o impacto da descoberta desta miríade de vieses em uma estrutura metodológica que apesar de não pressupor a existência diferenciada de sistemas psicológicos de raciocínio, identifica características comuns de nossas faculdades cognitivas que funcionam de maneiras semelhantes em dois contextos predominantes. Esta teoria categoriza os vieses cognitivos e os traduz diretamente para a linguagem da economia comportamental, formalizando de maneira técnica a interação entre esta área e a psicologia e, como veremos posteriormente, também com as neurociências. Apresentaremos aqui brevemente a base teórica desta perspectiva para então desenvolver sua análise nas neurociências comportamentais na próxima seção.

Ao que Tversky e Kahneman estabeleceram esta gama de fatores que podem influenciar direta ou indiretamente a maneira que pensamos e agimos, também ficou evidente que há dois tipos de sistemas predominantes que agem cada um de acordo com o contexto específico em que o processo deliberativo ocorre. À este modelo deu-se o nome de “sistema duplo” ou “processo duplo”, que sugere que há dois processos (ou sistemas) principais representativos do funcionamento cognitivo: o primeiro, Tipo ou sistema 1 (T1), ou, associado com

processos sumariamente intuitivos, afetivos, involuntários e inconscientes, o “rápido”; e o segundo, Tipo ou sistema 2 (T2), com processos cognitivos, voluntários e conscientes, o “devagar” (CLINE, 2017, p. 582–585; KAHNEMAN, 2013, p. 19–22). É necessário compreender que esta nomenclatura não procura denotar duas “entidades” ontológicas em nossos cérebros, e sim duas maneiras de se referir a processos cujas áreas ativadas exponencialmente tendem a se sobrepôr. Assim, ambos estão envolvidos no processo de tomada de decisão e cognição em geral; apenas tendem a estar mais ativos em determinados momentos. Sua divisão é meramente didática. (KAHNEMAN, 2013, p. 24–25)

O sistema T1 lida com nossas “intuições”, respostas imediatas, como a que surge ao sermos confrontados com o cálculo simples “ $1+1=$ \_\_”; nosso reflexo cognitivo padrão à estímulos em geral. T2, por outro lado, é mais relacionado à memória de trabalho, sendo uma atividade mais esforçada, lenta e deliberada, como precisar calcular “ $38 \times 419 =$ \_\_”. Como este exemplo, o agente tende a evitar seu uso a não ser que haja recompensa, bastante em jogo ou algum conflito com T1. É aqui que os autores se referem ao falar de “crença”: T1 ocupa-se automaticamente a todo o momento de mandar estímulos para os processos de T2, podendo, se por ele endossados, gerar crenças (KAHNEMAN, 2013, p. 24). Thaler e Sunstein, autores da teoria da arquitetura da escolha como vimos na seção sobre economia comportamental, também abraça a teoria do processo duplo como seu paradigma basilar, referindo-se a T1 como “sistema automático” e T2 como “sistema reflexivo” (THALER; SUNSTEIN, 2008, p. 19–22). Uma vez geradas, assim como no caso dos teóricos conspiracionistas do apocalipse alienígena, tais crenças são então cada vez mais reforçadas a ponto de tornar-se parte do T1 e fazerem parte intuitiva de nossa deliberação (KAHNEMAN, 2013, p. 24; THALER; SUNSTEIN, 2008, p. 30). Todavia, esta concepção de crença é problemática na filosofia, especialmente nos debates sobre cognitividade epistemológica. Veremos um pouco sobre isto no capítulo 7.

Esta teoria então evidencia como processos involuntários são predominantes no comportamento, uma vez que T2 só entra em cena em situações específicas por requerer razoável esforço em oposição à passividade de T1. Da mesma maneira, ao evocarmos T2, limitamos nossa atenção e capacidade de pensar adequadamente por ocuparmos então nossa memória de trabalho (KAHNEMAN 2013, p. 30). Aqui retornamos a um dos fatores principais

que delineiam limites para a racionalidade: o esforço. O *modus operandi* padrão de nossa estrutura cognitiva é atuar por meio de mecanismos automáticos que simplificam a deliberação a fim de otimizar a relação entre o esforço e a recompensa obtida. Como nos exemplos matemáticos, é bem provável que no primeiro exemplo quem esteja lendo este texto complete automaticamente, praticamente involuntariamente com a resposta a lacuna de “ $1+1=$ \_\_”, enquanto no segundo exemplo, “ $38 \times 419 =$ \_\_”, nem ao menos tente começar a realizar tal cálculo. Nas deliberações do cotidiano, portanto, o padrão também é nos apoiarmos em vieses que encurtam o esforço cognitivo em troca de uma acurácia imperfeita.

Um exemplo interessante proposto por Brendan Cline é o do juízo de perigo. Tal se caracteriza na crença motivada de que algo é perigoso. Para Cline, é possível termos crenças sobre perigo a partir de T2 – advertências do ministério da saúde, leitura de alguma informação sobre práticas perigosas etc. – que informam processos de T1. Todavia, também é possível que por trauma ou outra situação extraordinária tenhamos determinadas fobias que mesmo com exaustivo estímulo de T2 não alteram a motivação do indivíduo a agir. É o caso de pessoas, por exemplo, com fobia de voar de avião, que após ouvir todos os dados sobre o número comparativamente menor de acidentes, todos os mecanismos de segurança envolvidos etc. são capazes de afirmar e estarem convictos de uma crença de que é seguro viajar de avião; todavia, ainda assim não são capazes de fazê-lo. Isto não impede, no entanto, que exercícios – como sentar em um avião aterrissado e conscientemente passar por uma simulação de turbulência, fechar os olhos, escutar música e imaginar um lugar agradável – sejam capazes de paulatinamente influenciar T1 a fim de tranquilizar a pessoa a longo prazo. (CLINE, 2017, p. 587–588)

Independentemente do paradigma acerca do processo duplo ser correto ou não, ele nos ilustra de maneira clara que a racionalidade encontra-se então como secundária na tomada de decisão cotidiana. A psicologia evolutiva nos permite compreender de maneira mais ampla os motivos pelos quais nosso arcabouço cognitivo possui tais limitações: o padrão da natureza não é o da racionalidade, mas sim o da inanição – novamente, economia de energia é o foco. Basta considerarmos que o percentual de organismos no universo capazes de escolher entre mais de uma opção não-inteiramente instintiva é ínfimo, quase

desprezível em comparação à passividade predominante no vácuo e na matéria inanimada do universo inteiro e suas galáxias, além de extremamente recente na escala cronológica do cosmos como um todo. Assim, o fenômeno da possibilidade de escolher um curso de ação entre múltiplos engendra o desenvolvimento evolutivo de toda uma estrutura perceptiva, reflexiva e deliberativa do sistema nervoso do ser vivo em questão. Assim,

uma explicação causal de qualquer comportamento – racional ou não – necessariamente invoca teorias acerca da arquitetura destes aparelhos computacionais. A racionalidade de um comportamento é irrelevante para sua causa ou explicação.<sup>107</sup> (COSMIDES; TOOBY, 1994, p. 327, tradução nossa)

Ou seja, em suma, a pressuposição da exclusividade de uma objetividade deliberativa clássica cai por terra quando consideramos o contexto real em que agentes pensam, realizam escolhas e agem. Uma miríade de fatores entra em consideração e os critérios tradicionais não mais dão conta – e jamais dariam, à exceção de um agente plenamente onisciente – por si só nesta área, uma vez que

[o] mundo não é intrinsicamente dividido entre opções. Ao invés, um organismo deve processar estímulos de alguma maneira a fim de definir eventos, e, portanto, definir escolhas. Já que tal tomada de decisão é logicamente anterior à ação, podemos esperar que a seleção natural e sexual teria atuado para criar em organismos a capacidade de dividir o mundo desta maneira para facilitar o comportamento adaptativo.<sup>108</sup> (FESSLER, 2002, p. 191)

Isto reforça outro importante aspecto já mencionado a se considerar acerca de vieses cognitivos, incluindo a dissonância cognitiva e outros fatores supracitados. Em termos evolutivos, não são necessariamente uma falha: são características úteis que facilitaram a simplificação do processamento de informações (HASELTON; NETTLE; ANDREWS, 2005, p. 724–725). Se exauríssemos, como queriam os economistas clássicos ou mesmo os filósofos

---

<sup>107</sup> No original: “a complete causal explanation of any behavior – rational or otherwise – necessarily invokes theories about the architecture of these computational devices. The rationality of a behavior is irrelevant to its cause or explanation”.

<sup>108</sup> No original: “[t]he world is not intrinsically divided into options. Instead, an organism must parse stimuli in such a way as to define events, and thus to define choices. Because decision making is logically prior to action, we can expect that natural and sexual selection will have acted to create in organisms the capacity to divide up the world in this manner in order to facilitate adaptive behavior”.

racionalistas mais ferrenhos, todas as possibilidades relativas à possíveis cursos de ação de cada contexto específico de nossa vida cotidiana, iríamos não apenas desperdiçar quantidades desproporcionais de energia para as tarefas mais simples, como também estaríamos muito mais suscetíveis à predadores e acidentes, uma vez que nosso tempo de resposta a qualquer ameaça ou perigo seria consideravelmente mais elevado.

Vários destes elementos tratados de maneira breve aqui tornaram-se centrais não apenas para o estudo da psicologia cognitiva em geral, mas também para a economia comportamental, neurociências e quaisquer estudos sobre qualquer área acerca da tomada de decisão e outras temáticas tangentes à maneira que pensamos, agimos e que fatores os influenciam. Na próxima seção, veremos a manifestação última até então da culminação de todas estas áreas de estudos alinhadas ao desenvolvimento tecnológico das últimas décadas para um entendimento mais profundo do que nunca de nossos processos fisiológicos que subjazem nossos processos cognitivos: as neurociências comportamentais.

#### 4.4 Neurociências

O Homem não tem um Corpo distinto da Alma, pois aquilo que denominamos Corpo não passa de uma parte de Alma discernida pelos cinco sentidos, seus principais umbrais nestes tempos.

— William Blake, “O casamento do céu e do inferno”

As neurociências enquanto escola de pesquisa são, naturalmente, um resultado dos avanços nas pesquisas da psicologia cognitiva em seu viés experimental conforme vimos. Todavia, de certa maneira, também acaba por integrar todas as ciências apresentadas neste capítulo de maneira interdisciplinar – especialmente quando tratamos do tema do comportamento.

Já estabelecemos então os principais paradigmas nestas três grandes áreas científicas relevantes para esta tese: (i) a perspectiva experimental do naturalismo filosófico, que demonstra como o diálogo com as ciências naturais e afins podem revelar informações e aprofundar debates filosóficos; (ii) a biologia evolutiva, que destaca nosso comportamento como parte de um processo natural evolutivo análogo ao de outros animais, e não de uma esfera especial à parte da natureza; (iii) a economia comportamental, que traz à tona a

consideração que a deliberação raramente – quando muito – ocorre de maneira otimizada, objetiva e lógica, sendo sujeita à uma miríade de fatores externos; e (iv) o aprofundamento no estudo de nossas capacidades cognitivas pela psicologia revelando nossos viéses frequentemente ocultos em nossa tomada de decisão e sua categorização posterior em um paradigma de sistema duplo. Seguimos aqui então a fim de realizar a conexão entre todas estas áreas com o estudo comportamental sob o viés do sistema nervoso, que informa e aprofunda a concepção de duplo processo como um todo e nos apresenta visualmente informações acerca do tema em questão em diálogo direto com todas as áreas anteriores – inclusive a filosofia. Iniciaremos da mesma forma que os outros com uma breve introdução histórica a fim de contextualizar tais desenvolvimentos, para posteriormente termos finalmente a estrutura teórica apropriada para lidar com a interação entre as evidências que trataremos, seu contexto prático e seu impacto teórico de cunho filosófico.

A importância do estudo do sistema nervoso no tópico em questão se dá pelo fato que é o elemento mais básico de nossa fisiologia associado não apenas com a tomada de decisão, mas com todo o comportamento. Como explicam de maneira clara Cosmides e Tooby:

o cérebro é um aparelho computacional complexo, um sistema que toma a informação sensorial como estímulo, transforma-a em diversas maneiras, armazena-a, analisa-a, integra-a, aplica regras de decisão a ela e então traduz o resultado destas regras nas contrações musculares as quais chamamos de “comportamento”.<sup>109</sup> (COSMIDES; TOOBY, 1994, p. 328)

Não iremos abraçar nenhuma concepção específica da consciência além da rejeição absoluta do dualismo, como um entendimento mínimo dos estudos de neurociências sugere e espero que fique claro dentro desta seção. Os termos “mente” e “corpo” podem ser utilizados separadamente neste trabalho, todavia não pressupõe uma distinção metafísica de fato: ambos são diferentes manifestações – de uma maneira ou de outra – de uma mesma estrutura fisiológica.

---

<sup>109</sup> No original: “[t]he brain is a complex computational device, a system that takes sensory information as input, transforms it in various ways, stores it, analyzes it, integrates it, applies decision rules to it, and then translates the output of those rules into the muscular contractions that we call “behavior”.

#### *4.4.1 De buracos no crânio à ressonâncias eletromagnéticas*

Muito antes da filosofia associar o cérebro com a mente, ritos que associavam comportamentos estranhos com o cérebro já existiam há mais de 5000 anos atrás durante o período neolítico, principalmente na Europa e no Peru, em esboços primitivos de uma neurociência médica. O procedimento conhecido como “trepanação” envolvia a remoção de parte do crânio de indivíduos para uma série de possíveis fatores, tanto religioso e ritualístico como também médico e comportamental (VERANO; FINGER, 2009). Todavia, não foi até a união da anatomia, embriologia, fisiologia, farmacologia e da própria psicologia no século 19 que a área abrangente das neurociências evoluiu da neurocirurgia e da perspectiva do sistema nervoso como apenas outra parte do corpo como qualquer outra. O ponto crucial para sua união veio quando Camillo Golgi e Santiago Ramón y Cajal utilizaram o que foi uma revolução tecnológica para a ciência na época, o microscópio composto, para observar que as células neurais funcionavam de uma maneira diferente das de outros locais do corpo. Todavia, de fato, foi com a filosofia a partir de Descartes, Locke e Hume principalmente que o aspecto comportamental foi associado com o sistema nervoso. Mas Franz Gall que foi o primeiro a elaborar uma teoria neurológica do comportamento chamada de “frenologia”, associando partes específicas do cérebro independentemente a funções diferentes, e mesmo a traços de caráter específicos – todavia, mesmo em sua época já fora criticada por experimentos práticos demonstrando que o cérebro age em uníssono, com todas as áreas atuando em todas as funções, apenas variando em grau (KANDEL; SCHWARTZ; JESSELL, 1996, p. 6–8).

Com crescente avanço em metodologias e tecnologias de imageamento e medição de neurotransmissores dentre outros fatores, as neurociências têm estado cada vez mais na vanguarda de uma miríade de diversas áreas de pesquisa. Relevante ao tema deste trabalho, as neurociências vem a oferecer à psicologia cognitiva e às outras áreas relativas ao comportamento “informações importantes sobre regiões do cérebro envolvidas em comportamentos complexos específicos, e como estes comportamentos podem ser segmentados em operações mentais mais simples em regiões cerebrais interconectadas

específicas”<sup>110</sup> (KANDEL; SCHWARTZ; JESSELL, 1996, p. 19). Com estas novas informações, somos capazes de entender como nunca antes a arquitetura neural e fisiológica em geral relativa às medições comportamentais. O foco torna-se o sistema nervoso uma vez que é o órgão principal relativo às atividades relevantes na tomada de decisão, justamente por ser, em suma, “uma rede precisa de mais de 100 bilhões de células nervosas, interconectadas em sistemas que produzem nossa percepção do mundo externo, fixam nossa atenção e controlam o maquinário da ação”<sup>111</sup> (KANDEL; SCHWARTZ; JESSELL, 1996, p. 3).

Entender o funcionamento do sistema neural, assim como dos estudos da racionalidade limitada na economia e dos vieses na psicologia cognitiva, não significa reduzir todo o comportamento humano ao cérebro ou ao sistema nervoso como um todo somente. Apesar do naturalismo filosófico conforme tratado aqui pressupor uma redutividade no sentido de eliminar o sobrenatural do seu discurso, reduzindo a esfera dos problemas filosóficos apenas no que tange à fatos sobre o mundo empírico (IGANSI, 2014b, p. 132), não significa que haja uma redução da totalidade do comportamento meramente para elementos relativos ao cérebro. Afinal, mesmo as ciências neurológicas como um todo possuem para si como tarefa apenas

prover explicações do comportamento em termos de atividades do cérebro, explicar como milhões de células nervosas individuais no cérebro operam a produzir comportamento, e como, por sua vez, estas células são influenciadas pelo ambiente, incluindo o comportamento de outras pessoas.<sup>112</sup> (KANDEL; SCHWARTZ; JESSELL, 1996, p. 5)

Os termos “células” e “comportamento” nos remetem a duas áreas distintas: biologia e psicologia, respectivamente. A análise comportamental através de um viés neurocientífico, também conhecido com neurociências comportamentais, possui um outro nome que ilustra este seu teor interdisciplinar:

---

<sup>110</sup> No original: “*important information about regions of the brain involved in specific complex behaviors, and how these behaviors can be broken down into simpler mental operations in specific interconnected brain regions*”.

<sup>111</sup> No original: “*a precise network of more than 100 billion nerve cells, interconnected in systems that produce our perception of the external world, fix our attention, and control the machinery of action*”.

<sup>112</sup> No original: “*to provide explanations of behavior in terms of activities of the brain, to explain how millions of individual nerve cells in the brain operate to produce behavior and how, in turn, these cells are influenced by the environment, including the behavior of other people*”.

*biopsicologia*, ou *psicobiologia*. Assim como é pretendido ilustrar neste trabalho, os estudos comportamentais de ambos, da biologia e da psicologia, são engrandecidos exponencialmente ao dialogarem diretamente – mesmo sendo comumente tratadas como áreas tão destoantes tal como a filosofia o é delas. Aqui retornamos a um dos supostos fundadores da psicologia, Wundt, e seu intento de tratar de problemas cognitivos de maneira experimental na psicologia. Finalmente, pouco mais de um século depois, suas hipóteses – anteriormente limitadas ao campo introspectivo e conjectural – podem ser adequadamente testadas e medidas com o devido crivo da metodologia científica. A cognição, não mais uma entidade metafísica fora do escopo empírico como proposto pelos respectivos pesquisadores na psicologia, agora torna-se uma parte tangível de nosso arcabouço psicofisiológico através das neurociências: “funções cognitivas complexas ocorrem e podem ser entendidas em termos da rede de conexões entre neurônios”<sup>113</sup> (SOLSO; MACLIN; MACLIN, 2014, p. 66, tradução nossa). Assim, a união entre as diferentes áreas tratadas até agora fica cada vez mais evidente, já que

[a]qui é onde a integração de biologia evolutiva e ciência cognitiva provou-se tão útil. A aplicabilidade da biologia evolutiva é baseada em uma ideia simples mas poderosa. A forma segue a função: as propriedades de um mecanismo evoluído refletem a estrutura da tarefa que evoluiu para resolver.<sup>114</sup> (COSMIDES; TOOBY, 1994, p. 328, tradução nossa)

Após este contexto histórico e teórico do surgimento das neurociências e seu papel na integração da pesquisa sobre deliberação, na próxima subseção analisaremos brevemente como se dá então de fato sua anatomia e funcionamento básico. Posteriormente, procederemos então em investigar como há tal integração na análise de conceitos importantes para o estudo da tomada de decisão.

---

<sup>113</sup> No original: “Complex cognitive functions take place and can be understood in terms of the network of links among neurons”.

<sup>114</sup> No original: “[t]his is where the integration of evolutionary biology and cognitive science has proved so useful. The applicability of evolutionary biology is based on a simple but powerful idea. Form follows function: the properties of an evolved mechanism reflect the structure of the task it evolved to solve”.

#### 4.4.2 Neuroanatomia<sup>115</sup>

O sistema nervoso é um dos maiores órgãos do corpo humano, e arguivelmente seu mais importante. Do toque ao equilíbrio, da visão à audição, da alegria ao sofrimento, todas nossas sensações, pensamentos e reações são produzidos, interpretados e/ou no mínimo mediados por ele. Está presente nas terminações nervosas de nossas mãos que apreendem o toque, nos estímulos de órgãos e músculos para o funcionamento “passivo” dos mesmos e principalmente no encéfalo, comumente referido como “cérebro”, o maior responsável pelo processamento de informações e geração de ações a partir destes estímulos. Nesta subseção veremos brevemente as principais partes da anatomia do sistema nervoso relevantes para o tópico em questão.

O encéfalo, que engloba cérebro, tálamo, mesencéfalo, ponte do cerebelo e bulbo, em conjunto com a medula espinhal constituem o sistema nervoso central. O resto de suas conexões nervosas com o corpo constituem o sistema nervoso periférico. Todo o sistema nervoso é composto de dois tipos de célula: as células nervosas, conhecidas como “neurônios”, e as células da glia. As primeiras são caracterizadas pelos axônios, que são terminações dendríticas responsáveis pela transmissão de sinais através da diferença de potencial químico nas células através de diferentes cátions e íons de potássio, podendo ser sensoriais ou motoras. Esta transferência eletroquímica entre axônios se chama de sinapse e se dá por neurotransmissores depositados nas vesículas sinápticas nas extremidades dos dendritos destas células, liberados pelo neurônio pré-sináptico para o pós-sináptico. As células da glia, por outro lado, compõe a maior parte da massa do sistema nervoso garantindo sua estrutura física e manutenção, sendo responsável, por exemplo, pelo cuidado com as bainhas de mielina das células nervosas, cuja quantidade influencia na maior velocidade e qualidade dos sinais eletroquímicos enviados entre diferentes células. Alguns neurotransmissores conhecidos são a dopamina, relacionada à emoções e recompensa, e a serotonina, que regula o humor, sono, apetite e outras funções vitais (KANDEL; SCHWARTZ; JESSELL, 1996, p. 21–69; SANTOS, 2002, p. 12–13).

---

<sup>115</sup> É imensurável aqui meu agradecimento aos excelentíssimos pesquisadores que me auxiliaram na compreensão deste tópico tão complexo: Dra. Heloísa Allgayer, Dr. Matheus Silveira, Dr. João Scaini, Dr. Marcos Cordeiro, Dnda. Giuly Iturralde.

Assim, quando falamos de “informação”, falamos dos estímulos às terminações nervosas que é carregado por sinapses, estes padrões eletroquímicos que passam pelos dendritos neuronais às extremidades das células nervosas, geralmente atravessando a coluna vertebral e sendo interpretados pelo encéfalo. Estímulos na epiderme na ponta dos dedos, por exemplo, chegam em descarga elétrica no cérebro que é capaz de interpretar tal textura como uma superfície áspera, lisa, seca, úmida etc.; o qual devolve outro impulso após o processamento por trajetória similar informando se nossa mão deve manter-se ali ou ser retirada. A luz é apreendida pelo globo ocular e tal impulso é convertido em energia elétrica que estimula determinadas áreas encefálicas que o processam em imagem, e em contraparte elegendo movimentos corporais, maior foco ocular ou quaisquer outras ações que julgar necessárias naquele momento. A complexidade destes sinais da maneira como os entendemos é possível pela pura multiplicidade de combinações de conexões, articulações celulares e neurotransmissores possíveis, tendo capacidade de processamento muito maior do que mesmo os computadores mais complexos (KANDEL; SCHWARTZ; JESSELL, 1996, p. 29–40).

Estes sinais eletroquímicos, todavia, não alcançam a totalidade do encéfalo simultaneamente. Em um primeiro momento, são trazidos pelo tronco encefálico para a área chamada de sistema límbico<sup>116</sup>. Este sistema, conhecido como a união do “cérebro reptiliano” e “cérebro paleomamífero” pela categorização de MacLean em 1990, é associado com diferentes emoções, comportamentos associados às mesmas e memória de longo prazo (ROXO et al., 2011, p. 2433–2436). O motivo pela nomenclatura de MacLean é justamente sua extrema semelhança anatômica com vários outros animais vertebrados, que mesmo diferindo em tamanho e proporções mantém-se mecanicamente extremamente similares (KANDEL; SCHWARTZ; JESSELL, 1996, p. 78–81), além de historicamente prévios em nossa história anatômica evolutiva (WEBER; JOHNSON, 2008, p. 132). Afinal, esta compreensão do sistema nervoso não apenas se aplica para humanos, mas também para qualquer criatura vertebrada. Novamente, reforça-se a análise de Darwin acerca da comparação entre fisionomia, comportamento e faculdades de diferentes animais, uma vez que

---

<sup>116</sup> Há controvérsia acerca desta denominação na bibliografia; manteremos assim para fim didático.

“[n]o homem, assim como nos animais inferiores, muitas estruturas estão tão intimamente relacionadas, que quando uma parte varia, tal variação ocorre em outra também, sem podermos, na maioria do caso, atribuir alguma razão”<sup>117</sup> (DARWIN, 1889, p. 43, tradução nossa). Esta parte do sistema nervoso central, situada diretamente acima da coluna vertebral é, portanto, associada com instintos e intuições, nossas reações e ímpetos automáticos, usualmente relacionados com a sobrevivência, assim como a memória associada com estes elementos a longo prazo.

Uma das principais partes deste sistema é a amígdala cerebelosa<sup>118</sup>. Ela é responsável pelo processamento das emoções e associações com o hipocampo e com o neocórtex. Duas delas jazem em cada hemisfério do hipocampo como duas pequenas amêndoas. Ela será bem importante nos capítulos a seguir, pois é uma das áreas mais importantes para fatores emocionais. O hipocampo, por sua vez, localiza-se ao centro do sistema límbico, e é responsável pelas memórias a longo prazo, cujo conteúdo emocional é atribuído pelas amígdalas.

É importante aqui frisar, no entanto, que ter atividade em áreas associadas com determinados eventos não implica em uma exclusividade daquela área para tais. Mais de uma área pode ser responsável por determinada função, assim como áreas específicas podem atuar em diversas funções. O encéfalo, apesar de ser apenas um órgão dentre tantos no corpo e compreendendo apenas cerca de 2,5% da massa corporal, é responsável por consumir de oito a dez vezes mais energia calórica total do que qualquer órgão do corpo, totalizando cerca de 20% de toda energia que produzimos (PULIDO; RYAN, 2021, p. 1). Mesmo ao dormir ou consumir entretenimento para relaxar há uma miríade de atividades passivas no sistema nervoso ocorrendo através de diversos estímulos que muitas vezes nem percebemos. Aqui retornamos ao ponto da economia de energia: se fôssemos conscientemente processar cada som, sensação, variação de temperatura, estímulo na epiderme etc. não apenas esgotaríamos nossas reservas calóricas rapidamente como também não saberíamos discernir entre estímulos importantes de coisas banais. Afinal,

---

<sup>117</sup> No original: “*In man, as in the lower animals, many structures are so intimately related, that when one part varies so does another, without our being able, in most cases, to assign any reason*”.

<sup>118</sup> Não confundir com as tonsilas, as amígdalas palatinas localizadas na parte posterior da boca.

estariamos tendo que ter atenção ao processar a mínima informação adquirida pelas terminações nervosas, e não saberíamos diferenciar uma folha voando ao vento de um veículo vindo em nossa direção em alta velocidade. Assim, há uma série de mecanismos automáticos que funcionam passivamente sem ativar nossa atenção, memória e outras funções cognitivas importantes, deixando esta função para algo que venha a estimular a amígdala – tanto por instinto natural quanto por memória – para que estimule as emoções associadas com os processos responsáveis por lidar com a situação.

O córtex e o neocórtex, por outro lado, consistem na parte mais desenvolvida ao redor do sistema límbico, compreendendo quase todo o resto do volume cranial do encéfalo, representando 82% de sua massa (HERCULANO-HOUZEL, 2009, p. 3). O neocórtex, também conhecido como isocórtex, consiste nas seis camadas corticais mais exteriores do encéfalo, ou quatro em algumas como no paleocórtex, medindo em média 1-4mm. Neles estão localizadas as áreas associadas com o principal funcionamento de atividades de maior ordem como a linguagem e a reflexão. Cálculos matemáticos complexos, decidir uma rota de carro em um horário movimentado, imaginar complexas interações fictícias; grande parte dos sistemas envolvidos nestas atividades estão presentes nestas regiões. É uma das mais recentes no desenvolvimento evolutivo, e também uma das mais destoantes dos sistemas nervosos centrais de outros mamíferos por sua proporção muito maior na espécie humana. Usualmente, além da divisão clara entre o hemisfério direito e esquerdo, o córtex é dividido em quatro seções principais, chamadas de “lobos”: (i) o lobo frontal, que vai da região posterior aos olhos até pouco antes da metade, (ii) o lobo occipital, que jaz na área traseira do crânio próxima à coluna, (iii) o lobo parietal, na área central superior e (iv) o lobo temporal, na parte inferior da área central. Ambos hemisférios controlam certas funções diferentes, sendo a principal delas o aspecto motor em que o lado direito controla o lado esquerdo e vice-versa (KANDEL et al., 2021, p. 7–16; KANDEL; SCHWARTZ; JESSELL, 1996, p. 71–88, 347–351; SANTOS, 2002, p. 7–10).

A superfície destas regiões é irregular, “antigo artifício biológico para aumentar a área de superfície” (SANTOS, 2002, p. 8) com dobraduras cujas partes mais externas são denominadas de giros, e as reentrâncias entre elas de sulcos ou fissuras. Embora cada cérebro seja único em sua topologia, algumas

destas mantém-se comuns, como por exemplo o sulco lateral, também conhecido como “fissura de Sylvius”, uma das características mais proeminentes nos hemisférios laterais do encéfalo que parece dividir a área frontal do córtex da área parietal. Outra perspectiva que é importante de se salientar é a de que não há correlação entre a fisionomia do sistema nervoso e inteligência. Peso, tamanho das áreas, densidade neuronal, características dos giros e sulcos – nada possui relevância direta para as capacidades cognitivas individuais. Isto se dá mesmo em comparação com a anatomia de um encéfalo como de Albert Einstein, um nome frequentemente associado com genialidade e proeza intelectual, cujo cérebro é majoritariamente idêntico ao de qualquer um, variando apenas levemente em seu sulco lateral (WITELSON; KIGAR; HARVEY, 1999).

Por fim, evidentemente, o sistema nervoso não se desenvolve sempre da mesma maneira desde o nascimento. A neuroplasticidade, primeiramente formulada por um dos pais da neurociência moderna, Santiago Ramón y Cajal, é uma característica que implica na constante alteração de conexões sinápticas de acordo com experiências do agente (FURTADO et al., 2022; KANDEL et al., 2021, p. 71–72). Como veremos no caso de Phineas Gage no próximo capítulo, que teve seu crânio perfurado e seu cérebro atravessado por uma barra sólida de metal, lesões e outras influências fazem com que o sistema nervoso compense a ausência ou modificação de determinadas regiões redirecionando as sinapses a fim de seu contínuo funcionamento. Outro exemplo é da “dor fantasma”, quando alguém que perde alguma parte do seu corpo em algum acidente sente dores justamente onde não há mais nada ali, denunciando que as vias neuroquímicas ainda existem apesar de tal ausência (KANDEL et al., 2021, p. 474). Iremos nos aprofundar em outras áreas futuramente, mas por enquanto é suficiente esta compreensão superficial das principais regiões neurais. Vejamos agora brevemente também alguns mecanismos para seu mapeamento.

O diferencial de potencial nas cargas gerais de determinadas regiões neuronais do encéfalo é o que permite sua medida através da eletroencefalografia (EEG), que se utiliza de sensores no couro cabeludo capazes de medi-los a fim de detectar quais que possuem maior atividade em determinado momento. É um método não-invasivo, o que facilita seu uso e versatilidade de aplicação, e de custo menor que as outras alternativas. Todavia,

é sujeito a interferência eletromagnética de outros equipamentos, e seu resultado consta em resultados numéricos de tal diferença de potencial em gráficos em plano cartesiano, que apenas posteriormente podem ser traduzidos em visualizações fisiológicas. Depois temos a tomografia por emissão de pósitrons, chamada de “PET” por ser a sua sigla na língua inglesa, que se utiliza oxigênio-15 no corpo da pessoa para que torne suas partes internas visíveis por contraste a partir da captação de raios gama no decaimento de pósitrons após seu choque com electrons. A imagem gerada é bastante fiel ao órgão encefálico real, todavia requer que a pessoa tenha tal inserção de oxigênio-15 em seu corpo e deve manter-se deitada em um equipamento relativamente grande. De maneira semelhante, o imageamento por ressonância magnética e por ressonância magnética funcional gera a representação mais fiel do órgão de todos os equipamentos, oferecendo o equivalente a fatias acuradas do encéfalo de maneira horizontal e vertical. Isto é feito com o indivíduo também deitado, todavia dentro de um grande maquinário que emite sons bastante altos. Tal tecnologia conta com o uso de campos magnéticos e ondas de rádio do próprio corpo para gerar imagens quase em tempo real de suas seções, enquanto a (BABILONI et al., 2004; GREENE et al., 2001; KHALILI; MORADI, 2009; MICHEL; MURRAY, 2012; OHME et al., 2009; OLEJNICZAK, 2006; VECCHIATO et al., 2010; XU; PLATANIOTIS, 2012).

Embora esta seção tenha sido breve e de maneira alguma represente a riqueza anatômica do sistema nervoso como um todo, estabelece o básico acerca do principal órgão relacionado às neurociências e, por conseguinte, um dos mais importantes para compreender os próximos capítulos deste trabalho. Na próxima seção, analisaremos conceitos da psicologia cognitiva aprofundando então a conexão entre estas duas áreas, para então retornarmos à teoria do processo duplo enquanto localizado pelas ciências neurais – a última peça da estrutura deste trabalho para encerrarmos este capítulo.

#### *4.4.3 Integrando conceitos*

Alguns dos principais conceitos básicos da psicologia cognitiva – e mais importantes para entender o processo deliberativo como um todo – são percepção, atenção, memória, resolução de problemas e tomada de decisão, dentre outros (GROSS, 2020, p. 2). Embora separado, o conceito de tomada de

decisão envolve diferentes e ilimitadas integrações entre todos outros, apesar de isto não seja uma característica única. Dedicaremos esta subseção para brevemente ilustrá-los a fim de ter uma compreensão maior do uso destes termos ao resto do texto.

Apesar de associarmos percepção primariamente aos estímulos aos nossos sentidos como a visão, audição e olfato, a percepção na psicologia cognitiva refere-se a como processamos estes estímulos em informações. O famoso psiquiatra e neurocientista Oliver Sacks, por exemplo, sofria de prosopagnosia, um distúrbio neurológico que afeta como sua visão é processada. Seus olhos e centros nervosos associados com a visão estão praticamente intactos; todavia, seu giro fusiforme, a área responsável pelo reconhecimento facial localizada parcialmente no lobo occipital e no temporal, era danificado. Assim, ele era capaz de identificar sem dificuldade diferentes objetos e ter uma vida razoavelmente normal, com a exceção que ele não sabia diferenciar o rosto de entes queridos – ou mesmo o seu próprio – de qualquer outro rosto em multidões, tendo que apelar a outras estratégias como postura, cabelo, sobrancelhas, cor do olho, outras partes corporais ou comportamentos para inferir que pessoa que pode ser. Ele via rostos e cada parte que os compõe, mas era incapaz de registrá-los – inclusive pedindo desculpas à homens grandes e barbudos por quase chocar-se com eles quando na verdade tratava-se dele mesmo no espelho (KATZ, 2010).

A faculdade cognitiva de percepção, então, é constituída em como processamos os estímulos sensoriais externos e como somos capazes de compreendê-los (GROSS, 2020, p. 242), e é central para a experiência humana como um todo e para o desenvolvimento do conhecimento (KANDEL et al., 2021, p. 381). Várias áreas do sistema nervoso são associadas com a percepção. Vamos tratar brevemente dos cinco sentidos clássicos, visão, audição, olfato, paladar e tato, e desenvolvê-los um pouco para mostrar não apenas seu funcionamento e neuroanatomia como também que possuem papéis bem maiores do que esta denominação sugere. Uma das mais importantes é a visão, cuja apreensão de iluminação pelo globo ocular através da retina para os fotorreceptores que a traduzem em impulsos elétricos para o córtex visual, localizado no lobo occipital (KANDEL et al., 2021, p. 381, caps. 21–22).

O sistema auditivo, por sua vez, é responsável por interpretar vibrações no ar através de uma rede complexa de partes nas orelhas e passar esta informação pelo nervo coclear até os lados superiores do lobo temporal, mais especificamente no sulco lateral englobando os giros transversais temporais e o giro temporal superior. Não apenas o sistema auditivo capta sons, mas ajuda na localização espacial através deles e inclusive no nosso próprio equilíbrio (KANDEL et al., 2021, p. 382, caps. 26 e 28).

Já o paladar e o olfato correspondem à compreensão química do mundo por diferentes interações entre sabores e odores de diferentes maneiras em receptores especializados. Mamíferos em geral, como nós, possuem uma quantidade imensa de receptores olfatórios de diferentes tipos, oferecendo uma sensibilidade e versatilidade única para detecção e identificação de diferentes tipos de odores em diferentes graus. Os receptores do sistema gustativo, por sua vez, majoritariamente localizados na língua em suas papilas gustativas, detectam cinco submodalidades de sabor que refletem os requerimentos necessários de dieta alimentícia, cada uma com seu tipo de receptor sensorial e célula. O córtex olfativo localiza-se no uncus, cujo nome significa “gancho” em latim, aludindo ao seu formato na região paralela ao hipocampo (para-hipocampal), e compreende o córtex piriforme, o tubérculo olfatório, o giro para-hipocampal e também a amígdala. O córtex gustativo, por outro lado, consiste de duas subestruturas: o opérculo frontal no giro frontal inferior do lobo frontal e a ínsula anterior no lobo insular, que jaz dentro do sulco lateral (KANDEL et al., 2021, p. 382, cap. 29).

Por fim, o sistema somatosensorial abarca informações de estímulos como pressão, vibração e temperatura, assim como pela sinalização de dor advinda de alguma forma por estes. Seus receptores jazem na epiderme, vísceras e músculos e transmitem informações para o córtex somatossensorial primário, localizado no lobo parietal, mais especificamente no giro pós-central, dentre diversas outras áreas cerebrais como o córtex insular, envolvido na consciência corporal, dor etc. e a área parietal ventral, relacionada com o córtex pré-motor e memória somatossensorial (KANDEL et al., 2021, p. 382, caps. 18–20).

A atenção é um dos conceitos principais na psicologia cognitiva. Novamente reiterando o fato de que somos sujeitos à torrenciais de estímulos a

cada momento da existência, e que, portanto, é importante sabermos filtrar a quais devemos prestar atenção para nossa sobrevivência e qualidade de vida, tal é o papel da atenção. Ela divide-se entre a vigilância atenta voluntária, relacionada ao foco intencional em aspectos escolhidos do ambiente ou atividade, e a captura atenta involuntária, que é uma resposta automática de estímulos incomuns ou inesperados que nos levam a prestar atenção em algo. A neurofisiologia associada com a atenção remete a diversas áreas afins da percepção, uma vez que pressupõe o uso dos sentidos, todavia é predominantemente associada ao córtex cingulado posterior (GLIMCHER; FEHR, 2014, p. 455–456, 459). A amígdala também possui papel na atenção, uma vez que serve como ponte entre estímulos, emoções, comportamento instintivos do sistema límbico e o uso do córtex, sendo por exemplo vital na ativação de um estado de medo (FEINSTEIN et al., 2011).

O córtex parietal soma um número de sinais diferentes – motores, visuais, cognitivos – para criar um mapa de prioridades do campo visual. O sistema motor utiliza-se deste mapa para escolher alvos para o movimento. O sistema visual utiliza o mesmo mapa para encontrar o local da atenção visual. (KANDEL et al., 2021, p. 596)<sup>119</sup>

Ainda também parcialmente no sistema límbico temos outro conceito crucial para a tomada de decisão: a memória. Como vimos, o hipocampo é relacionado com memória de longo prazo, mas este é apenas um tipo de memória, que por si só se divide em memória explícita e memória implícita. A primeira, também chamada de declarativa, possui duas classes: memória episódica, relativa a experiências individuais, e semântica, sobre conhecimentos e fatos. Implícita, por sua vez, como o nome sugere, refere-se a habilidades aprendidas ou instintivas motoras ou perceptivas, hábitos reforçados e afins. A memória de curto prazo, também chamada de memória de trabalho, consiste no armazenamento temporário de informações usualmente tendo em mente algum objetivo. Ambos os tipos de memória se utilizam de áreas neurofisiológicas diferentes. As de longo prazo requerem o sistema do lobo temporal para seu armazenamento inicial, sendo o acesso das explícitas envolvendo o lobo

---

<sup>119</sup> No original: “*The parietal cortex sums a number of different signals—motor, visual, cognitive—to create a priority map of the visual field. The motor system uses this map to choose targets for movement. The visual system uses the same map to find the locus of visual attention*”.

temporal medial assim como redes frontoparietais também relacionadas com atenção e controle cognitivo. Implícitas, por outro lado, utilizam-se muito dos mesmos mecanismos perceptivos e motores, assim como do corpo estriado, também relacionado com a conexão do córtex com o mesencéfalo, e, novamente, da amígdala (KANDEL et al., 2021, cap. 52, 56).

A resolução de problemas, embora soe semelhante à tomada de decisão, é equivalente a um passo anterior no processo. Retomando a perspectiva computacional da economia comportamental, a resolução de problemas envolve um sentido amplo em que todo pensamento envolve-a quando há algum tipo de bloqueio para um objetivo. Em contraste com a tomada de decisão, como veremos, não há opções ou alternativas cujas quais deve-se priorizar ou descartar. Outra maneira de se pensar é que este conceito surgiu na psicologia, que sempre tendeu a tratar da mente de uma maneira mais generalizada, enquanto a tomada de decisão ocorreu na economia, que direciona uma perspectiva do âmbito mental para escolhas definidas. Também tem sido utilizada amplamente no estudo de inteligência artificial, uma vez que problemas considerados mínimos por nós exige bastante direcionamento de programação.

Um exemplo tradicional é o do xadrez, jogo no qual há uma miríade de maneiras de lidar com jogadas e estratégias a um prazo maior do que apenas uma escolha: o objetivo é vencer, todavia o caminho até lá deve ser resolvido conforme a situação se desenvolve. Evidentemente, este processo envolve praticamente todos os conceitos já tratados: percepção da situação, atenção para os critérios que importam, memória de trabalho no momento e explícita/implícita para auxílio. Ainda assim, há predominância de atividade no córtex parietal, especialmente no giro angular, além da parte superior do sistema parietal posterior, do sulco horizontal intraparietal, do córtex pré-frontal dorsolateral (CPFdl) e do anterior, dentre outros (GROSS, 2020, p. 326–331; ROBERTSON, 2017, p. 11–13, cap. 9).

Por fim, e mais importante para nós neste trabalho, temos a tomada de decisão. Como vimos, ela funciona como uma especialização da capacidade de resolução de problemas, e surge quando estamos em situações específicas e necessitamos escolher alguma opção em detrimento de outras. De acordo com Kandel e Shandlen,

[u]ma decisão é um comprometimento com uma proposição, ação ou plano – dentre opções – baseada em evidências, conhecimento prévio, e resultados esperados. O comprometimento não necessita ação imediata ou quaisquer comportamentos, e pode ser modificado. (KANDEL et al., 2021, p. 1415)<sup>120</sup>

Estas evidências consistem em informações adquiridas por estímulos sensoriais, como os que vimos na percepção, e o conhecimento prévio pela memória. Uma das maiores diferenças com os mecanismos gerais de resolução de problemas é o fator “tempo”, ou o elemento da contingência como Kandel se refere; ou seja, fato de que uma decisão neste sentido pode ter um atraso até julgarmos ser prudente realizá-la ou de fato nunca ocorrer (KANDEL et al., 2021, p. 1392–1393). Outro fator importante para tomada de decisão, e para a maneira que iremos lidar com ela neste trabalho, é o “ruído”. Ao utilizarmos qualquer mecanismo cognitivo, há a presença de ruído: estímulos fora daquilo o qual é seu foco. Ao escolhermos entre uma bebida que gostamos, porém é cara, e outra que não temos muito apreço mas é de bom preço, há uma miríade de fatores além desta simples decisão que são enviadas para o encéfalo. Além daqueles esperados como memória de quanto dinheiro temos no banco, atenção para os valores e nosso desejo no momento e outros, há também aqueles que não possuem nenhuma relação direta com a situação mas podem influenciá-la de qualquer maneira. Se a pessoa que lhe atende parece estar com pressa pois o local está movimentado, se há bastante barulho na volta dificultando a reflexão sobre a situação, se mais cedo você chutou a quina de uma mesa e seu pé está doendo – todos estes estímulos procedem em algum grau junto com o resto das informações acerca das evidências envolvidas na situação na qual a tomada de decisão é requisitada. Isto pode acionar os mecanismos heurísticos como os que vimos acima e gerar os casos em que a pessoa, por exemplo, escolhe algo inconsistente com seu histórico, uma vez que o ruído gerado por estes fatores externos dificultou a análise objetiva desta situação (KANDEL et al., 2021, p. 1393–1404). Embora não seja o momento neste capítulo, não é difícil imaginar que em decisões de cunho moral, que normalmente envolvem deliberações muito mais difíceis do que sobre qual bebida comprar e igualmente muito mais

---

<sup>120</sup> No original: “A decision is a commitment to a proposition, action, or plan—among options—based on evidence, prior knowledge, and expected outcomes. The commitment does not necessitate immediate action or any behavior, and it may be modified”.

ruído sobre fatores emocionais, normas sociais e afins, tal objetividade ou consistência normativa fique ainda mais longe de ser um objetivo alcançável.

Claramente, a tomada de decisão envolve uma compreensão holística acerca de nossas capacidades, oferecendo

uma janela sobre a neurociência da cognição. Ela modela comportamento contingente e operações mentais que estão livres das demandas imediatas do processamento sensorial e do controle da musculatura do corpo. (KANDEL et al., 2021, p. 1415)<sup>121</sup>

Desta maneira, é difícil delinear especificamente que áreas estão envolvidas na tomada de decisão, uma vez que pode depender de uma variedade extrema de fatores. Mas principalmente podemos apontar para o córtex pré-frontal, associado com funções executivas em geral, planejamento, linguagem, personalidade e comportamento social, dentre muitas outras coisas (GROSS, 2020, p. 326, 331–334; KANDEL et al., 2021, cap. 56; ROBERTSON, 2017, p. 258), especificamente sua parte medial; o segmento anterior do cíngulo, um conjunto de fibras nervosas de axônios, associado com emoções, o corpo estriado e o corpo caloso (FURTADO et al., 2022), que

[...] é composto de aproximadamente 190 milhões de axônios e participa não apenas no processo cognitivo mas também na integração de informação entre hemisférios. Sua função permite a ativação de regiões análogas ou sua inibição em algumas instâncias para garantir a efetividade da função cerebral.<sup>122</sup> (FURTADO et al., 2022, tradução nossa)

Quando a tomada de decisão envolve também visão, além das suas áreas previamente vistas associadas, o córtex lateral intraparietal, localizado no sulco intraparietal, associado com sacadas, ou movimentos sacádicos, que são movimentos rápidos dos olhos como quando em análise visual, e com a memória de trabalho (KANDEL et al., 2021, p. 1405–1408). A queda nos neurotransmissores de serotonina afeta negativamente decisões cooperativas,

---

<sup>121</sup> No original: “*Decision-making provides a window on the neuroscience of cognition. It models contingent behavior and mental operations that are free from the immediate demands of sensory processing and control of the body’s musculature*”.

<sup>122</sup> No original: “[...] *is composed of approximately 190 million axons and participates not only in the cognitive process but also in the integration of information between the hemispheres. Its function allows the activation of homologous regions or its inhibition in some instances to warrant the effectiveness of brain function*”.

assim como aumento de ocitocina o fazem positivamente. Os centros de codificação de recompensas como o sistema mesoencefálico de dopamina, assim como neurônios de dopamina e outros no corpo estriado também relacionam-se fortemente com a tomada de decisão também em um nível social, assim como a punição que está associada com núcleo caudado, nos núcleos da base, também associado com aprendizagem e memória (GLIMCHER; FEHR, 2014, p. 72–75).

Para encerrar, retornamos brevemente à teoria de processo duplo, uma vez que convergem diretamente desta compreensão neurofisiológica com o que vimos até então de psicologia cognitiva. Como vimos, processos de T1 são predominantemente automáticos e relacionado ao aspecto emocional. Assim, tais “[s]entimentos fortes como prazer, dor, medo e raiva envolvem uma rede sócio-emocional de de regiões cerebrais, em particular estruturas límbicas e paralímbicas” (WEBER; JOHNSON, 2008, p. 132) e o corpo estriado ventral (REYNA; BRAINERD, 2011, p. 181) como por exemplo, de

(a) circuitos envolvidos na o aprendizado de recompensas e seleção de ações (e.g. núcleos da base, núcleos dopaminérgicos do mesencéfalo) e (b) circuitos envolvidos no estímulo e motivação emocional (e.g. amígdala, hipotálamo, córtex insular anterior) até (c) o [córtex pré-frontal ventromedial].<sup>123</sup> (CLINE, 2017, p. 589)

Assim, o cerne de T1 é relacionado com áreas intuitivo-afetivas – principalmente (a) e (b) como circuitos responsáveis pelo viés da motivação e estados mentais análogos ao desejo. Ao entrar em contato com (c) o córtex pré-frontal ventromedial (CPFvm), tornam-se conscientes e disponíveis para a memória de trabalho. Cline identifica diretamente este processo e funções relacionadas também com juízos normativos, ou como ele os entende, crenças ordinárias intimamente conectadas com a motivação, todavia não de maneira essencial, uma vez que também depende de elementos de T2. Vamos investigar isto um pouco mais a fundo nos capítulos 6 e 7.

Enquanto há maior associação de T1 a mecanismos do sistema límbico, as áreas relacionadas aos processos de T2 correspondem majoritariamente as

---

<sup>123</sup> No original: “(a) circuitry involved in reward learning and action selection (e.g., the basal ganglia, midbrain dopaminergic nuclei) and (b) circuitry involved in emotional arousal and motivation (e.g., the amygdala, hypothalamus, anterior insular cortex) to (c) the [ventromedial prefrontal cortex]”.

áreas neocorticais, em especial o córtex pré-frontal, especificamente o CPFdl e o córtex do cíngulo anterior (CCA). Estes circuitos são relacionados com o controle cognitivo e memória de trabalho (CLINE, 2017, p. 590; REYNA; BRAINERD, 2011, p. 181; WEBER; JOHNSON, 2008, p. 132).

Neste capítulo procuramos cimentar a relação entre psicologia cognitiva, economia comportamental, biologia evolutiva e o papel das neurociências em sua unificação através de conceitos comportamentais associados com os principais aspectos cognitivos da ação relacionados ao tópico geral deste trabalho. Todas estas áreas possuem, cada uma, recursos inestimáveis para o entendimento da ação humana, e com esperança estabelecemos as bases para esta compreensão neste capítulo. No próximo, veremos dados mais práticos de pesquisas que envolvem vários aspectos de cada uma destas disciplinas, aplicando o conhecimento aqui estruturado na compreensão do que este paradigma nos traz acerca da ação humana.

## 5. IRRACIONALIDADE EM DADOS

Nas seções anteriores nos adentramos em diferentes áreas não estritamente filosóficas para entender como elas tratam à sua maneira aspectos que se interconectam com problemáticas que interessam à epistemologia moral e à filosofia de maneira abrangente. A biologia evolutiva, economia comportamental, psicologia cognitiva e neurociências, como vimos, apesar de diferenças e distanciamentos históricos, andam de mãos dadas ao tratar de problemas os quais interessam igualmente todas elas. Como dito na introdução, o objetivo deste trabalho é aproximar as concepções predominantes acerca do processo deliberativo na filosofia do estado-da-arte das pesquisas nestas outras áreas que possam informá-la a fim de oferecer conteúdo prático e real às conjunturas filosóficas. Assim, armados de uma compreensão básica dos preceitos de nossa cognitividade em um contexto neurocomportamental sobre nossos vieses, trabalharemos neste capítulo diversas evidências acerca de como tais afetam o processo de tomada de decisão como um todo.

### 5.1 Nojo e outros fatores externos

O nojo é uma emoção elementar, uma reação predominante em quase todo ser vivo: é uma aversão instintiva, emocionalmente movida, frequentemente associada com raiva ou medo (CALDER; LAWRENCE; YOUNG, 2001, p. 352, 355). Em outras palavras, o nojo consiste em “uma emoção tradicionalmente definida como um estado aversivo que motiva ao afastamento à substâncias ofensivas tais como produtos animais e certas comidas”<sup>124</sup> (BORG; LIEBERMAN; KIEHL, 2008, p. 1529, tradução nossa). Hume identificava que, literalmente, “juízos morais eram comparáveis a juízos de gosto, e [ele] entendia a noção de gosto bem literalmente: as pessoas avaliam algo como moralmente errado quando elícita nojo físico”<sup>125</sup> (ESKINE; KACINIK; PRINZ, 2011, p. 295, tradução nossa).

Até o início do milênio, era comum associar nojo ao medo (CALDER; LAWRENCE; YOUNG, 2001, p. 352), embora ambos se dissociassem nas áreas

---

<sup>124</sup> No original: “*an emotion traditionally defined as an aversive state that motivates withdrawal from offensive substances, such as animal products and certain foods*”.

<sup>125</sup> No original: “*moral judgments were comparable to judgments of taste, and [he] understood the notion of taste quite literally: People assess something as morally wrong when it elicits physical disgust*”.

cerebrais nas quais possuem maior atividade: o medo, relativo à amígdala, e o nojo, relativo aos núcleos da base insular (BORG; LIEBERMAN; KIEHL, 2008, p. 1541) – associados à T1. Porém, ao trabalhar diferentes tipos de nojo, expondo os participantes à memorização de frases sobre (i) patogênicos (doenças e afins), (ii) incesto, (iii) moral não-sexual, além de (iv) frases neutras, Borg et al. (2008) encontraram que a supracitada base insular teve maior atividade apenas em resposta à (ii), enquanto a amígdala foi mais ativa em todos os três exercícios de nojo do que no neutro. Estes resultados são apoiados por outros estudos mais recentes que de Calder *et al.* (2008, p. 1541–1542). Assim, o nojo é um sentimento parcialmente moral assim como parcialmente não-moral, dependendo da situação.

Tomemos dois estudos sobre nojo e alteração de juízos morais: Eskine et al., sobre sensação de gosto amargo (2011), e Schnall et al. (2008) que trata de quatro experimentos com a deliberação ocorrendo, respectivamente, com cheiros ruins, em um quarto sujo, relembrando uma experiência nojenta e em indução audiovisual. No primeiro, os participantes deveriam beber rapidamente uma das três bebidas disponíveis aleatoriamente designadas, amarga, doce ou neutra (água), ao observar uma vinheta sobre algo tido como moralmente desaprovável, como incesto, corrupção e roubo. Ao ingerir bebidas de gosto mais amargo, em oposição à doce ou neutra, a avaliação desaprovadora “piorava” em cerca de 20% a mais do que na ingestão das outras duas, que não parecem ter influenciado significativamente o resultado. (ESKINE; KACINIK; PRINZ, 2011, p. 296)

No segundo estudo, o mesmo padrão se repete (SCHNALL et al., 2008). As mesmas vinhetas do experimento anterior são apresentadas em quatro situações diferentes. No primeiro experimento, um odor desagradável é colocado próximo à alguns participantes, que tal como no estudo anterior, demonstraram um aumento na desaprovação moral das vinhetas em questão em comparação aos participantes que não foram influenciados por nenhum odor diferenciado. (2008, p. 3–5) No segundo experimento, alguns participantes assistiram as vinhetas anteriores, mais algumas um pouco mais “ofensivas”, contendo por exemplo canibalismo e zoofilia, assim como outras mais suaves como sobre políticas públicas etc.; além de alguns questionários com situações morais sem teor desgostante, nominalmente sobre um homem que não devolve uma carteira

que achou na rua, sobre alguém falsificando seu currículo e o dilema do bondinho. Alguns participantes o fizeram em uma sala com odor desagradável, sujeira e lixo espalhado, e o grupo controle em uma sala limpa comum. Outro diferencial foi a auto-avaliação de “consciência sobre o corpo privado”, que orientaria os pesquisadores acerca do quão suscetíveis à nojo em relação ao corpo os participantes teriam e o quanto afetaria os resultados. E, de fato, aqueles com maiores notas de consciência sobre o corpo privado tiveram simultaneamente maior desaprovação por nojo, e o padrão daqueles submetidos a situações estimulantes de nojo continuou superior aos em situação normal; com exceção das vinhetas mais leves, que não foram influenciadas. (2008, p. 5–9)

O terceiro foi semelhante ao segundo apenas em exceção à situação da sala, e que receberam uma escala para designar o estado emocional pré-teste. O fator de nojo fora induzido em alguns ao serem requeridos de lembrar de uma memória que os deixassem “fisicamente doentes ou enjoados do estômago”<sup>126</sup> (2008, p. 9, tradução nossa). Os resultados do experimento 2 foram replicados (2008, p. 9–11).

O quarto experimento altera os fatores: ao invés de focar em nojo, focou-se em tristeza, para ver se a desaprovação moral aumentava com outros sentimentos negativos ou era algo mais regular do nojo. Um vídeo com um banheiro imundo foi mostrado para o fator nojo, e uma cena triste de um filme onde um garoto vê seu pai falecer foram utilizados, além de um trecho emocionalmente neutro de um documentário sobre baleias para controle (2008, p. 11–13). Os resultados foram surpreendentes: não apenas a tristeza não afetou de forma negativa como o nojo, como também fez o caminho contrário em “suavizar” o julgamento (majoritariamente naqueles com maior consciência sobre o corpo privado).

Estes resultados são reveladores. O nojo por elementos não-morais pode, de fato, afetar razoavelmente avaliações morais de forma negativa. Porém, alguns são menos influenciados do que outros, e como Schnall et al. colocam, baseados em mais bibliografias, que “pode ajudar ser lembrado que em algumas situações o sentimento de nojo simplesmente não tem nada a ver com o juízo

---

<sup>126</sup> No original, “*physically ill or sick to the stomach*”.

em questão”<sup>127</sup> (SCHNALL et al., 2008, p. 15). A premissa humeana do nojo como base do moralmente errado mostra-se novamente verdadeira aqui, uma vez que aqueles que não eram suscetíveis à este sentimento através dos métodos utilizados mantiveram-se aparentemente inalterados; e os que o eram claramente incorporaram tal em seus julgamentos. Parece bastante claro, aqui, que a influência de fatores quaisquer que afetem sistemas relacionados a T1 podem gerar crenças morais sem a mediação de T2. Novamente as evidências parecem condizer, ao menos parcialmente, com o modelo de processo duplo.

Um dos meios onde a tomada de decisão é considerada de suprema importância é no meio jurídico – especialmente tratando da deliberação de juízes em casos que podem definir o resto da vida de indivíduos. Os pesquisadores Danziger, Levav e Avnaim-Pesso (2011) acompanharam um grupo de oito juízes em 1.112 sessões de deliberação acerca da liberdade condicional de vários detentos pela duração total de 50 dias em um período de 10 meses. Foi feito um acompanhamento de inúmeros fatores para fins de controle: dos prisioneiros, sexo, idade, nacionalidade, etnia, teor e grau do crime, tempo e comportamento na prisão. Dos juízes, experiência na profissão, sexo, idade, etnia, nacionalidade e pausas para descanso e alimentação. A única variação notável foi no último fator, a alimentação, que acabou por influenciar majoritariamente todo e qualquer julgamento em um padrão quase invariável repetido em todas as amostragens por todo o período. Cada dia analisado consistia em três sessões de julgamentos com dois intervalos para refeições. Quase invariavelmente, no início do dia e após cada refeição, a média de sentenças favoráveis vai de ≈65% e cai de maneira regular a zero até antes de cada intervalo ou o fim do dia. Isto se repetiu em todos os períodos em todos os dias do estudo. Assim, ironicamente, “[d]e fato, a caricatura que justiça é o que o juiz comeu no café da manhã talvez seja uma caricatura apropriada da tomada de decisão humana em geral”<sup>128</sup> (DANZIGER; LEVAV; AVNAIM-PESSO, 2011, p. 6892, tradução nossa). Isto não significa que as decisões nos períodos tardios foram necessariamente “desfavoráveis”: os autores sugerem que “quando juízes realizam julgamentos

---

<sup>127</sup> No original: “*it might help to be reminded that in some situations, the feeling of disgust simply has nothing to do with the judgment at hand*”.

<sup>128</sup> No original: “[i]ndeed, the caricature that justice is what the judge ate for breakfast might be an appropriate caricature for human decisionmaking in general”.

repetidos, demonstram uma maior tendência de julgar em favor do *status quo*” (2011, p. 6892, tradução nossa)<sup>129</sup>.

## 5.2 Neuropatologias cognitivas

Como vimos anteriormente, o sistema nervoso abrange o cerne de toda nossa cognitividade e comportamento. Assim, a investigação patológica de diferentes maneiras em que sua fisionomia ou funcionamento podem ser afetados nos trazem informações vitais acerca dos processos que temos como padrão na tomada de decisão.

### 5.2.1 Atravessando o cérebro de Gage

Um dos mais populares exemplos acerca da relação entre comportamento e neurofisiologia é o icônico caso de Phineas Gage (CORDEIRO, 2011; DAMASIO et al., 1994; GREENE; HAIDT, 2002). Trabalhador em vias férreas, aos 25 anos envolveu-se em um acidente com explosivos em 1848 teve seu crânio perfurado por uma barra de metal de 109cm, que do chão penetrou entre sua mandíbula e cavidade ocular, atravessou seu córtex medial pré-frontal e saiu pelo topo de sua cabeça. O resultado? Um breve momento de confusão. Ainda conseguia falar e mesmo caminhar com a ajuda de seus colegas mesmo após sofrer um acidente que comumente associaria-se ao falecimento instantâneo. Gage sobreviveu quase intacto, exceto pelas cicatrizes no rosto e no topo de sua cabeça, e também pelo principal aspecto de sua fama: sua personalidade mudou quase completamente após o evento. Após meses de recuperação Gage estava fisicamente e intelectualmente recuperado. Todavia, socialmente, não era mais o mesmo. Segundo relatos,

tornou-se irreverente e inconstante. Seu respeito por convenções sociais pelas quais outrora seguia desapareceu. Sua profanidade abundante ofendia àqueles à sua volta. Talvez o mais perturbador fora que afastou-se do seu senso de responsabilidade. Não poderia ser confiado para honrar seus compromissos. Seus empregadores consideravam-no o homem “mais eficiente e capaz” no seu trabalho mas agora tinham que dispensá-lo. [...] Nas palavras de suas amizades

---

<sup>129</sup> No original: “*when judges make repeated rulings, they show an increased tendency to rule in favor of the status quo*”.

e conhecidos, “Gage não era mais Gage”.<sup>130</sup> (DAMASIO et al., 1994, p. 1102, tradução nossa)

De acordo com Greene e Haidt (2002, p. 518, 521) e também Damasio, o acidente danificou em parte áreas próximas do supracitado giro medial frontal e do córtex frontal ventromedial/orbitofrontal. Isto condiz, portanto, com sua reportada mudança de comportamento, já que estas áreas são responsáveis respectivamente pela relação entre emoção, decisões e planejamento, e pela representação do valor de punição/recompensa. Danos nestas áreas são associados, da mesma forma, à um juízo prático empobrecido, e em casos semelhantes, agressão reativa e empatia e conhecimento social diminuídos.

O caso de Phineas Gage não é único. Damasio trabalhou com outros semelhantes, e Greene e Haidt citam outros estudos com pacientes em situações análogas (GREENE; HAIDT, 2002, p. 518), além de outros casos e estudos referentes às mesmas áreas e propriedades (2002, p. 520). É interessante observar aqui uma instância indisputável de um elemento não-moral que influencia o comportamento e decisões morais. Um elemento bastante grande na verdade, nominalmente, uma barra de ferro de mais de um metro de comprimento.

### 5.2.2 Síndrome do Cérebro Dividido

É comum, em um mundo onde o prefixo “neuro-” é atrelado à uma multiplicidade de nomenclaturas independentemente de verificação acadêmico-científica, a imagem do cérebro “dividido” em duas áreas principais associadas com tipos de comportamento diferente: esquerda e direita, razão e emoção, criatividade e lógica, instinto e reflexão. Embora, como vimos, o cérebro funcione de maneira unificada, em oposição à associação prévia de áreas específicas com comportamentos ou traços específicos, há indubitavelmente maior ação de determinadas áreas dependendo do comportamento em questão. Todavia, isto não ocorre da maneira que esta perspectiva popular postula, uma vez que

---

<sup>130</sup> No original: “*he had become irreverent and capricious. His respect for the social conventions by which he once abided had vanished. His abundant profanity offended those around him. Perhaps most troubling, he had taken leave of his sense of responsibility. He could not be trusted to honor his commitments. His employers had deemed him “the most efficient and capable” man in their “employ” but now had to dismiss him. [...] In the words of his friends and acquaintances, ‘Gage was no longer Gage’.*”

diferentes funções ocorrem em locais semelhantes, assim como funções semelhantes em diferentes locais. A maior diferença entre os hemisférios cerebrais é que o esquerdo controla o lado direito do corpo e vice-versa (KANDEL et al., 2021, p. 16). Mas este não foi sempre o consenso, e muito do debate posterior veio dos estudos que veremos nesta subseção.

O que aconteceria, então, se fisicamente dividíssemos ambos os hemisférios cerebrais? Embora soe como uma prática digna de filmes de terror, isto é, de fato uma intervenção clínica existente chamada *calosotomia*. Este procedimento é nomeado de acordo com o corpo caloso, área cerebral que une ambos hemisférios esquerdo e direito, e era utilizado para o tratamento de epilepsia em pacientes que não respondem a medicamentos. Esta área de pesquisa surgiu nos anos 70 a partir do processo cirúrgico que iniciou nos anos 40, e paulatinamente vem trazendo importantes informações sobre nossos processos cognitivos. Os estudos iniciaram com outros animais, mas não demonstravam grandes diferenças entre aqueles com cérebros normais e os calosotomizados (GAZZANIGA, 2005, p. 653–654).

Nos humanos, no entanto, tais são gritantes. Há um bloqueio imediato da comunicação entre os hemisférios no que tange informações de cunho visual, coordenação e percepção espacial, processamento emocional e linguagem. Assim, torna-se evidente que há tarefas que determinado hemisfério tem maior maestria do que outro: o esquerdo especializa-se na linguagem falada e processamento da linguagem escrita. Apesar do hemisfério direito também ter alguma capacidade limitada para a leitura, em contraste especializa-se no processamento visuoespacial, como a detecção de movimento aparente, discriminação de imagens em espelho, relações espaciais etc.; embora ambos hemisférios consigam lidar com tarefas visuais sem componentes espaciais e também com vocabulário, retornando ao caso da linguagem (GROSS, 2020, p. 66). A causalidade também é processada de maneira diferente em ambos, sendo o esquerdo exclusivamente capaz de realizar inferências causais simples e de ordem superior, enquanto o direito é o único de ambos capaz de identificar a natureza causal de eventos de colisão (GAZZANIGA, 2005, p. 654, 656–657).

Mas muito além do que especializações cognitivas específicas, a noção de identidade e autonomia complicam-se ao estudar cérebros calosotomizados, pois cada hemisfério parece ter um comportamento próprio. Um pode estar

agitado e conflitante enquanto o outro calmo, como a mão esquerda, dominada pelo hemisfério direito, estar brincando com um objeto e o hemisfério esquerdo estar confuso quanto a tal comportamento (GAZZANIGA, 2005, p. 656–657). Mais informações interessantes surgem em estudos estimulando visualmente ambas as áreas diferentemente.

Em um dos experimentos, participantes focavam em um ponto numa tela que ocasionalmente expunha determinadas palavras em um curtíssimo período de tempo em algum dos lados, a ponto de serem apenas inconscientemente apreendidas por um dos hemisférios. Após o estímulo com a palavra “chave” à sua esquerda na tela, sendo inconscientemente percebida pelo hemisfério direito, pedia-se que escolhessem entre algum objeto dos variados disponíveis em sua frente, em que invariavelmente escolhiam a chave. Todavia, nenhum participante sabia explicar o motivo de tê-lo feito. Quando a palavra “heart” (“coração” em inglês) aparecia separada com “he” (“ele”) e “art” (“arte”) à esquerda e à direita respectivamente, os participantes dirão “arte”, mas quando solicitados a escolher um dos cartões com a mão esquerda onde cada uma das partes da palavra está escrita, escolhem o “he”, por sua projeção no hemisfério direito.

De maneira mais extrema, mostrando rapidamente metade de um rosto de um idoso e metade de um jovem separando a visão de ambos os olhos, o mesmo ocorre (GROSS, 2020, p. 67). Todavia, as evidências sobre uma independência factual são inconclusivas, além da questão da identidade própria onde o hemisfério direito é especializado na identificação acurada de rostos, enquanto o esquerdo pela velocidade de identificação mesmo com graus bem inferiores de acurácia, que com outras evidências sugerem uma necessidade de integração entre ambos para o desenvolvimento da identidade própria (GAZZANIGA, 2005, p. 657).

Esta ausência do corpo caloso pode também ocorrer no nascimento, chamado de “agênese do corpo caloso” (FURTADO et al., 2022), gerando problemas cognitivos associados com o espectro autista e com aprendizado, mas da mesma maneira que vimos acima, pode também resultar em indivíduos com uma integração normal de informação entre os hemisférios, mostrando que esta conexão através do corpo caloso seja fruto de um período de neuroplasticidade (FURTADO et al., 2022; GROSS, 2020, p. 68). Em uma meta-

análise de 11 estudos sobre esta condição de nascença, verificou-se que a tomada de decisão nos indivíduos é afetada dada a ausência de comunicação entre os hemisférios dos lobos frontais. Tarefas de tomada de decisão relacionadas a comportamento de liderança foram comprometidas, apontando menor ativação da região insular anterior, também associada com funções emotivas, cognitivas e motivacionais além de nojo e outras regiões de emoções homeostáticas como sede, dor e fadiga (FURTADO et al., 2022; MAYER, 2011). Apesar da inteligência geral não ser impactada, as habilidades de tomada de decisão em jogos monetários também eram bastante afetadas, demonstrando grande desvio com a população controle (cérebros sem nenhuma forma de agênese do corpo caloso ou calosotomização) e mais perdas por erros estratégicos (FURTADO et al., 2022).

A unificação entre hemisférios cerebrais, ou falta dela, nos revela bastante acerca de nossas capacidades cognitivas, identidade e processos de tomada de decisão em geral. Claramente há diversos processos heurísticos que dominam a maneira que navegamos o mundo e sobre as quais somos predominantemente inconscientes, como sugerido pela teoria do processo duplo.

### 5.2.3 Psicopatia

Ao tratarmos de moralidade e racionalidade dentro de um viés científico, raros casos são mais interessantes do que o da psicopatia. A alcunha de “psicopata” é amplamente conhecida, sendo utilizada no cotidiano e representada de diversas maneiras na mídia. É frequentemente associada à assassinos da pior estirpe, frios e calculistas, monstros incapazes de sentir remorso ou simpatia pela dor alheia. Embora seja uma asserção parcialmente correta, há muita confusão sobre as motivações e comportamento de psicopatas. Em primeiro lugar, o transtorno psicológico da psicopatia não engendra necessariamente uma pessoa assassina, ou mesmo criminosa. De acordo com Robert Hare, criador do principal teste para identificação da psicopatia, o *Psychopathy Checklist, Revised* (PCL-R) (BABIÁK; HARE, 2007, p. 22–23),

Alguns psicopatas vivem em sociedade e não tecnicamente violam a lei – embora possam chegar perto de fazê-lo, com comportamento que normalmente é bastante desagradável para aqueles à sua volta. Alguns podem levar vidas aparentemente normais, sem machucar pessoas em maneiras que atraem atenção, mas causando problemas

de qualquer jeito em maneiras ocultas econômicas, psicológicas e emocionalmente abusivas.<sup>131</sup> (BABIÁK; HARE, 2007, p. 19–20)

De acordo com o Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM<sup>132</sup>) da Associação Psiquiátrica Americana (APA) dos Estados Unidos da América, manual padrão de categorização de transtornos mentais utilizado no mundo todo por profissionais da psicologia e psiquiatria, a psicopatia não é uma classificação primária. Ela aparece como uma característica específica que pode estar presente e é abrangida pelo transtorno de personalidade antissocial (TPA). Na quinta edição do manual (DSM-5, de 2013), o TPA está classificado nas categorias “Transtornos Disruptivos, do Controle dos Impulsos e de Conduta” e “Transtornos de Personalidade”, e é definido como “uma falha a conformar-se com comportamento ético e dentro da lei, e uma falta de preocupação pelos outros insensível e egocêntrica, acompanhada por desonestidade, irresponsabilidade, manipulação e/ou comportamento arriscado”<sup>133</sup> (AMERICAN PSYCHIATRIC ASSOCIATION, 2013, p. 763, tradução nossa), além de diversos outros traços de personalidade e dificuldades (2013, p. 659, 763–765).

A psicopatia, então, é classificada como uma variante que parcialmente sobrepõe-se, mas não se iguala, ao TPA. Caracteriza-se então como

uma falta de ansiedade ou medo e por um estilo interpessoal ousado que pode mascarar comportamentos mal-adaptativos (e.g. fraudulência) [...] níveis baixos de ansiedade (domínio da Afetividade Negativa) e afastamento (domínio do Desprendimento) e altos níveis de busca por atenção (domínio do Antagonismo). Alta busca por atenção e baixa busca por afastamento capturam o componente da potência social (assertivo/dominante) da psicopatia, enquanto baixa ansiedade captura o componente da imunidade ao stress (estabilidade emocional/resiliência).<sup>134</sup> (AMERICAN PSYCHIATRIC ASSOCIATION, 2013, p. 765, tradução nossa)

---

<sup>131</sup> No original: “Some psychopaths live in society and do not technically break the law—although they may come close, with behavior that usually is very unpleasant for those around them. Some may lead seemingly normal lives, not hurting people in ways that attract attention, but causing problems nonetheless in hidden economic, psychological, and emotionally abusive ways”.

<sup>132</sup> Do original “*Diagnostic and Statistical Manual of Mental Disorders*”.

<sup>133</sup> No original: “a failure to conform to lawful and ethical behavior, and an egocentric, callous lack of concern for others, accompanied by deceitfulness, irresponsibility, manipulateness, and/or risk taking”.

<sup>134</sup> No original: “a lack of anxiety or fear and by a bold interpersonal style that may mask maladaptive behaviors (e.g., fraudulence) [...] low levels of anxiousness (Negative Affectivity domain) and withdrawal (Detachment domain) and high levels of attention seeking (Antagonism domain). High attention seeking and low withdrawal capture the social potency

Em suma, indivíduos psicopatas são “sem consciência e incapazes de empatia, culpa ou lealdade para qualquer um além de si mesmos”<sup>135</sup> (BABIÁK; HARE, 2007, p. 19, tradução nossa).

Neurologicamente, a psicopatia foi recentemente associada também em meta-análise e novo estudo com o corpo estriado de maior tamanho entre adultos, confirmando o foco na busca de estímulos e impulsividade, características altamente associadas com esta área (CHOY; RAINE; SCHUG, 2022). Disfunções na amígdala também são associadas com esta patologia, uma vez que

debilita o condicionamento aversivo, aprendizado instrumental e o processamento de expressões de medo e tristeza. Estas deficiências interferem com a socialização a tal ponto que o indivíduo não aprende a evitar ações que causa dano a outros. (BLAIR, 2006, p. 19)

Isto explica muito do que o DSM define. Há reiteração disto também em outras pesquisas de Robert Blair (BLAIR, 1995; BLAIR et al., 1997), igualmente associando com uma incapacidade de diferenciar entre o que são transgressões morais, como a violência interpessoal, e transgressões convencionais, como falar durante uma aula (BLAIR, 1995).

Outra área cuja disfunção parece influenciar este diagnóstico, segundo Blair, é o córtex frontal orbital, parcialmente associado à amígdala por também estar associado com o aprendizado, mas também de tarefas de controle de resposta, associadas com tarefas sem expectativas claras de recompensa (BLAIR, 2006). A variação da condutância da pele (SCR) é uma atividade eletrodérmica, ou seja, da diferença de potencial detectada na pele, que reflete atividade do sistema nervoso autônomo em estado de atenção e emoções relacionadas (OHME et al., 2009). Em psicopatas, sua ausência em contextos sociais também é comum em comparação com não-psicopatas; embora para perigos tradicionais como quando defrontados com cobras ainda esteja presente (BLAIR et al., 1997).

---

*(assertive/dominant) component of psychopathy, whereas low anxiousness captures the stress immunity (emotional stability/resilience) component”.*

<sup>135</sup> No original: “without conscience and incapable of empathy, guilt, or loyalty to anyone but themselves”.

Outro termo comumente associado aos anteriores é a sociopatia, que também não é formalmente uma categorização psicológica técnica, mas que igualmente se considera parcialmente inserida no TPA. Caracteriza-se em indivíduos que se associam a subculturas que subvertem normas e códigos morais considerados corretos ou padrões, tipicamente normalizando ou exaltando comportamento criminal ou antissocial. Sociopatas, ao contrário de psicopatas, ainda são capazes de apresentar remorso, empatia e lealdade a outrem (BABIÁK; HARE, 2007, p. 19).

Estes dados nos trazem informações importantes sobre o funcionamento geral cognitivo dos psicopatas em comparação aos considerados normais: considerando as áreas neurofisiológicas e aspectos comportamentais, “a capacidade defeituosa para o juízo moral na psicopatia parece não derivar de um déficit racional, mas sim de um déficit no sistema afetivo”<sup>136</sup> (NICHOLS, 2002, p. 301). A psicopatia como aqui caracterizada, a princípio, não ofereceria desafio algum aos eticistas tradicionais. Deontologistas afirmariam que o psicopata instrumentalizaria outrem e agiria conforme um imperativo hipotético, no máximo, em terminologia kantiana. Utilitaristas também afirmariam que suas ações vis de maneira alguma maximizariam o bem para outros além de si mesmo, com custos possivelmente exponencialmente maiores do que o prazer egoísta – para utilizar um eufemismo – que o psicopata teria. Eticistas da virtude, da mesma maneira, afirmariam que seriam indivíduos que jamais poderiam agir de acordo com virtudes que não acidentalmente, uma vez que seu caráter seria incapaz de aprender a prudência necessária.

Todavia, a amoralidade do psicopata levanta uma questão crucial para a análise da agência na esfera ética, especialmente para os racionalistas no que tange o papel da razão. Se ela é um critério central para a moralidade, como pode um indivíduo deficiente em qualificação emocional a ponto de tratar relações sociais de maneira mecanizada – ou seja, que de certa maneira age significativamente mais de maneira racional do que a população “normal” – ser completamente amoral? Afinal, são indivíduos que não se importam com outros, são manipuladores, não sofrem com ansiedade, medo ou stress e são ausentes de empatia. Afinal,

---

<sup>136</sup> No original: “*the defective capacity for moral judgment in psychopathy seems not to derive from a rational deficit, but rather from a deficit to an affective system*”.

“[o] psicopata é frequentemente considerado a epítome do mal, e agora os fatos sobre psicopatas parecem oferecer um problema sério para os meios mais promissores para assegurar a objetividade moral. Assim, os próprios indivíduos cujas ações elicitam nossa mais forte condenação provêem evidência contra teorias que nos permitiriam observar violações morais como objetivamente erradas.” (NICHOLS, 2002, p. 301)

Como vimos no capítulo 2, esta presunção de uma objetividade da razão, de uma neutralidade do processo racional é central para teorias racionalistas. A figura do psicopata parece ameaçar esta pressuposição, no entanto. Algumas objeções podem ser levantadas, como que este representaria um caso extremo, ou que muitos racionalistas ainda atribuem às emoções um papel no mínimo secundário para guiar a razão para o que julgam ser os juízos morais corretos.

Neste capítulo nosso foco é o de compreender as diferentes facetas do funcionamento de nossos mecanismos decisivos associados à moralidade, então não me adentrarei em debates mais aprofundados sobre o tópico. Mas é suficiente dizer que, apesar de ser de fato um exemplo extremo, é o mais próximo do ideal racionalista que temos. Além disso, nos provê uma perspectiva singular acerca do funcionamento da razão e suas interações emocionais. Devemos nos lembrar que a pessoa psicopata não é ausente de emoções: a diferença com indivíduos neurotípicos é a maneira como os sistemas afetivos interagem com nossos processos cognitivos como um todo. Novamente, é importante notar que não há, em termos biológicos, uma separação forte entre áreas associadas com emoções e outras com processos de decisão. Ambos aspectos se integram e compartilham áreas com diferentes capacidades, efeitos e propriedades.

A racionalidade não age sozinha; especialmente no que tange a moral. Mas no que parece ser o máximo que conseguimos conceber e analisar de sua independência, sugere que há mais da moralidade do que este único aspecto. O tema da psicopatia é fruto para inúmeros debates e argumentos interessantes em outros lugares, especialmente nos trabalhos de Blair e nos de Nichols. Para nós, é suficiente esta investigação por hora para compreender a incapacidade da racionalidade como independente, de acordo com o que é possivelmente um dos melhores e únicos casos para esta análise.

### 5.3 Preconceito e cognição

Fear discriminates.  
– Kahan *et al.*, 2007

Um fato inquestionável que infelizmente perdura na sociedade contemporânea é o de discriminação contra minorias. Uma vez que estamos tratando de vieses cognitivos, infelizmente fugiria ao escopo deste trabalho estabelecer os motivos de seu surgimento, o interesse de determinados grupos para sua manutenção, e os impactos em larga escala que o racismo, o sexismo e diversas outras maneiras de segregação e preconceito exercem. Temas de extrema importância e sobre os quais há uma miríade de publicações importantíssimas inclusive de pessoas diretamente afetadas por estes aspectos. Tentaremos, portanto, nos limitar a representar nesta seção evidências das influências de vieses racistas, sexistas e afins, para retornarmos ao debate epistemológico mais profundo no próximo capítulo.

Racismo, sexismo e outras discriminações ocorrem em diversos graus. De uma ação específica individual a políticas públicas, a maneira como diferentes grupos interagem, tratam-se e percebem-se é profundamente influenciada por preconceitos. Como sugerimos até aqui e veremos mais adiante, prescritividade com base em algum juízo de valor associado a aspectos biológicos não possuem embasamento filosófico científico algum, e ainda assim até hoje permeiam todos os níveis da sociedade, intencionalmente ou não, diretamente ou não. Nesta seção, investigaremos como estas características sujeitas à discriminação afetam e são afetadas em um grau cognitivo em todo seu espectro prático: como a percepção e atitude individual é afetada por vieses cognitivos e como políticas públicas afetam e são afetadas por estes mesmos.

Tomemos como um exemplo primário o famoso “Quociente de Inteligência”, ou QI. Como o nome sugere, pretendia-se como método para medir de maneira objetiva e neutra a inteligência individual através de uma miríade de testes lógico-matemáticos. A tentativa arbitrária de criar um critério objetivo para a medição das capacidades intelectuais não é nova, assim como já vimos que estas pretensões de neutralidade são no mínimo problemáticas. Não adentraremos aqui em como um teste deste teor, mesmo sendo confiável e

reprodutível, é incapaz de abarcar a totalidade de inteligências e habilidades possíveis de pessoas ou de evitar que haja treinamento para melhorar o escore.

O conceito de QI disseminou-se na década de 90 com a publicação do livro *The Bell Curve*<sup>137</sup> do cientista político Charles Murray e do psicólogo Richard Herrnstein (SAINI, 2019, p. 103), uma obra contemporaneamente análoga à crítica (todavia mais merecedora) de G. E. Moore sobre Herbert Spencer em que há uma tentativa de justificar preconceito e discriminação com uma fraca teoria pretensa a cientificidade. Nela, associa-se uma diferença de QI a um fator genético, utilizando uma definição arbitrária de raça para denotar que negros seriam biologicamente inferiores a brancos e asiáticos, além de uma retórica reforçando o reconhecimento dessa diferença como necessária e útil para a sociedade (SAINI, 2019, p. 113). Ironicamente, ao sugerir que classe, acesso a universidades, bons empregos etc. não mais sejam definidas mais de acordo com fatores contingentes como onde a pessoa nasceu, em que família e afins, os autores estipulam este critério arbitrário fortemente enviesado racialmente limitando tal acesso com base em pseudociência racista.

Há inúmeras outras críticas possíveis à metodologia do trabalho, dentre elas a falta de revisão por pares dos dados apresentados, do foco em colaboradores de uma única revista de baixo impacto e frequentemente associada com retórica racista neonazista – meio no qual tal trabalho frequentemente é citado positivamente – (SAINI, 2019, p. 103–104), além das supracitadas da redução de todo o espectro de inteligência à um mero número e outras. Todavia, ignoremos por enquanto a ausência de validade prática de uma medida como essa ou o contexto em que ela é apresentada e justificada, e foquemos no por que estas diferenças de QI de fato ocorrem.

De fato, há diferença racial de QI entre brancos, com resultados maiores em média, e negros, menores (NEEDLEMAN, 1990; SCHROEDER et al., 1985). Ignorando o fato histórico de abuso e discriminação de negros, que sem dúvida resultaria numa comunidade em condições desiguais para seu desenvolvimento pleno, temos o aspecto de que esta etnia – muito por resultado destes fatores, no entanto – possui proporcionalmente menor poder aquisitivo e ficam delegadas

---

<sup>137</sup> “A curva em sino” em tradução nossa, referência ao conceito estatístico de distribuição normal, de Gauss ou de Laplace-Gauss em referência aos seus idealizadores, que em gráfico tem um formato semelhante ao de um sino.

a bairros mais pobres. Isto tem resultados práticos além de condições de vida abaixo do ideal em geral, como também a dificuldade de adequação a leis de saúde como por exemplo a que veio após a descoberta dos diversos males que o chumbo pode trazer para o desenvolvimento como um todo. Extremamente difundido em encanamentos e tintas de parede de locais mais antigos, o baixo poder aquisitivo de muitas destas comunidades implica na demora, ou mesmo na incapacidade até os dias de hoje, de retirar ou trocar estas fontes tóxicas (FLYNN; SLOVIC; MERTZ, 1994; YETER; BANKS; ASCHNER, 2020). Sendo bairros mais pobres geralmente com maior foco industrial, poluição e com menor atenção sanitária em geral, as minorias acabam majoritariamente sujeitas a este tipo de insegurança ambiental (FLYNN; SLOVIC; MERTZ, 1994).

Os resultados são claros: crianças que crescem com constante contato com chumbo desenvolvem dificuldades cognitivas, de atenção e de controle emocional (CHIODO et al., 2007; NEEDLEMAN, 1990; SCHROEDER et al., 1985; YETER; BANKS; ASCHNER, 2020). Portanto, comunidades negras são mais afetadas nestes critérios, claramente importantes para a realização de um teste padronizado como este; felizmente, após o afastamento do contato com o chumbo – para aqueles que conseguem – permite que tais sequelas desapareçam (SCHROEDER et al., 1985; YETER; BANKS; ASCHNER, 2020). A este tipo de impacto que acidentalmente ou não afeta apenas uma etnia chama-se de “racismo ambiental”, e é frequente no abandono do cuidado de minorias neste tipo de debate sobre saúde pública (FLYNN; SLOVIC; MERTZ, 1994). Assim, à análise já fraca de Murray e Herrnstein soma-se a incapacidade de entender fatores históricos, sociais e ambientais para sua análise ao inverter causa e efeito através de racionalizações de viés prévio racista.

É interessante que, por serem mais sujeitos a exposição de desastres tóxicos, a comunidade não-branca tem uma percepção significativamente maior de risco deste tipo de evento (FLYNN; SLOVIC; MERTZ, 1994, p. 1101). Além disso, a percepção entre pessoas do sexo masculino e feminino de não-brancos é semelhante, enquanto nas populações de brancos são menores no geral; todavia, pessoas do sexo masculino brancos menosprezam este tipo de risco exacerbadamente, enquanto pessoas do sexo feminino brancas ficam em um meio termo entre as minorias e os brancos masculinos. Isto implica claramente não em uma diferença biológica de um viés perceptivo da realidade ambiental,

mas sociopolítica, onde o controle patriarcal branco da sociedade oferece maior segurança para um determinado grupo em detrimento de outros (FLYNN; SLOVIC; MERTZ, 1994). Na verdade, a presença estatística da menor percepção sobre risco em geral de brancos do sexo masculino é tão comum que possui nomenclatura desde o fim do século passado: o “efeito do homem branco”<sup>138</sup> (FINUCANE et al., 2000; KAHAN et al., 2007; OLOFSSON; RASHID, 2011). Outras atitudes associadas com esta etnia são

[...] maior simpatia à perspectivas hierárquicas, individualistas e anti-igualitárias, maior confiança em gestores da tecnologia, menor confiança no governo, e menor sensibilidade à estigmatização potencial de comunidades por perigos. Estas posições sugerem maior confiança em especialistas e menor confiança em processos sociais de gestão pública. (FINUCANE et al., 2000, p. 170, tradução nossa)<sup>139</sup>

E estas diferenças mantém-se mesmo controlando para nível educacional, variando mesmo dentro de cientistas especialistas da mesma área. O argumento machista que apela para o instinto maternal das mulheres por se preocuparem mais também não dá conta da diferença entre raças e de homens brancos e afrodescendentes. Os estudos de Kahan *et al.* sugerem que, além da destoaância racial, uma cognição motivada semelhante à Simon e Kahneman, todavia baseada no ímpeto de auto-preservação da identidade cultural destes grupos, é o elemento universal em todas as etnias. Assim, o contexto sociopolítico dá origem a diferentes concepções culturais de identidade, que por si motivam nossa cognição a juízos alinhados à nossa percepção de mundo. Isto explica a resistência de aceitação de perspectivas destoantes à de seu grupo e da dificuldade de debate de pautas que interferem com as identidades, especialmente em homens brancos individualistas e hierárquicos, mais tementes à leis que podem afetá-los do que dos riscos que elas pretendem tratar (KAHAN et al., 2007), configurando forte dissonância cognitiva.

O aspecto cognitivo do racismo aprofunda-se mais ainda. Uma das maneiras em que tal se manifesta cognitivamente é como falta de empatia. Uma das maneiras de se medir a empatia de maneira prática é através da observação

<sup>138</sup> No original: “*white male effect*” (WME).

<sup>139</sup> No original: “[...] *more sympathetic with hierarchical, individualistic, and anti-egalitarian views, more trusting of technology managers, less trusting of government, and less sensitive to potential stigmatization of communities from hazards. These positions suggest greater confidence in experts and less confidence in public-dominated social processes.*”

da dor: ao ver outrem experienciando dor, nossas redes nervosas agem como se estivéssemos sentindo-a nós mesmos, variando em intensidade de acordo com personalidade, experiências prévias semelhantes e outros fatores. Primeiramente, pessoas de grupos étnicos brancos e negros realizaram o Teste de Associação Implícita (TAI)<sup>140</sup>, que testa a associação entre palavras positivas e negativas e imagens de negros e brancos, avaliando não apenas a média associativa como também o tempo de demora em tais associações. Posteriormente, analisaram as reações dos participantes observando dor em pessoas do mesmo grupo étnico e de fora. Em ambos os grupos, houve forte redução de conectividade empática para os de fora. Acrescentando também um exemplo de modelos de cor violeta, ou seja, sem denotar etnias conhecidas (pintados como tal utilizando como modelos pessoas de ambas as etnias), verificou-se que apesar de serem de fato notados como menos familiares a ambos os grupos, houve maior ativação empática a eles do que aos de fora do grupo étnico dos participantes. Isto sugere que a falta empática para aqueles de outra etnia advém de efeitos estereótipos e preconceito (AVENANTI; SIRIGU; AGLIOTI, 2010).

Em outro estudo por ressonância magnética funcional e conjunto com análise de sinais de susto, como a resposta de piscar com os olhos em determinado período de tempo, investigou-se o papel da amígdala em avaliações sociais de pessoas do mesmo grupo étnico contra os de fora. Neste caso, o mesmo TAI fora utilizado previamente e posteriormente avaliou-se que há maior estranhamento de brancos ao olharem negros do que o caso contrário. Todavia, ao se olhar figuras famosas como Martin Luther King Jr. e Will Smith da etnia negra e Jerry Seinfeld e Tom Cruise da etnia branca, não houveram diferenças significativas. Isto reforça que há um preconceito cultural tão enraizado que influencia implicitamente brancos a terem receio com sujeitos negros (PHELPS et al., 2000).

Vimos acima que que a resistência de homens brancos a leis que podem afetá-los também existe na esfera jurídica. Afinal, se a alimentação pode afetar a idoneidade dos julgamentos de juízes, o que então o racismo pode causar? O impacto desta disparidade é bem menor conforme mais recentemente

---

<sup>140</sup> No original, "*Implicit Association Test*" (IAT)

analisamos, todavia ainda existe. Considerando dados do início deste século para negros e brancos em situações semelhantes nos EUA, há alguns exemplos como em Connecticut, em que a fiança para negros chega a ser 25% maior do que para brancos; que a nível federal sentenças eram 12% mais longas para negros; e quanto à pena de morte, pessoas negras e/ou responsáveis pela morte de alguém branco eram mais prováveis de recebe-la (RACHLINSKI et al., 2008, p. 1195–1196).

Não é de se surpreender que, em experimento realizado com juízes de ambas as etnias, haviam vieses racistas implícitos em suas decisões judiciais. Todos realizaram o TAI para o registro de seus vieses prévios e executaram julgamentos fictícios com uma miríade de exemplos de criminosos e vítimas de ambas etnias. Semelhante ao experimento sobre empatia, não apenas juízes brancos deram sentenças mais severas para réus negros e mais leves para brancos como os juízes negros também o fizeram com as etnias invertidas; mas apesar de presente em todos os participantes, os vieses implícitos e sua reflexão nos julgamentos eram muito mais significativos nos juízes brancos em todas as versões do experimento. A boa notícia é que, apesar disso, o esforço consciente para lidar com estes vieses parece funcionar para manter a imparcialidade na hora do julgamento (RACHLINSKI et al., 2008). Infelizmente é um esforço ainda necessário, mas com esperança não precisará sê-lo num futuro próximo.

Com o sexismo, ocorre algo bastante semelhante. Em um estudo com 28 professores, cada um analisou um perfil para um programa de jovens talentos cuja única variação era se a pessoa descrita era masculina ou feminina. A versão masculina demonstrou significativa vantagem para obter tal programa em detrimento da versão feminina, deixando claro o forte viés de gênero (BIANCO et al., 2011). Outro estudo ainda analisou o processo de tomada de decisão sobre a avaliação de personalidade de mulheres de acordo com sua altura. Pessoas masculinas e femininas julgaram igualmente que aquelas mais altas foram caracterizadas como mais inteligentes, assertivas e ambiciosas; a única diferença foi que homens julgaram as mulheres mais baixas como mais expressivas, algo que não ocorreu com as mulheres que avaliaram a mesma situação. Este estudo pode parecer supérfluo, todavia traz também que o critério da altura em mulheres influencia diretamente a percepção alheia de seu sucesso

e até mesmo reflete positivamente, ainda que pouco, em diferenças de salário (CHU; GEARY, 2005).

O impacto do viés de gênero, todavia, assim como o racista, vai muito além da percepção pessoal. No meio acadêmico, por exemplo, implica no cerceamento por gênero de determinadas áreas, como física para homens e enfermagem para mulheres, com pouca representatividade do gênero minoritário (BEZERRA; STANISCUASKI; BARBOSA, 2022); em semelhante dificuldade de mulheres para entrarem em posições de poder e influência (SILVA et al., 2021); e também no abuso sistêmico de minorias (BRITO et al., 2022).

Retornaremos aos impactos dos dados vistos neste capítulo e suas implicações filosóficas para o tema em questão no próximo capítulo. Por enquanto, é suficiente dizer que o preconceito não apenas existe, como perpassa grande parte da experiência humana. É inconcebível pensar em uma racionalidade alheia ao conhecimento deste fato que não seja por ele influenciado: independentemente da etnia, classe, gênero ou o que for, vieses implícitos e explícitos moldam nossas decisões em diversos níveis.

## 6. A AMORALIDADE DA MORAL

*Obter progresso com questões difíceis na teoria ética muito frequentemente requer um encontro prolongado com as ciências humanas empíricas.*

– Doris e Plakias, 2008, p. 327, tradução nossa.

O efeito Knobe, nomenclatura posterior ao efeito colateral/secundário identificado por Joshua Knobe (2003), é baseado inicialmente em dois experimentos de pensamento aplicados na prática que revelou importantes fatores acerca de nosso processo de ajuizamento moral, atribuição de intencionalidade e motivação. Neles, transeuntes em Manhattan eram questionados sobre uma de duas variações em dois cenários e inquiridos sobre a intencionalidade de uma ação que tem como efeito colateral/secundário algo positivo, que “ajuda”, ou algo negativo, que “causa dano”. No primeiro, 78 pessoas foram aleatoriamente questionadas sobre um destes dois casos: num deles, o líder de uma empresa é informado que um plano de ação irá trazer lucros, porém ao custo de causar danos ao meio-ambiente. O mesmo coloca que não se importa com os efeitos secundários desta ação, desde que de fato resulte em lucro. Ao responderem, 82% das pessoas afirmaram que o dano à natureza pelo líder da empresa foi intencional. Mas no outro caso, quando altera-se o efeito colateral de negativo para positivo, ou seja, ao invés de danificar o meio-ambiente, ajuda-o – mesmo com a idêntica afirmação do líder de não se importar com as consequências de sua ação desde que gere lucros, 77% dos indivíduos afirma que o líder da empresa *não* teve intenção de ajudar a natureza (KNOBE, 2003, p. 191–192).

E o mesmo se repete no segundo experimento com 42 pessoas onde, ao invés de um ambiente corporativo, postula-se um cenário de guerra no qual um tenente ordena um sargento a levar seus homens para tomar uma colina. No primeiro caso, esta colina representa o fronte de combate, implicando na morte certa de soldados; no segundo, eles estão no fronte e tal colina é um local seguro, em que eles estariam sem dúvida a salvo, evitando quaisquer mortes. Novamente, o líder, no caso o tenente, não se importa com as vidas dos soldados, apenas com o controle tático da colina. Desta maneira, ao responderem sobre a intencionalidade do tenente, o padrão se confirma: 77% atribui culpa à ele no caso que gera dano aos soldados e 70% afirma que ele

não tem responsabilidade por tê-los salvo no segundo caso (KNOBE, 2003, p. 192–193).

Este experimento, embora simples, captura uma faceta vital acerca do debate ético. Imputabilidade faz parte central de debates acerca de responsabilidade e culpabilidade de ações, por exemplo, e neste caso, há um desvio gritante entre a associação acerca da agência do mesmo indivíduo com a mesma motivação – no caso, não se importar com as consequências de seus atos independentemente do cenário em questão. Não iremos nos adentrar neste trabalho sobre as especificidades do debate acerca da responsabilidade moral uma vez que nosso propósito é uma compreensão da esfera deliberativa como um todo; portanto, a apresentação desta pesquisa serve para denotar que tal compreensão é complexa e frequentemente evidencia que nossas teorias clássicas não dão conta de explicá-la.

Em ambos os capítulos anteriores versamos sobre as raízes científicas do estudo da tomada de decisão e dos vieses cognitivos responsáveis por influenciá-lo, e posteriormente mergulhamos em inúmeras evidências das mais diversas formas e contextos em que isto ocorre. Vimos que a figura de que a racionalidade é ultimamente sujeita à fatores externos, ou seja, não necessariamente associados à uma razão fria e calculista, fortifica-se. Assim, o título provocativo deste capítulo de “amoralidade da moral” relaciona-se aos experimentos de Knobe em que a linha que divide o que consideramos “moral” e “não-moral” é bastante tênue, e frequentemente um lado influencia o outro. Embora muitos dos exemplos anteriores já sejam indissociáveis da esfera de fenômenos que associamos com a ética, neste capítulo entraremos propriamente nas evidências e entendimento cientificamente informado sobre como juízos morais são formados e também como são influenciados por fatores que comumente não se associa à esta esfera de estudo. Assim, teremos uma imagem mais clara de como a racionalidade de fato se relaciona o processo de ajuizamento moral com um todo, servindo ao propósito desta tese de entender o quanto da tomada de decisão na ética é racional.

Iremos inicialmente apresentar como conceitos morais de maneira geral entendem-se dentro destes contextos retomando a compreensão neurocomportamental do processo duplo. Posteriormente, visualizaremos os debates que esta perspectiva possui com as principais linhas normativas na ética

e os desafios que trazem a elas; todavia, o foco será em ilustrar a compreensão do nosso processo de tomada de decisão e do arcabouço psicológico envolvido nele dentro do contexto destas teorias. Não caberá a este texto refutar Kant, os eticistas da virtude ou quaisquer outras teorias, apenas ilustrar como uma metodologia experimental no empreendimento filosófico pode nos trazer informações que situam o debate sobre temas morais em outro patamar além do da mera hipótese de poltrona.

### **6.1 Situando a moralidade**

Em uma revisão de mais de 60 livros, artigos e meta-análises de publicações relativas à aplicação de métodos e tecnologias de neuroimagem à temáticas morais, Joshua Greene e Jonathan Haidt (2002) mapearam as áreas de atividade neurológica associadas com estados e processos mentais relacionados à moralidade. Com estas informações, procuraram identificar que áreas cerebrais estão mais comumente ativas quando relacionadas a determinados tipos de situações morais e não-morais, patologias sociais e a que uso normalmente se associam. O resultado é praticamente um breve “manual de introdução” ao viés experimental neurocientífico do naturalismo moral contemporâneo.

Segundo sua prática tabela (GREENE; HAIDT, 2002, p. 4–5), por exemplo, atividade na área cerebral do giro frontal medial, localizada na área central do córtex pré-frontal (a alguns centímetros da testa na junção entre ambos hemisférios cerebrais), é associada com as tarefas morais como as de juízos de perdão, juízos simples e com as tarefas não-morais de atribuir intencionalidade à formas e personagens de desenhos. Em conjunto com o giro angular bilateral, ativam-se de maneira mais presente quando há uma proximidade pessoal envolvida no juízo moral (GREENE et al., 2001, p. 2017). O dano físico nesta área é associado à um julgamento prático empobrecido e agressão reativa, assim como empatia e conhecimento social diminuídos. No geral, a área é associada com a integração da emoção para a decisão e planejamento assim como com uma teoria da mente. Isto, claro, é uma simplificação a fins ilustrativos de todos os fatores envolvidos de vinte artigos citados nesta única entrada na tabela, nominalmente desta região cerebral específica, com um total de oito áreas referenciadas.

Em uma revisão de 75 publicações acerca de populações carcerárias associadas com alto índice de violência e propensão à agressão, Davidson et al. (2000) identificaram importantes fatores que parecem estar associados com este comportamento. Apesar das causas para estes fatores variarem em termos de propensão genética e ambiente de criação, é invariável a predominância da disfunção dos neurocircuitos associados à regulação das emoções, nominalmente, da produção de serotonina relacionada ao desenvolvimento de algumas áreas cerebrais que podem ser familiares à nós: a amígdala, córtex frontal orbital e o córtex anterior cingular (2000, p. 592). Ao retornarmos à revisão de Greene e Haidt (2002, p. 520–521), vemos que ambos a amígdala e o córtex frontal orbital estão relacionados com elementos morais como o processamento sócio-emocional, identificação de figuras morais e realização de juízos morais simples; e disfunções nestas áreas confirmam os achados de Davidson *et al.* com empobrecimento do julgamento prático, além de aumento da reação agressiva, diminuição de empatia e conhecimento social e maior dificuldade com tarefas mais avançadas sobre teoria da mente.

Uma vez que o modelo de processo duplo coloca os processos de tomada de decisão como sistemas-cerne de qualquer funcionamento cognitivo (KAHNEMAN, 2013, p. 24–25), temos harmonia com os achados de Haidt e Bjorklund no desenvolvimento de seu intuicionismo social. Segundo eles,

a mente está sempre avaliando, sempre julgando tudo que vê e escuta dentre uma dimensão de “bom-mau” [...]. Não importa se estamos olhando para rostos de homens, listas de aperitivos ou palavras em turco; o cérebro possui um tipo de medidor (ocasionalmente chamado de “gostômetro”) que está constantemente indo para trás e para frente. (HAIDT; BJORKLUND, 2008, p. 7, tradução nossa)<sup>141</sup>

Há constante funcionamento de ambos em qualquer momento desperto (KAHNEMAN, 2013, p. 24) e portanto, de certa forma, há um fator afetivo (emocional) em toda nossa crença – ou como Greene coloca, “estou inclinado a concordar com Hume [...] que todo juízo moral deve ter algum componente

---

<sup>141</sup> No original: “*the human mind is always evaluating, always judging everything it sees and hears along a “good-bad” dimension [...]. It doesn’t matter whether we are looking at men’s faces, lists of appetizers, or Turkish words; the brain has a kind of gauge (sometimes called a “like-ometer”) that is constantly moving back and forth*”.

afetivo”<sup>142</sup> (GREENE, 2007, p. 64, tradução nossa), independente do contexto inserido.

Ao tratar do juízo moral motivado, ou seja, do juízo normativo, dentro do viés do processo duplo, Brendan Cline apresenta uma extensa bibliografia relativa às áreas neurais específicas envolvidas no processo de tomada de decisão (CLINE, 2017, p. 588–592). Como vimos no item 4.4.3, segundo ele, os processos associados com T1 envolvem circuitos neurológicos presentes (a) na escolha de ações e no aprendizado de recompensas como os núcleos da base e os núcleos dopaminérgicos do mesencéfalo e (b) no estímulo e motivação emocional como a amígdala e o hipotálamo (CLINE, 2017, p. 589). E ao entrar em contato com o CPFvm (c), (a) e (b) tornam-se conscientes e disponíveis para a memória de trabalho e serem expressas em locuções normativas. Estes circuitos relacionados com o controle cognitivo e memória de trabalho, (d) o CPFdl e o CCA, como já vimos, são relacionados com os processos de T2, e “isto explica a conexão íntima entre crença normativa e motivação”<sup>143</sup> (2017, p. 589, tradução nossa). Todavia, é mais difícil do caminho oposto ocorrer, sendo os circuitos de T2 mais raros em processos de valoração, (2017, p. 590) o que apoia a ideia externalista de juízos normativos sem motivação. A bibliografia de patologias nestas regiões e sua influência em testes de mesmo cunho apoia estas conclusões. (2017, p. 590–592)

O entendimento do CPFvm, associado com (c), condiz com os achados de Cline (2002, p. 517–518, 520): é associado à tarefa moral de juízos morais simples, e às tarefas não-morais de, p. ex., recompensa e punição e reconhecimento de faces com raiva (aspectos intuitivos-afetivos). Da mesma maneira, patologias sociais advindas de danos nessa área rendem juízos práticos enfraquecidos, agressão reativa e, especialmente em casos de desenvolvimento, empatia e conhecimento social diminuídos, além de dificuldade com tarefas avançadas de teoria da mente. A amígdala, parte de (b), por sua vez, é relacionada com identificação visual de imagens emotivas, rostos e filmes tristes, e dano nesta área rende o enfraquecimento do juízo social baseado nestas identificações visuais, tendo um “papel crucial no

---

<sup>142</sup> No original: “*I am inclined to agree with Hume [...] that all moral judgment must have some affective component*”.

<sup>143</sup> No original: “*that explains the intimate connection between normative belief and motivation*”.

processamento sócio-emocional”<sup>144</sup> (2002, p. 522, tradução nossa). O CPFdl de (d), também de fato é associado com a memória de trabalho e tarefas cognitivas, assim como o juízo moral de cunho impessoal, que envolve menor fator emocional dado o afastamento do agente (2002, p. 519, 521).

Retornando a sujeitos calosotomizados, indivíduos que não possuem o corpo caloso, parte do encéfalo responsável pela conexão entre os hemisférios, notamos mais aspectos importantes acerca do funcionamento do juízo moral. Já estabelecemos que tal diagnóstico afeta a tomada de decisão, por dificultar a comunicação de ambos os lados de regiões como o córtex frontal, mas como isto se relaciona com a moralidade? Uma outra área associada com a atribuição de intencionalidade e crenças em outros (teoria da mente) é a junção parietal temporal, localizada na área posterior do sulco lateral, responsável por incorporar informações do tálamo e sistema límbico assim como do sistema visual, auditório e somatossensorial. Ela, como a maioria das partes encefálicas, está presente em ambos os hemisférios, todavia sua parte direita parece mais relacionada com juízos morais envolvendo a atribuição de responsabilidade em casos acidentais.

Em uma pesquisa com pessoas completamente calosotomizadas e outras ainda com partes do corpo caloso que permitiriam o contato entre ambos os hemisférios dessa região neural ao menos, além do grupo controle com o corpo caloso intacto, pediu-se para que avaliassem verbalmente alguns cenários como “permissíveis” ou “proibidos”. Os cenários compunham diferentes combinações de agentes fazendo algo para outra pessoa: intencionalmente ou acidentalmente causando dano ou não a outrem, como por exemplo uma pessoa pegando um pó branco que é uma substância tóxica de cientistas que ela acredita ser açúcar e colocando no café de alguém, causando sua morte, ou sabendo que era tóxico, mas não causando sua morte por ser açúcar mesmo, e afins. Predominantemente, ambos os tipos de pessoas calosotomizadas consideraram mais as consequências da ação do que a intenção do agente no caso dos acidentes, julgando como “proibidos” mesmo atos neutros como acidentes numa piscina onde seria proibido crianças brincarem e tendo dificuldade de identificar

---

<sup>144</sup> No original: “*crucial role in social-emotional processing*”.

casos de *faux pas*, ou pessoas passando vergonha socialmente, resultados consideravelmente diferentes do grupo controle.

Este tipo de comportamento condiz com o de crianças, que guiam sua moralidade de maneira muito mais social do que na avaliação própria da crença do agente, mesmo com as pessoas calosotomizadas demonstrando compreensão e maturidade adequada na análise das perguntas. Uma das hipóteses é que, sem o acesso devido a áreas de T1, o juízo moral falha em considerar crenças alheias. O interessante é que alguns destes participantes notaram certa estranheza em sua resposta imediata, o que a bibliografia sugere ser uma característica comum de quem perpassa este tipo de procedimento, e também criaram racionalizações sobre seus juízos justificando-se, por exemplo, na inocuidade de certas situações, aprovando-as independente da intenção do agente nela de causar danos (MILLER et al., 2010). Esta atitude é semelhante também a crianças altamente funcionais no espectro autista que “[...] usam processos como linguagem e raciocínio lógico para compensar déficits em raciocínios de crença ou teoria da mente”<sup>145</sup> (MILLER et al., 2010, p. 2220, tradução nossa).

Como vimos anteriormente, a grande maioria de regiões correlacionadas com a tomada de decisão em geral são praticamente idênticas às da geração do juízo moral – e não são poucas. Crescentemente, a imagem se forma de que nossos processos cognitivos de tomada de decisão e de ajuizamento moral estão intimamente conectados. De fato, “[...] parece agora que muitos dos mecanismos neurais implicados no pensamento moral são majoritariamente os mesmos daqueles implicados no pensamento avaliativo geral e na tomada de decisão”<sup>146</sup> (CLINE, 2017, p. 584, tradução nossa). É inútil falarmos de “cérebro moral” e tentarmos apontar exatamente aqueles presentes na moralidade pois ela possui “poucos, se é que existem, mecanismos neurais próprios”<sup>147</sup> (GREENE; YOUNG, 2020, p. 1005), uma vez que “o cérebro moral é, mais ou menos, o cérebro inteiro, aplicando seus poderes computacionais a problemas

---

<sup>145</sup> No original: “[...] use cognitive processes like language and logical reasoning to compensate for deficits in belief reasoning or theory of mind”.

<sup>146</sup> No original: “[...] it now appears that many of the core neural mechanisms implicated in moral thought are largely the same as those implicated in general evaluative thought and decision-making”.

<sup>147</sup> No original: “few, if any, neural mechanisms of its own”.

que nós, por razões não-neurocientíficas, classificamos como ‘moral’<sup>148</sup> (2020, p. 1005). Assim, a necessidade de uma interdisciplinaridade conforme viemos clamando desde o princípio do texto torna-se mais central do que nunca: não há maneiras de referir-se a um aspecto de nosso funcionamento sem envolver algo que engloba as outras.

## 6.2 Desacordo, virtudes e situacionismo<sup>149</sup>

*“Ethos anthropo daimon”*<sup>150</sup>  
– Heráclito, fragmento 119

No segundo capítulo, figuramos algumas definições de racionalidade tendo por contexto como lidam com o famoso problema do desacordo moral. Aqui, iremos adentrar-nos neste último mais aprofundadamente, investigando diversas evidências que versam sobre sua existência e quais as consequências para teorias morais. Primeiramente, faremos um apanhado geral de como houve a integração entre análises culturais, principalmente antropológicas, com teorias filosóficas, sendo fonte para algumas das posições filosóficas mais relevantes do século XX como de Alasdair MacIntyre e de J. L. Mackie, além de o ser para a discussão acerca da relevância de uma ética das virtudes como condizente com esses aspectos. Finalizaremos esta seção com um debate entre os últimos com psicólogos que defendem a ausência de características duradouras o suficiente para chamarmos de “caráter” – fonte normativa principal dos eticistas da virtude – com base em uma série de experimentos que trabalharemos ao longo do texto, encerrando com a crítica a ambos de Prinz, procurando conciliar ambas as posições.

Apesar do diálogo entre casos exóticos de valoração moral e o interesse filosófico possa ser rastreado até Heródoto e Sexto Empírico (ALFANO; LOEB, 2017, p. 31), não foi até o início do século XX que filósofos tomaram a antropologia em mãos e trabalharam este problema diretamente, sem depender de antropólogos sem treinamento filosófico. Edward Westermarck, William

<sup>148</sup> No original: *“the ‘moral brain’ is, more or less, the whole brain, applying its computational powers to problems that we, for nonneuroscientific reasons, classify as ‘moral’”*.

<sup>149</sup> Sou extremamente grato ao prof. Dr. João Hobuss por toda a discussão neste capítulo.

<sup>150</sup> “O caráter do ser humano é o seu destino” (Cf. DE OLIVEIRA, 2013, p. 83)

Graham e Franz Boas com suas estudantes Ruth Benedict, Melville Herskovits e Margaret Mead foram os primeiros a estabelecerem o conceito de relativismo moral, “a muito grosso modo a perspectiva que a verdade moral depende nas crenças morais de vários grupos ou indivíduos”<sup>151</sup> (ALFANO; LOEB, 2017, p. 32). Todavia, os dados e interpretações eram relacionados de maneira falha, sem a devida nuance. Richard Brandt e John Ladd, já na metade do século, aprenderam com os erros de seus antecessores, que

apesar de documentarem notável diversidade avaliativa, não apoiaram inferências confiantes acerca do desacordo moral sob condições ideais, amplamente por frequentemente prover orientação limitada sobre o quanto do desacordo moral pode ser rastreado ao desacordo acerca de temas factuais que não são morais por natureza, tais como aqueles relacionados com perspectivas religiosas ou cosmológicas.<sup>152</sup> (DORIS; PLAKIAS, 2008, p. 314)

Assim, com maior rigor científico e filosófico, produziram evidências com maior atenção a detalhes e correlações baseados nas convivências com comunidades nativo-americanas Hopi e Navajo, respectivamente. Brandt, especificamente, considera o caso de crianças hopis “brincarem” com pássaros amarrando-os em suas mãos – brincadeira que normalmente dura pouco pois os pássaros falecem com a violência envolvida. Os hopis entrevistados não negaram que os animais sofriam e não acreditavam que eram recompensados no pós-vida; nem que eram mecânicos ou insensíveis – de fato, acreditavam que os animais eram “mais próximos à humanidade do que o homem branco comum” (2008, p. 314). Todavia, Brandt acaba caindo em generalizações muito amplas, como o supracitado “homem branco comum”. A permissibilidade de sofrimento animal na sociedade contemporânea não é tão difundida quanto esta generalização, vide as leis de proteção animal e as inúmeras ONGs e grupos de proteção animal, ambiental ou meramente vegetarianos e afins – para citar alguns exemplos do problema desta linha de arguição.

---

<sup>151</sup> No original: “*very roughly the view that the moral truth depends on the moral beliefs of various groups or individuals*”

<sup>152</sup> No original: “*although documenting remarkable evaluative diversity, do not support confident inferences about moral disagreement under ideal conditions, in large measure because they often give limited guidance regarding how much of the moral disagreement can be traced to disagreement about factual matters that are not moral in nature, such as those having to do with religious or cosmological views*”.

Mas apesar de todas as falhas inerentes à novas metodologias e intersecções entre diferentes áreas de pesquisa, estes grupos estabeleceram algumas das principais bases para o diálogo entre ciências empíricas e teoria filosófica – segundo Prinz, estabelecendo o relativismo moral de aspecto descritivista como um fato empírico (2007a, p. 173–174). A partir destas análises, então, foram desenvolvidos métodos mais eficazes, ferramentas estatísticas e demográficas, instrumentos de alta precisão e conhecimento legal que receberam exponencial progresso desde a época de Brandt, manifestando-se no campo da psicologia cultural, que parte para analisar processos emocionais e cognitivos de diversas culturas (DORIS; PLAKIAS, 2008, p. 316).

Outro desenvolvimento do diálogo crescente entre antropologia e filosofia foi o pensamento de Alasdair MacIntyre. Apesar de ter raízes predominantemente continentais e posteriormente abraçar um posicionamento mais associado com Tomás de Aquino, sempre teve um ímpeto muito semelhante ao espírito da filosofia experimental. No prefácio da edição de 2007 de seu livro *Depois da Virtude*<sup>153</sup> de 1981 revela o seguinte:

[e]ste livro emergiu da reflexão extendida sobre as insuficiências de meu próprio trabalho prévio em filosofia moral e de uma crescente insatisfação com a concepção de 'filosofia moral' como uma área independente e isolada de investigação. Um tema central de [meu] trabalho anterior [...] fora que precisamos aprender da história e antropologia da variedade de práticas, crenças e esquemas conceituais morais. A noção que o filósofo moral pode estudar os conceitos da moralidade meramente refletindo, numa poltrona como fazem em Oxford, no que ele ou ela e aqueles à sua volta dizem é improdutivo.<sup>154</sup> (MACINTYRE, 2007, p. xvii)

Esta passagem poderia estar presente em qualquer trabalho de filosofia experimental, tamanho seu reconhecimento da importância interdisciplinar no estudo da ética. Todavia, nesta obra, procura retornar às raízes de outro autor que, como comentamos, condiz com o intuito desta escola de pensamento: Aristóteles.

---

<sup>153</sup> No original: “*After Virtue*”.

<sup>154</sup> No original: “[t]his book emerged from extended reflection upon the inadequacies of my own earlier work in moral philosophy and from a growing dissatisfaction with the conception of ‘moral philosophy’ as an independent and isolable area of enquiry. A central theme of much of [my] earlier work [...] was that we have to learn from history and anthropology of the variety of moral practices, beliefs and conceptual schemes. The notion that the moral philosopher can study the concepts of morality merely by reflecting, Oxford armchair style, on what he or she and those around him or her say and do is barren”.

MacIntyre procura desenvolver uma ética das virtudes adaptada para o mundo contemporâneo, participando do retorno desta linha de filosofia moral na segunda metade do século XX junto com Elizabeth Anscombe, Philippa Foot e Bernard Williams, procurando a normatividade ao invés da descritividade da metaética naturalista. Não iremos nos adentrar aqui nem em como desenvolvem cada um sua versão desta teoria nem em nenhum aprofundamento significativo da ética das virtudes como um todo; no entanto, o debate sobre o caráter que tal linha ética andou tendo com a psicologia social nas últimas décadas é de significativa importância para este trabalho. Isto se dá por que o caráter não apenas é um conceito crucial para a ética das virtudes, como também é um conceito psicológico analisável empiricamente e que pode ter influência na maneira que agimos, da mesma maneira que diversos outros fatores como os vários tratados anteriormente podem fazê-lo.

De uma maneira bastante generalizada, a virtude na ética das virtudes trata-se de uma disposição do caráter, ou seja, são “capacidades racionais ou habilidades de racionalidade prática”<sup>155</sup> (PRINZ, 2009, p. 122) para se agir conforme um fim, a saber, a *eudaimonia* como vimos com Putnam, predominantemente traduzida simplesmente como “felicidade”. Ou seja, há a manutenção de traços de personalidade que se desenvolvem através da prática de ações virtuosas, aquelas que representam uma ação prudente própria à situação entre o exagero ou inação, que constituem ações viciosas. Este exercício da sabedoria prática, ou prudência, é chamado de doutrina da mediedade. Patricia Churchland apresenta-a de maneira bastante simplificada como uma

Média de Ouro (não confundir com a Regra de Ouro: ‘faça aos outros como queira que façam para ti’). A Média de Ouro nos aconselha que o meio termo é geralmente bom: não se deve ser nem temeroso nem tímido, mas apropriadamente corajoso; nem mão-fechada nem mão-aberta, mas apropriadamente generoso; nem totalmente indulgente nem completamente abstinente, mas moderado; e assim por diante. Como ser apropriado não é algo definido pela aplicação de uma regra, de acordo com Aristóteles; requer sabedoria prática, adquirida através da experiência e reflexão.<sup>156</sup> (CHURCHLAND, 2008, p. 113)

<sup>155</sup> No original: “*rational capacities or practical reasoning skills*”.

<sup>156</sup> No original: “*Golden Mean (not to be confused with the Golden Rule: “Do unto others as you would be done by”). The Golden Mean counsels us that the middle ground is generally good: one should be neither reckless nor timid, but appropriately courageous; neither tightfisted nor openhanded, but appropriately generous; neither wholly indulgent nor utterly abstemious, but*

Ao exercer tal prudência, o indivíduo então age virtuosamente, podendo alcançar a felicidade. Haveria então uma responsabilidade moral do agente baseada em atitudes reativas características de seu caráter virtuoso. A psicologia moral do agente, portanto, deve pressupor traços duradouros de personalidade, e uma certa universalidade deste “bem-estar” e das virtudes correspondentes. Segundo os situacionistas, cuja própria teoria veremos a seguir, o caráter funciona da seguinte maneira para os eticistas da virtude: para John Doris, há a necessidade de um “globalismo” do caráter que deve ser (i) consistente dentre uma miríade de condições, (ii) estável em repetições destas, e (iii) cujos traços devem estar integrados com outros relacionados. G. Harman define que tais traços devem ser “abrangentes” (*broad-based*), no sentido semelhante à consistência (i) de Doris, que sejam aplicáveis em vários contextos distintos e não apenas em conjuntos de situações particulares. Merrit fala também de que pressupõe uma autossuficiência motivacional, ou seja, não apenas é capaz de identificar o curso de ação virtuoso, mas é por si só motivador da ação. Há diversos problemas nas interpretações dos situacionistas sobre a ética das virtudes, e entraremos em algumas críticas posteriormente, mas para nosso fim pragmático de trabalhar apenas outra faceta da tomada de decisão nos é suficiente por enquanto.

Os situacionistas, então, predominantemente da área da psicologia social como Doris, Harman, Merrit e Alfano, defendem que não há evidências para tal coisa como o caráter. Segundo eles, elementos da situação em que o agente está inserido são única e exclusivamente responsáveis por definir seu curso de ação. Além disso, como sugerimos através deste trabalho, defendem que há uma miríade de fatores externos envolvidos na deliberação que pesam mais do que quaisquer traços de personalidade da pessoa que está para tomar uma decisão. Os traços de personalidade que temos, que seriam a base do caráter, se são eficazes causalmente, a ponto de terem efetividade prática na ação, não são globais, variam demasiadamente para serem considerados virtuosos; ou vice-versa, se são globais não possuem a eficácia causal necessária para serem relevantes. E ao contrário dos eticistas da virtude, os situacionistas trazem

---

*moderate; and so forth. How to be appropriate is not something that is settled by applying a rule, according to Aristotle; it requires practical wisdom, acquired through experience and reflection”.*

inúmeras evidências científicas para apoiar seu caso. Veremos a seguir quatro das principais.

O primeiro e mais famoso caso é o experimento do choque de Milgram de 1974, onde uma pessoa opera um mecanismo que supostamente dá choques em outra pessoa, que não a vê, caso erre determinadas perguntas do experimentador. Esta vítima, todavia, é um ator, e o experimentado real é a pessoa operando o botão. Originalmente, aproximadamente 65% dos participantes administraram o que seriam doses ostensivamente letais do choque, e 100% dos participantes deram doses altíssimas – mesmo quando o “alvo do teste” já não mais gritava. Foi amplamente utilizado para analisar como pessoas comuns poderiam ser coagidas por autoridades a ferir outrem.

O segundo é o de Darley e Batson de 1973, também amplamente conhecido, no qual alunos a caminho de uma palestra importante deparam-se com uma pessoa que requer ajuda: 63% dos que estavam adiantados auxiliaram a pessoa, todavia, no grupo de alunos que estavam atrasados, apenas 10% o fizeram. Em terceiro, temos o experimento de Isen e Levin de 1972, em que pessoas que encontraram uma moeda na cabine telefônica eram significativamente mais inclinadas a ajudar um transeunte que deixa cair seus papéis do que outros que não encontravam tal moeda. Por fim, em um estudo de 1928 sobre honestidade com 11.000 crianças por Harthshorne e May, encontrou-se apenas baixa correlação entre 33 comportamentos nos quais a honestidade poderia ser medida, sugerindo que não é um traço globalmente presente de maneira igual. Todos estes estudos, dentre muitos outros, condizem com as hipóteses dos situacionistas: caso houvessem traços globais estáveis de personalidade eles seriam consistentes independentemente de fatores alheios à situação nos casos apresentados. Todavia, não foi o caso (PRINZ, 2009).

Poderíamos pensar que o debate se encerrou aqui, com os situacionistas saindo vitoriosos após a apresentação destas evidências. Mas embora as críticas dos eticistas da virtude sejam mais abstratas e filológicas, atendo-se à detalhes técnicos das definições de caráter das diversas formulações destas teorias éticas através da história, as mais significativas sem foram dos próprios experimentalistas. Jesse Prinz critica os situacionistas e os eticistas da virtude respondendo e elaborando críticas a ambos através dos métodos dos próprios situacionistas: com pesquisa empírica. Não nos adentraremos aqui em todos os

aspectos do debate, porém focaremos nas argumentações de Prinz que possuem maior relevância para nós neste trabalho.

Em primeiro lugar, traços estáveis existem. Como sugere João Hobuss, “você acha que as pessoas que você conhece atuam por situação ou com base em quem são?”<sup>157</sup> Prinz defende que de fato existe uma espécie caráter, trazendo diversas evidências para apoiar esta perspectiva; todavia um caráter *empírico*, que, como veremos, compromete não apenas o situacionismo como a própria ética das virtudes. Se pensarmos em grupos como hippies e metaleiros, conservadores e socialistas, interiorano e urbano e exemplos afins como concepções de si, atitudes políticas e mesmo regionais, temos traços de personalidade que se mantém obedecendo todos os critérios estipulados pelos situacionistas (PRINZ, 2009). Voltaremos então brevemente para a análise antropológica contemporânea além de reproduções de estudos já citados a fim de ter uma concepção melhor do que compreende essa noção revisada de caráter.

Richard Nisbett e sua equipe, por exemplo, investigaram o fenômeno da “cultura da honra” nos Estados Unidos. Tal conceito é baseado nos preceitos de que homens, majoritariamente brancos não-hispânicos e do sul dos EUA, possuem uma “honra” inerente, cujo insulto implica diretamente no dever da retaliação. Há uma miríade de estudos sobre este fenômeno, concluindo que membros deste este estereótipo (i) são muito mais inclinados do que outros homens brancos de outras regiões a se envolver em homicídios resultantes de discussões (embora não de roubo ou outros crimes), (ii) acham a violência “extremamente justificada” em uma série de casos, e a falta de seu uso dada provocação torna a pessoa “menos homem”, (iii) esta região permite maior violência na auto-defesa e na defesa de propriedade que estados do norte, (iv) no envio de currículo com carta de intenções de um trabalhador exemplar fictício de 27 anos cujo único crime foi ser condenado pelo homicídio de alguém que teria se gabado de ter relações com sua mulher, onde empregadores do sul não apenas eram mais propícios a contratar tal trabalhador, como uma das respostas inclusive dizia que qualquer um poderia ter agido como ele e que não deveria ser um problema – enquanto carta alguma de empregador do norte fora tão

---

<sup>157</sup> Fala (parafrazeada) do prof. Dr. João Hobuss na disciplina de “Virtude e Felicidade” no segundo semestre de 2018 do PPG em Filosofia da UFPel.

simpática; e por fim, (v) em um exame de cortisol e testosterona, participantes dentro deste grupo obtiveram resultados muito maiores e significativos do que outros a um insulto por um experimentador que lhes dá um encontrão no corredor – tudo devidamente programado e controlado sem o conhecimento dos participantes (2008, p. 317–319).

Teoriza-se que tal comportamento advém da herança de povos cuja economia era centrada em rebanhos, em contraste com aqueles cuja economia era baseada na agricultura. Segundo as pesquisas de Nysbett e outros, povos centrados em cuidar de rebanhos (como muito era no sul dos EUA), pela ausência do Estado em manter a lei, deveriam ficar atentos para que outros não roubassem seu produto, gerando maior desconfiança e cultuação de personalidades vingativas. Povos centrados na agricultura, por outro lado, não sofreriam com a constante competição e tenderiam a se inclinar mais a estratégias pacíficas de convívio, não fomentando então esta “cultura da honra” (2008, p. 316–317).

Outro caso interessante na ética contemporânea acerca do desacordo moral é o do magistrado e da multidão: “deve a polícia indiciar e punir um único bode expiatório inocente para prevenir protestos que levarão a substantiva destruição de propriedade e perda de vidas?”<sup>158</sup> (2008, p. 322). Tecnicamente, os utilitaristas deveriam afirmar positivamente: o sacrifício de um inocente que irá prevenir maiores mortes e sofrimento de uma quantidade muito maior de pessoas deveria ser moralmente justificado. Todavia, não é o caso. Utilitaristas em geral, alguns mesmo dos mais ferrenhos, como Bloomfield, Shafer-Landau, e Smart, negam veementemente que este seja o caso – o sacrifício não é justificado. Evidências apontam mesmo que há uma tendência forte de se condenar o sacrifício. Anscombe mesmo chega a afirmar que quem defende o sacrifício nesta situação possui uma “mente corrompida” (2008, p. 322–323).

Poderia parecer que temos aqui nossa primeira instância de acordo moral, e ainda que parece trespassar fortes barreiras teóricas. Todavia, não é o caso. Na aplicação de Nysbett, e posteriormente Peng *et al.*, com populações orientais, o resultado muda drasticamente. Participantes chineses foram significativamente mais propensos a aceitar o sacrifício individual em prol do bem maior da

---

<sup>158</sup> No original: “should the police prosecute and punish a single innocent scapegoat to prevent rioting that will lead to substantial destruction of property and loss of life?”

comunidade (2008, p. 323–325). Teoriza-se que esta diferença está relacionada a traços de auto-concepção individualistas no ocidente, “que foca nos atributos pessoais que podem ser caracterizados com referência limitada ao contexto social”<sup>159</sup> (2008, p. 323, tradução nossa) em contraste a uma auto-concepção coletivista oriental que entende as pessoas em relação ao seu papel em grupos.

Há também várias reproduções posteriores dos experimentos de Milgram. Ao controlar todos os outros fatores, os resultados parecem alterar-se somente no quesito nacionalidade, gerando uma diferença praticamente quatro vezes maior entre alemães, obedientes 85% das vezes, contra mulheres australianas, obedientes apenas 16% das situações. No experimento da moeda em uma cabine telefônica, o percentual de pessoas que ajudaram após encontra-la foi de 31% em Nova Iorque, situada em um país rico, contra 100% dos brasileiros no Rio de Janeiro, em um país mais pobre, dentre outros exemplos. Estes experimentos sugerem também que participantes de países mais pobres possuem atitudes mais coletivas, e de países mais ricos atitudes mais individualistas.

Apesar do situacionismo não se demonstrar uma teoria que dê conta adequadamente das evidências acerca do caráter, a ética das virtudes também deixa a desejar. A crítica de Prinz pressupõe-se na característica da ética das virtudes de basear-se em uma natureza humana, e não na normatividade. Há exceções, mas Prinz baseia-se no argumento aristotélico enquanto padrão de argumentação, onde então há uma tentativa de embasamento no caráter humano, e não na ação *per se*. Tal posição postula determinados traços de caráter específicos que hipoteticamente seriam mais adequados que outros para a ação, traços estes que seriam virtuosos. Prinz defende que as evidências empíricas, como concorda Churchland, apoiam a perspectiva que a ação é mais dependente do contexto em que ela se dá do que traços de caráter específicos (CHURCHLAND, 2008, p. 128). Algumas objeções seriam de que a virtude seria difícil de alcançar, o que não é empiricamente relevante, e que tal também seria referente ao raciocínio prático, e não como reação instantânea, além de que crenças errôneas em outros fatores, por exemplo nos cientistas envolvidos no experimento, etc., levariam a ações viciosas. Estas postulações também

---

<sup>159</sup> No original: “*which focuses on personal attributes that may be characterized with limited reference to social context*”.

contradizem evidências que corroboram a visão casuística da ação moral, como por exemplo o famoso experimento citado acima de Milgram.

Na verdade, de fato, há estudos na psicologia que tratam de traços de caráter. Porém, dos 5 a 9 traços de caráter presentes majoritariamente em uma meta-análise, variando a quantidade conforme diferentes concepções, poucos são condizentes com a perspectiva aristotélica. Além disso, a ética das virtudes contemporânea parece ignorar tais estudos. Assim, o problema da ética das virtudes jaz em que seus postulados e justificações são muito problemáticos ao correlacioná-los com evidências empíricas da natureza humana. A variedade é um fator crucial para o paradigma evolucionista – que como vimos permeia o paradigma científico, e, portanto, naturalista contemporaneamente –, e concepções populares de virtude variam de acordo com diferentes povos. Afinal, “Alasdair MacIntyre [...] descreve como a lista de virtudes de Benjamin Franklin difere de Jane Austen, cuja lista difere de Aristóteles, cuja lista, devemos notar, difere de Confúcio”<sup>160</sup> (ZIMMERMAN, 2010, p. 25): a própria ética das virtudes não consegue manter uma constância própria de formulações de virtudes, sendo evidentemente sujeita, também, às variações culturais.

E ainda que tomemos traços virtuosos como moralmente bons com base em nossos sentimentos frente a tais, a ética das virtudes ainda assim deveria dar lugar a teorias emotivistas para seu tratamento devido. Churchland até aceita a perspectiva que, conforme Aristóteles, uma vida virtuosa geraria um governo virtuoso, mas que as definições de virtudes, vida virtuosa, etc., seriam vagas em demasia para sua validade. Os sentimentos, por sua vez, segundo Prinz, não apenas são a base dos traços de personalidade, como são o alicerce de qualquer normatividade. Traços de caráter, como vimos, por sua vez, são inconstantes, irregulares e variáveis demais, e portanto “o problema com a ética das virtudes não é que traços de caráter não existem, mas sim que eles não podem formar a fundação de uma teoria normativa”<sup>161</sup> (PRINZ, 2009, p. 142). Traços de caráter surgem a partir de ações, e não o contrário.

---

<sup>160</sup> No original: “Alasdair MacIntyre [...] describes how Benjamin Franklin’s list of virtues differed from Jane Austen’s, whose list differed from Aristotle’s, whose list, we might note, differed from Confucius’s”.

<sup>161</sup> No original: “The problem with virtue ethics is not that character traits do not exist, but rather that they cannot form the foundation of a normative theory”.

Talvez a informação mais importante desta seção é que, apesar dos pesares, há uma constância de caráter. Até então, temos visto inúmeras evidências que apoiam a influência inconsciente de inúmeros fatores completamente alheios a nossa racionalidade na tomada de decisão, sugerindo quase nosso descontrole próprio sobre nossa capacidade de ação. A existência de traços de personalidade constantes permite que tenhamos algum critério de estabilidade cognitiva nos inúmeros processos de tomada de decisão que perpassamos na vida.

### **6.3 Experimentos e Pensamento**

Se você tivesse que escolher entre salvar cinco pessoas ao puxar uma alavanca desviando um bondinho descontrolado indo na direção delas para que mate apenas um trabalhador no trilho anexo, ou salvar o trabalhador e matar as cinco pessoas, o que você faria? Todos aqueles minimamente familiarizados com estudos em ética conhecem os famosos “dilemas do bondinho” (*trolley problems*, no original) foram formulados originalmente por Philippa Foot, que procurava defender sua concepção de ética das virtudes neoaristotélica com sua formulação, e foram posteriormente expandidos por Judith Thomson para trabalhar sua posição contra consequencialistas. Mas uma breve pesquisa bibliográfica mostra que uma multitude de autores se utilizaram destes dilemas para endossar ou criticar diferentes teorias até atualmente, sendo amplamente popular a ponto de ser conhecida fora da academia como em séries de televisão.

Joshua Greene, apesar de abertamente consequencialista, resolveu aproximar-se deste problema em seu viés experimental. Em uma primeira instância, realizou dois estudos de enquetes acerca de diferentes formulações deste dilema acompanhadas de imagenologia de ressonâncias magnéticas funcionais dos participantes (GREENE et al., 2001). Posteriormente, somou sua experiência nestes estudos à uma análise bibliográfica para uma investigação mais aprofundada do que os resultados podem nos dizer sobre nossas capacidades morais e sobre alguns posicionamentos tradicionais filosóficos (GREENE, 2007). Os resultados foram, no mínimo, interessantes.

Uma diferenciação conceitual importante que o autor faz é quanto à “pessoalidade” de uma ação moral, ou seja, qual a proximidade física do agente moral em determinada ação. Uma ação impessoal é como a do início desta

seção, em que há uma distância entre o agente e o(s) objeto(s) de sua ação – no caso, está distante das pessoas cujas vidas terá que deliberar sobre ao manejar uma alavanca à distância. Uma ação pessoal, por outro lado, envolve a ação direta do agente sobre os objetos: muda-se a situação, por exemplo, para que ao invés de realizar a ação à distância, envolva o agente empurrar ou não alguém de uma ponte sobre os trilhos, custando a vida desta pessoa, para que o bondinho pare.

Objetivamente, ambas as situações envolvem o dilema quanto ao sacrifício de uma pessoa para salvar cinco. Todavia, é consenso ser moralmente aprovável este sacrifício no primeiro caso, enquanto no segundo, não. (2007, p. 41–42) O único fator variável nestes casos é o da personalidade, da proximidade entre o agente e a possível vítima. Para entendermos melhor esta relação, é necessário que observemos os processos neuronais envolvidos.

No caso dos dilemas pessoais, como o exemplo da ponte, há uma maior atividade quase instantânea em três áreas relacionadas à emoção: no córtex posterior cingular, no córtex pré-frontal medial e na amígdala, assim como no sulco temporal superior, associada com cognição social em humanos e outros primatas. Já nos dilemas impessoais, como no primeiro exemplo, envolveram maior atividade em duas áreas mais cognitivas, ou seja, associadas com raciocínio e cálculo passivo: no córtex pré-frontal dorsolateral e no lobo frontal parietal. E ao contrário do caso da ponte, houve uma demora maior no tempo de resposta dos participantes (2007, p. 43–50; GREENE et al., 2001, p. 2106–2107; cf. GREENE; HAIDT, 2002, p. 520–521). O fator da proximidade, então, é traduzido em um maior engajamento emocional com o alvo da ação, modificando a escolha na decisão em dois cenários que, sob qualquer outra perspectiva, são idênticos – mais um elemento que, ao menos sob a perspectiva tradicional filosófica, nunca fora relevante para tratar das ações morais, ou seja, não-moral.

Ainda relevante quanto à tradição filosófica, é como, historicamente, os autores se posicionariam em situações semelhantes. Autores deontologistas, segundo Greene, como Kant e Tomás de Aquino, argumentariam “que é errado causar danos a alguém como meio de ajudar outra pessoa”<sup>162</sup> (GREENE, 2007, p. 42, tradução nossa). O caso e resposta geral do dilema da ponte, portanto,

---

<sup>162</sup> No original: “that it is wrong to harm someone as a means to helping someone else”.

seria o que mais se aplicaria à suas teorias. Por outro lado, os consequencialistas (utilitaristas, hedonistas, etc.) estariam associados mais à reação mais comum do primeiro caso, por concluir em um resultado que houve um sacrifício menor pelo bem maior, ou uma dor menor pelo prazer maior, seja a terminologia que quiser usar.

O que é mais interessante aqui é como estes resultados contrastam com a perspectiva arraigada que temos das tradições. Como vimos, Kant e os deontologistas em geral como Rawls são associados à um grau maior de racionalidade na deliberação, dando pouco papel à emoção, enquanto a alcunha de consequencialistas normalmente é utilizada para sentimentalistas e emotivistas como Hume, Bentham e Sidgwick. Mais tarde em seu texto Greene trata também sobre punição, e como consequencialistas como Bentham priorizam tratamento e readaptação de infratores contra uma punição imediata severa equivalente ao crime, enquanto deontologistas como Kant defendem o contrário (GREENE, 2007, p. 50–58). É evidente que aqui é realizada outra simplificação quanto à posicionamentos gerais teóricos, todavia, com estas análises neurofisiológicas torna-se claro há um equívoco no senso comum sobre estas teorias tradicionais: deontologistas tendem a defender o mais emocionalmente estimulante (proteger invariavelmente a vida de alguém, punir à altura quem comete crimes), enquanto os consequencialistas envolvem-se em procedimentos mais calculistas, que necessitam deliberação a longo prazo (princípios de utilidade, foco no resultado geral da ação). Reitero: isto não é uma afirmação universal sobre estas teorias, porém nos dá ferramentas de análise que mostram como este tipo de investigação pode trazer informações e novas perspectivas à posicionamentos tradicionais.

Esta seria então a ironia secreta da alma de Kant, segundo Greene, ao citar Nietzsche: “ele escreveu contra os acadêmicos em apoio ao preconceito popular, mas para os acadêmicos e não para o povo”<sup>163</sup> (GREENE, 2007). Kant estaria tentando explicar e formular academicamente o senso moral da pessoa comum, dando força a juízos e axiomas de cunho fortemente emocionais e intuitivos através de uma roupagem racional, acadêmica. Ou seja, Kant

---

<sup>163</sup> No original: “*He wrote against the scholars in support of popular prejudice, but for scholars and not for the people*”

procura organizar e justificar o que ele vê como boas morais e não está focado em mudar as opiniões das pessoas sobre problemas morais controversos. Ele quer explicar por que as coisas erradas do cotidiano são erradas, derivando de princípios primeiros a imoralidade inerente de mentir, roubar e matar [...] (e também de se masturbar [...]). Seu objetivo é colocar sua versão preferida de moralidade de senso comum em uma fundação dedutiva sólida, análoga aos princípios fundacionais da matemática.<sup>164</sup> (GREENE, 2017, p. 73)

Encerrarei esta seção com outro experimento de Greene em seu artigo “O imperativo ratagórico: intuição moral e os limites do aprendizado afetivo”<sup>165</sup> (2017). Como o nome sugere, Greene continua na sua linha de pesquisa tendo em vista entender o fenômeno moral por trás de conceitos normativos tradicionais na filosofia moral através da filosofia experimental; desta vez sobre a possibilidade de intuições morais confiáveis pela neurociência comportamental com dados etológicos sobre ratos. Ele parte do que vimos anteriormente neste trabalho de que intuições morais, como intuições em geral, apesar de serem úteis por facilitarem a vida dentro de determinados grupos, estão inclinadas a serem falhas por dependerem de mecanismos heurísticos que ignoram diversas questões de probabilidade e afins, além de chocar-se com intuições morais de grupos externos. Mesmo tendo a possibilidade do aprendizado afetivo, ou seja, guiado emotivamente, de novas intuições através de mecanismos sofisticados flexíveis e integrativos, elas estão fadadas ao erro.

Ele exemplifica isto através de um experimento de pensamento com ratos baseado em pesquisas reais: há um habitat onde o ato de puxar uma alavanca recompensa os ratos com comida. Apesar de não fazer parte de sua natureza, com o tempo aprendem a utilizá-la e torna-se parte comum de seu cotidiano: aprenderam afetivamente (influenciados emotivamente pela recompensa da comida) através de um sistema sofisticado uma nova intuição que os afeta positivamente. Todavia, em algum momento, eles são retirados da gaiola e permitidos a comer sem limites. Ao voltar, continuam pressionando a alavanca deixando acúmulos de comida sem interesse algum. Três ratos filósofos então

---

<sup>164</sup> No original: “*aims to organize and justify what he regards as good morals and is not focused on changing people’s minds about controversial moral issues. He wants to explain why the wrongs of everyday life are wrong, deriving from first principles the inherent immorality of lying, stealing, and killing [...] (And also masturbating [...]). His goal is to put his preferred version of commonsense morality on a solid deductive foundation, analogous to the foundational principles of mathematics*”.

<sup>165</sup> No original: “*The rat-a-gorical imperative: Moral intuition and the limits of affective learning*”.

poderiam teorizar acerca do ato de puxar a alavanca: um atribui nosso ímpeto de puxar a alavanca com a própria existência do mecanismo como uma questão de caráter. Outro pode racionalizar que o faz por sentir vontade, pelo exercício, ou simplesmente por que tem vontade, mas outro ainda pode ser mais nobre e afirmar que o faz por um sentimento de dever, e “se tal rato for inspirado pelo sucesso da matemática, pode tentar derivar este imperativo ratagórico dos princípios da razão pura roedora”<sup>166</sup> (GREENE, 2017, p. 75). Por fim, um último rato pode observar a futilidade da tarefa, e propor que se pare a atividade mesmo identificando ser contra sua intuição, que clama para que se continue o processo. Mas ele sabe que está sofrendo com a experimentação de outrem, e que então seus sentimentos não necessariamente são confiáveis.

A parábola é simplista e, francamente, bastante boba, deliberadamente arquitetada tendo em vista a agenda consequencialista do autor. Todavia, reflete aspectos reais sobre nossa cognição moral. De fato, os exatos mesmos mecanismos que se ativam quando rechaçamos determinadas escolhas tradicionalmente deontológicas ativam-se também em ratos ao rechaçar aproximar-se de alavancas que já lhe deram choque. Como temos visto desde o a psicologia evolutiva, atravessando vieses cognitivos até aqui, nossa estrutura cognitiva, por mais sofisticada que seja, ainda é anacrônica com o meio atual em que vivemos. Ela fora selecionada centenas de milênios antes da palavra escrita – quiçá milhares ou milhões dependendo de qual antecessor e qual comportamento fora selecionado –, e a mais tempo ainda dos problemas morais práticos com os quais nos deparamos cotidianamente. Novamente, não temos aqui uma “refutação” da deontologia ou da ética das virtudes, mas uma parábola para entendermos que, conforme nossa compreensão contemporânea do funcionamento de nossa mente, talvez os pressupostos de muitas coisas na tradição não sejam nada mais do que instintos pré-históricos atuando fora do contexto original em que foram selecionados.

#### **6.4 Preconceito e Epistemologia**

A justiça não é cega.<sup>167</sup>  
– Rachlinski et al

<sup>166</sup> No original: “*And if such a rat were inspired by the success of mathematics, he might attempt to derive this rat-a-gorical imperative from principles of pure rodent reason*”.

<sup>167</sup> No original: “*Justice is not blind.*” Tradução nossa.

Embora ambos raça e gênero mereçam individualmente atenção individuais exclusivas, é indisputável que os métodos de discriminação são razoavelmente comuns ao tratar-se de raça e gênero que não seja especificamente a do homem branco, como vimos no capítulo 5. Além disso, características compartilhadas de membros destas categorias que participam de mais de uma delas auxiliam em ainda maior discriminação, o que demanda uma perspectiva interseccional ao tratar destes temas. De acordo com Kimberlé Crenshaw (1989), tratar ambas categorias como mutualmente excludentes é um desserviço para teses feministas e antirracistas, pois limita a análise do impacto que sofrem por acabar priorizando os membros mais “privilegiados”<sup>168</sup> de ambos os grupos, respectivamente homens negros e mulheres brancas (CRENSHAW, 1989, p. 139–140). Isto acaba por deixar de lado por exemplo mulheres negras, que são discriminadas em ambos contextos por serem “dois passos distantes da norma do homem branco”<sup>169</sup> (CRENSHAW, 1989, p. 143 em nota de rodapé), que por sua vez não é vista como uma “classe composta” pois são os padrões nos quais a norma é implicitamente definida, e portanto definem a posição de “neutralidade” pelas quais as outras são analisadas<sup>170</sup>. Assim, “mulheres Negras são protegidas apenas ao passo que suas experiências coincidem com aquelas daqueles de algum dos dois grupos”<sup>171</sup> (CRENSHAW, 1989, p. 143, tradução nossa). Torna-se claro aqui que isto se dá uma vez que

branquitude e masculinidade são pressupostos [...] pois estão intrinsecamente conectados na crença errônea na objetividade, na racionalidade, na, como Catherine Mackinnon coloca, ‘ponto-de-desperspectividade’ da perspectiva branca, masculina. Pois esta perspectiva não é articulada como branca e masculina (pois não precisa ser), por ser a norma, é presumida não ser subjetiva. É

---

<sup>168</sup> O termo “privilégio” tem sido utilizado de maneira ampla e fora de contexto tão frequentemente que sua definição é problemática. Utilizo aqui no sentido de Crenshaw em que se refere no “quão distante” a pessoa localiza-se em distância da norma de “homens brancos”, como veremos a seguir. Portanto as aspas, a fim de clarificar que é privilégio no sentido que “ a experiência interseccional é maior que a soma de racism e sexism” (no original: “*the intersectional experience is greater than the sum of racism and sexism*”) (CRENSHAW, 1989, p. 140, tradução nossa).

<sup>169</sup> No original: “*two steps removed from a white male norm*”.

<sup>170</sup> Isto não é apenas uma conjectura idiossincrática da autora, uma vez que em casos jurídicos em que envolviam pessoas de diversas raças e gêneros, nega-se às mulheres negras justiça mesmo que claramente, com evidências além de quaisquer questionamentos, estavam mais defasadas e tratadas mais desigualmente que seus pares “mais privilegiados”, ao ponto hediondo de juízes acusá-las de querer criar “super categorias” ou “novas classes de minorias protegidas” em seu favor (CRENSHAW, 1989, p. 141–143).

<sup>171</sup> No original: “*Black women are protected only to the extent that their experiences coincide with those of either of the two groups*”.

presumida que seja objetiva. Universal, até.<sup>172</sup> (PEREZ, 2019, p. 33, tradução nossa)

O tópico também não é do lugar de fala do autor desta tese, e não pressuponho representar a pesquisa acadêmica ou sequer parte da luta que estes grupos perpassam constantemente. Tentarei aqui, no entanto, dar voz à representantes destes grupos uma vez que é constante sua ausência em trabalhos que lidam diretamente com temas relacionados aos mesmos, especialmente na área geral de filosofia.

O foco desta seção será acerca do conceito de “neutralidade”, frequentemente associado aos poderes racionais para nos abstrairmos de nosso contexto e pensarmos de maneira “objetiva” na tradição filosófica. Ao pressupormos uma racionalidade capaz de atingir este posicionamento, atribuímos a ela a capacidade de funcionar sem nenhum fator externo influenciando-a – percepção esta que já vimos nas seções anteriores ser no mínimo altamente improvável. Iremos iniciar esta seção situando o problema, e então procederemos para versar primeiramente sobre estes viéses cognitivos que inconscientemente influenciam nossos juízos e decisões no que tange nossa percepção racial e de gênero, e posteriormente o impacto que esta perspectiva errônea de neutralidade acata na pesquisa filosófica com colocações da epistemologia feminista.

#### 6.4.1 *Racismo, ciência e filosofia*

Sempre que alguém lhe diz ‘sou objetivo, sou apolítico’,  
essa é a hora para cuidar sua carteira, pois você está  
prestes a ser roubado.<sup>173</sup>

– Angela Saini

Esta objetividade citada acima, então, remete ao tópico em questão acerca da neutralidade, uma vez que pretende ser uma perspectiva ausente de viéses – algo que repetidamente vimos que está além do controle de qualquer

---

<sup>172</sup> No original: “[...] *whiteness and maleness go without saying [...] because it is intrinsically linked to the misguided belief in the objectivity, the rationality, the, as Catherine Mackinnon has it, ‘point-of-viewlessness’ of the white, male perspective. Because this perspective is not articulated as white and male (because it doesn’t need to be), because it is the norm, it is presumed not to be subjective. It is presumed to be objective. Universal, even*”.

<sup>173</sup> No original: “*Whenever anybody tells you ‘I am objective, I am apolitical,’ that is the time to watch your wallet, because you’re about to have your pocket picked*”.

um. O espectador imparcial de Smith, o véu da ignorância na posição original de Rawls, as “condições ideais” dos realistas morais, dentre incontáveis exemplos na história da filosofia inadvertidamente convidam este paradigma. Assim, “a filosofia, no decorrer de sua história, muito comumente se colocou como puramente do reino das ideias, pretendendo-se isenta e atemporal, verdadeira para todos, independentemente de particularidades de sexo, raça e classe” (ARBO, 2020b, p. 107).

Isto nos traz de volta diretamente ao comentário realizado anteriormente sobre o viés profundamente racista e misógino de Kant. Os amplos trabalhos de Kant sobre geografia e antropologia – os temas principais onde ele destila sua pseudociência discriminadora – em mero volume tornam seus escritos filosóficos em geral apenas uma fração de sua produção acadêmica (ABUNDEZ-GUERRA, 2018, p. 118). Ele considera ambas “áreas *gêmeas*, geografia estudando os aspectos corporais e externos do homem, e antropologia estudando os aspectos psicológicos e internos do homem”<sup>174</sup> (ABUNDEZ-GUERRA, 2018, p. 118, grifo do autor, tradução nossa), dando o contexto geral de onde ele invoca nossas supostas faculdades intelectuais que regem os aspectos filosóficos de sua teoria:

[n]o pensamento de Kant, a antropologia ocupa uma posição central, uma vez que estuda a natureza do homem, e tal investigação subsume as três questões remanescentes que marcam o campo da filosofia: o que posso saber? O que devo saber? Pelo que devo ter esperança?<sup>175</sup> (MARGUTTI, 2013, p. 67)

O autor alemão identifica arbitrariamente (como o montante de sua perspectiva) quatro “raças” principais de seres humanos de acordo com a cor da pele e atribui a elas talentos inatos referentes ao seu contexto sociogeográfico nativo, e situa o homem branco europeu no topo do domínio das capacidades, enquanto as outras raças conseguem no máximo possuir talentos de maneira imperfeita. Acredito ser desnecessário adentrar-me nas especificidades do racismo pseudocientífico kantiano, mas é evidente que ele associa um poder de raciocínio inferior a qualquer membro externo à raça branca. Da mesma maneira,

<sup>174</sup> No original: “*twin sciences, geography studying the bodily and external aspects of man, and anthropology studying the psychological and internal aspects of man*”.

<sup>175</sup> No original: “[i]n Kant’s thought, anthropology occupies a central position, since it studies the nature of man and such an inquiry subsumes the three remaining questions which mark the field of philosophy: what can I know? what ought I know? what may I hope for?”.

Kant diminui o papel de mulheres em que não são portadoras dos mesmos poderes cognitivos de homens. A mulher na perspectiva kantiana resume-se a um papel secundário ao do homem, submetida a sua proteção e guia, sendo apenas capazes de procurar reproduzir aquilo inerente a eles de maneira limitada – como, no caso, a razão.

[...] Kant [reforça] a ideia do exercício da razão e da intelectualidade como sendo característica intrinsecamente masculina, e que uma mulher que busque conhecimento não está buscando o exercício da razão teórica e sim a masculinidade em si. (ARBO, 2020b, p. 111)

Novamente, racismo e sexismo andam de mãos dadas em uma tentativa de manter uma neutralidade ou objetividade de um discurso filosófico. Sem dúvida, isto não é exclusivo de Kant, uma vez que nesta época autores que defendiam algum tipo de direito das mulheres ou de negros e outras minorias eram a exceção, sendo mesmo os mais progressivos como Rousseau favoráveis a uma assimetria de direitos entre esses grupos e quem eles mesmos representavam. Em suma, os pensadores no período iluminista realizavam análises e generalizações sem sair de sua poltrona geográfica, contando apenas com representações exageradas ou superficiais de povos com os quais nunca interagiram adequadamente – ou seja, “definiram a humanidade sem de fato ter muito uma ideia de como a maioria da humanidade vivia ou como parecia”<sup>176</sup> (SAINI, 2019, p. 24), irônico uma vez que o objetivo original do iluminismo de trazer de volta como foco do empreendimento filosófico e científico a humanidade em contraste com o divino do período medieval.

O iluminismo reforçou a ideia de unidade humana, de uma qualidade essencial biológica que elevava os humanos sobre todas as outras criaturas. [...] [Q]uando pensadores europeus definiram o padrão do que consideravam o humano moderno, muitos construíram tal de acordo com suas próprias experiências e ao que acontecia de serem valor culturalmente naquela época.<sup>177</sup> (SAINI, 2019, p. 24 tradução nossa)

<sup>176</sup> No original: “*defined humanity without really having much of an idea how most of humanity lived or what it looked like*”.

<sup>177</sup> No original: “[t]he Enlightenment reinforced the idea of human unity, of an essential biological quality that elevated humans above all other creatures. [...] [W]hen European thinkers set the standard for what they considered a modern human, many built it around their own experiences and what they happened to value culturally at that time”.

E assim,

embora a racionalidade, enquanto conceito, não tivesse gênero em si, as condições para a prática do que se concebia por racionalidade eram, sim, generificadas: o acesso à educação era exclusividade masculina. A razão, portanto, tornava-se ilusoriamente uma característica masculina, enquanto as mulheres eram erroneamente associadas ao oposto da razão: ao frívolo, ao fútil, ao sentimento puro. (ARBO, 2020a, p. 126)

Foi neste momento também, que o conceito de “raça” surgiu em um âmbito científico<sup>178</sup>. Em 1798, o famoso e influente biólogo Carl Linnaeus<sup>179</sup> foi um dos primeiros a tentar adaptar o conceito de “raça” utilizado anteriormente em outros animais e seres vivos para humanos, resultando numa explosão posterior de tentativas de categorizar diferentes etnologias em grupos estáticos com características perenes (SAINI, 2019, p. 43). Sem dúvida, tal como no caso de machismo e sexismo, o *status quo* europeu fora extrapolado para confirmar viéses prévios de relações sócioeconômicas em diversas formulações pseudocientíficas sob a alcunha neutra e objetiva da ciência moderna.

A ciência de raças sempre situou-se na intersecção entre ciência e política, entre ciência e economia. Raça não era apenas uma ferramenta para classificar diferença física, mas também uma maneira de medir progresso humano, de passar julgamento às capacidades e direitos de outros.<sup>180</sup> (SAINI, 2019, p. 54)

Contemporaneamente, após séculos de tentativas de cientificação de preconceito com outros povos e modos de vida, sabe-se que raça não existe enquanto categoria científica para representações fenotípicas; tendo sido abandonada já no início do século passado e refutada profundamente em humanos por um dos primeiros proponentes da síntese moderna na biologia, Theodosius Dobzhansky, que urge à antropologia e outras áreas o uso de “raça” como representativa apenas de populações geográficas (1959, cap. 6; DUNN; DOBZHANSKY, 1949; GANNETT, 2013, p. 8). Milênios de escravatura,

<sup>178</sup> O prof. Paulo Margutti afirma que Kant foi quem estabeleceu tanto “raça” quanto “branquitude” na comunidade científica da época (MARGUTTI, 2013, p. 67), todavia ele não cita fontes específicas nem pude encontrar outras para embasar.

<sup>179</sup> Originalmente Carl Linnæus, ou posteriormente Carl Linné, estilizado amplamente no Brasil como Carlos Lineu.

<sup>180</sup> No original: “*Race science had always sat at the intersection of science and politics, of science and economics. Race wasn’t just a tool for classifying physical difference but was also a way of measuring human progress, of placing judgement on the capacities and rights of others.*”

supressão de direitos e todo o aparelhamento legal e social para mantê-los contra minorias gerou apenas exageros de uma suposta neutralidade que culminou no mínimo em maior segregação e opressão desses grupos e no máximo nas catástrofes eugenistas da segunda guerra até políticas públicas que ainda consistentemente os afetam.

A politically correct scientific terminology emerged. Researchers began referring to groups as “populations,” and occasionally as “ethnic groups.” But the departure from the old race science wasn’t quite as complete as it might have been. Although the parameters of research had changed, the racial categories were still alive in people’s minds. (SAINI, 2019, p. 79–80)

Mas não é eliminando diferenças entre o padrão de homem branco e quaisquer grupos minoritários que se acaba com o racismo na ciência; afinal, isto retornaria ao pressuposto inexistente de neutralidade. Raça, contemporaneamente, pode ser entendida como um constructo social adequado se considerado como representativo de determinados grupos étnicos em diálogo com as mesmas, podendo gerar resultados positivos para suas comunidades (DE WOLFE et al., 2021; JACKSON et al., 2019).

Reconhecer as particularidades de diferentes grupos sem privilegiar um sobre os outros é especificamente a tarefa da epistemologia feminista, que promulga que esta perspectiva é capaz de nos prover as maneiras mais ricas de acesso ao conhecimento, novamente contra uma suposta neutralidade normativamente discriminatória. Veremos brevemente algumas perspectivas na próxima subseção afim de cimentar a importância da necessidade da rejeição de ditas “imparcialidades” da racionalidade no campo da epistemologia.

#### *6.4.2 Localizando a epistemologia<sup>181</sup>*

John Dewey fora o primeiro auto-intitulado “naturalista” que reconheceu o papel da mulher como um constructo social associado ao seu contexto, e não à uma natureza inerente, e que, portanto, seria merecedora de direitos iguais à educação e sufrágio. Mas muito antes dele o feminismo já buscava tais

---

<sup>181</sup> Sou imensamente grato à Jade Arbo, Bruna Schneid, Gustavo Teles e Guilherme Meneghello, além dos grupos de pesquisa de Filosofia e Contemporaneidade e de Epistemologia Feminista pelos debates frutíferos que originaram esta seção.

reparações dentre muitas outras, que podem ser rastreadas à Mary Wollstonecraft<sup>182</sup> no século XVIII, reivindicando

que sejam as mulheres consideradas como criaturas morais e racionais de fato, como fazendo parte da universalidade do gênero humano proposta por seus contemporâneos, e que lhes fossem permitido buscar e alcançar uma educação igual à recebida pelos homens. (ARBO, 2020b, p. 114)

A história do feminismo é uma história de luta e subversão desta normatividade ética e epistemológica implicitamente excludente que está presente nos preceitos da sociedade tradicional. Segundo Longino, “o feminismo é um conjunto de posições e investigações caracterizado por interesses sociopolíticos comuns centrados na abolição da desigualdade de sexo e gênero” (LONGINO, 2008, p. 505). Nesta subseção veremos brevemente como autoras como Helen Longino, Miranda Fricker, Donna Haraway e outras desenvolveram este debate epistemológico sobre o feminismo e qual a importância desta perspectiva para nosso trabalho.

A epistemologia feminista surgiu a partir dos debates da epistemologia social, que procura trazer materialidade à discussão sobre o conhecimento elicitando a importância dos conhecimentos de diferentes comunidades para uma perspectiva mais completa para o conhecimento como um todo. Vimos extensivamente no capítulo anterior como preconceito e discriminação atuam em diversos níveis da vida humana afetando desde nossa cognição até o âmbito da lei. Segundo Longino, a importância da epistemologia feminista é identificar quando esta suposta neutralidade utilizada para embasar verdades ou valores epistêmicos positivos na verdade é unilateralmente desenvolvida e subserviente à homens brancos heterossexuais. Todavia, não procura afirmar que há, em contraste, uma importância comparativamente maior ou uma qualidade melhor no juízo epistemológico feminino. Reconhecendo o paradigma da metodologia científica de evitar a influência de vieses implícitos na produção, leitura e interpretação de dados – motivo pelo qual valoriza e enfatiza práticas como reprodutibilidade e avaliação por pares a fim de diminuir tal influência –, o objetivo da epistemologia feminista é o de apontar que ignorar tais vieses

---

<sup>182</sup> Também mãe de Mary Wollstonecraft Shelley, autora de uma das obras mais importantes de ficção científica e horror acerca de temas humanistas frente a um uso irreflexivo da ciência.

predominante sexistas das dinâmicas estatísticas e de poder de, como equipes menos diversas, por exemplo, está cedendo-se a estes vieses de maneira que afetam diretamente a qualidade dos empreendimentos científicos e epistemológicos (HARAWAY, 1988; LONGINO, 2008).

Donna Haraway desenvolve esta perspectiva com a noção de “saberes localizados”<sup>183</sup>, título também de influente artigo seu (1988). Segundo ela, a neutralidade seria análoga a um critério de conhecimento incorpóreo, além da tangibilidade do mundo real e da responsabilidade sobre o que pressupõe. Assim, segundo ela, a epistemologia feminista tem o papel de “situar” o conhecimento de maneira empírica através de sua corporificação (ARBO, 2020a, p. 126), uma vez que essa neutralidade etérea nada mais é do que o pressuposto do homem branco heterossexual como padrão. Quem não se adequa a este padrão “não [tem] a permissão de *não* ter um corpo, uma perspectiva finita, e portanto um viés inevitavelmente desqualificador e poluidor em qualquer discussão de consequência fora de nossos próprios pequenos círculos”<sup>184</sup> (HARAWAY, 1988, p. 575, grifo da autora, tradução nossa). O conceito de “injustiça epistêmica” surge neste contexto com Miranda Fricker, trazendo a epistemologia ainda mais para uma prática aplicada, onde esta falta de diversidade populacional no testemunho científico não apenas rende uma teoria mais pobre como também injusta, uma vez que as agências epistêmicas das minorias e afins são negadas e suas perspectivas ignoradas.

Assim, efetivamente, “epistemólogas feministas naturalizadas não demandam nenhuma quebra radical com os comprometimentos internos da ciência empírica” (ANDERSON, 1995, p. 51, tradução nossa), apenas a atenção e reconhecimento dos limites já pressupostos de maneira geral pela metodologia, mas que evidentemente têm falhado acerca da representatividade de minorias. O que isto levanta é que o processo de racionalidade, teórica ou não, é inerentemente social. Tanto a racionalidade no âmbito individual como no científico envolve a aquisição de evidências e postulação de crenças para agir. Especialmente no meio da pesquisa, independente da área, tais evidências “devem ser publicamente acessíveis, e em contextos experimentais,

---

<sup>183</sup> No original: “*Situated Knowledges*”.

<sup>184</sup> No original: “*not allowed not to have a body, a finite point of view, and so an inevitably disqualifying and polluting bias in any discussion of consequence outside our own little circle*”.

reprodutíveis por outrem” (ANDERSON, 1995, p. 53). Assim, as maneiras em que se dão as relações sociais influenciam diretamente a construção de teorias, e invariavelmente vieses cognitivos tem um papel neste processo. Tendo variedades de perspectivas, etnias, gêneros e diversidade em geral, é mais provável que haja a oportunidade de manter estes vieses em cheque. A ciência só tem a ganhar em seus critérios metodológicos ao abraçar uma perspectiva como a da epistemologia feminista (ANDERSON, 1995).

Este trabalho também. Temos visto que é possível, então, oferecer resistência aos fatores externos que podem influenciar nossos processos cognitivos. Greene sugere que o conhecimento sobre nossas intuições falhas permite-nos superá-las para agir de maneira mais bem-informada, juízes podem tentar manter seu racismo implícito de influenciar suas decisões, e é possível cultivar uma personalidade que nos mantenha menos inclinados a ceder ao contexto situacional. Nos esforçarmos a analisar nossos preconceitos implícitos e nos permitir aprender mais sobre outras culturas, vivências e modos de vida pode nos ajudar muito a tirarmos um pouco das limitações que nossa racionalidade possui a fim de decisões e juízos mais prudentes.

Na próxima seção, faremos uma última análise acerca do papel das intuições e das emoções no juízo moral após toda a investigação realizada neste trabalho. Por fim, desenvolveremos o montante da discussão realizada até então no próximo capítulo para encerrar com os resultados desta pesquisa.

### **6.50 papel das intuições**

O intuicionismo social segue a tese da racionalidade restringida na economia comportamental como vista anteriormente onde o juízo moral é majoritariamente intuitivo, um reflexo emotivo do meio em que o agente está situado.

Juízos morais são portanto definidos como avaliações (bom vs. mau) das ações ou caráter de uma pessoa que são feitos de acordo com um conjunto de virtudes tomados como obrigatórios por uma cultura ou subcultura.<sup>185</sup> (HAIDT, 2001, p. 817)

---

<sup>185</sup> No original: “*Moral judgments are therefore defined as evaluations (good vs. bad) of the actions or character of a person that are made with respect to a set of virtues held to be obligatory by a culture or subculture*”.

Este paradigma baseia-se em uma série de experimentos da psicologia social que procura identificar os aspectos irracionais do juízo moral, muitos dos quais já vimos. A influência do nojo, o efeito Knobe, o “gostômetro” e diversos outros paradigmas e experimentos foram influenciados ou mesmo diretamente desenhados por intuicionistas sociais. Difere-se do intuicionismo clássico, portanto, ao relegar o papel da razão como secundária no processo de elaboração do juízo moral. O aspecto social surge uma vez que não há então uma individualidade da razão para adquirir, processar e realizar um juízo moral; é, na verdade um processo social. Pelo fato de que somos animais sociais que vivem em grupos, muitos dos mecanismos de ajuizamento moral condizem primariamente com instintos de sobrevivência para convivência de grupo, intimamente relacionados com as dinâmicas sociais envolvidas nesse contexto.

O modelo dos intuicionistas sociais para o processo de ajuizamento moral é composto por quatro processos: (i) através de mecanismos como de T1, ou seja, de maneira automática e imediata, avaliamos intuitivamente algo como moralmente aprovável ou desaprovável – o juízo moral de fato. Posteriormente, há (ii) a racionalização *post hoc*, na qual nosso sistema cognitivo como em T2 procura justificar a existência do juízo moral gerado por T1 como se intencionalmente o tivesse realizado após tais considerações. Assim como no exemplo das pessoas calosotomizadas, não necessariamente temos ciência de que realizamos esta racionalização. Então, há (iii) o processo de persuasão fundamentada, onde há a produção verbal deste juízo moral e de sua justificação, que pode afetar outrem apesar de, pela própria natureza do juízo moral conforme os intuicionistas, é rara a persuasão de fato entre indivíduos que discordam. Por fim, há (iv) o processo da persuasão social, onde esta manifestação verbal da racionalização do juízo moral alinha-se com o viés familiar social do grupo inserido, podendo de fato influenciá-los diretamente mesmo sem a necessidade de uma defesa ostensiva lógica ou racional. Evidentemente, este último passo não ocorre tão facilmente com pessoas as quais não estão inseridas em seu grupo social, principalmente pela psicologia de equipe que elicitava aversão aos de fora do grupo (HAIDT, 2001).

Claramente, conforme temos visto através deste trabalho inteiro, há forte suporte experimental para a tese do intuicionismo social. Retornando também ao exemplo de Cline sobre o juízo de perigo que vimos no capítulo 4,

confirmamos aqui a herança de Hume: a racionalidade é insuficiente para a ação por si só, mas pode orientar as distinções morais dos sentimentos; ou no caso, crenças de T2 sem estímulo intuitivo-afetivo de T1 não são suficientes para motivar a ação, enquanto, como Prinz concorda, “podemos formar a crença que algo é moralmente errado apenas tendo uma emoção negativa direcionada para tal. Neste sentido, emoções são suficientes para a avaliação moral”<sup>186</sup> (PRINZ, 2006, p. 31). Assim, poderíamos concluir que a razão não tem papel na deliberação moral, encerrar este trabalho e parar de nos preocupar com grandes racionalizações filosóficas acerca de instintos primitivos. Afinal,

há evidência que o juízo moral é comprometido por emoções incidentais, heurísticas enganosas e confabulação. Filósofos utilizaram tais evidências para desenvolver argumentos detratores de acordo com os quais áreas-chave do juízo ético de senso comum são epistemicamente podres – são baseadas em processos epistemicamente defeituosos.<sup>187</sup> (NICHOLS, 2021, p. 5, tradução nossa)

Todavia, como vimos na seção anterior, há elementos que temos notado em diversos contextos de que, apesar da fraqueza da razão em nos orientar única e objetivamente a fim de uma otimização da geração do juízo moral, ela não é de todo inútil.

Quando falamos sobre psicopatia, vimos que Shaun Nichols fora um dos pesquisadores que enfatizou o papel das emoções no juízo moral, apontando como o caso da racionalidade intacta do psicopata contradizia a premissa das teorias racionalistas em geral. Todavia, em 2020, quando indagado em entrevista acerca da efetividade do programa experimental para lidar com problemas filosóficos e psicológicos, comentou sobre muito que envolve a pesquisa do intuicionismo social e dados relacionados:

[...] o trabalho que fiz inicialmente foi bastante motivado pelo papel da emoção no juízo moral, e ainda penso que a maior parte daquele trabalho mantém-se relevante, mas agora parece que o papel das emoções no juízo moral não explica tanto quanto gostaríamos. Na primeira onda da intersecção entre filosofia e psicologia [...] estávamos

<sup>186</sup> No original: “we can form the belief that something is morally wrong by simply having a negative emotion directed towards it. In this sense, emotions are sufficient for moral appraisal”.

<sup>187</sup> No original: “There is evidence that moral judgment is compromised by incidental emotions, misleading heuristics, and confabulation. Philosophers have used such evidence to develop debunking arguments according to which key areas of common-sense ethical judgment are epistemically rotten—they are based on epistemically defective processes”.

empurrando estas teorias sobre o juízo moral baseadas em emoções, e penso que muito das evidências para tais apenas não são adequadas às afirmações pois os efeitos das emoções são muito mais fracos do que parecia que seriam [...]. (NICHOLS, 2020, 4:52, tradução nossa)

Como vimos, Greene já procurava defender que não há apenas uma dicotomia entre intuição e racionalização, mas que é possível uma orientação da tomada de decisão a partir da análise racional de nossos vieses implícitos e do conhecimento de nossos processos cognitivos e como podem ser influenciados. Os intuicionistas sociais delegavam sistemas de T2 a um papel puramente racionalizante dos ímpetus de T1 – como diria Nichols, um papel *auxiliar* (NICHOLS, 2021, p. 6) –, mas Greene sempre defendeu uma concepção de processo duplo que englobava o papel da racionalidade na deliberação moral. Assim, mesmo que a análise das inúmeras evidências que Greene trouxera parecessem enfatizar uma espécie de intuicionismo social, concluía sempre destoando-se no que tange nossos sistemas reflexivos em T2 e sua capacidade de influenciar ou circumventar as intuições (GREENE, 2007, 2017; GREENE et al., 2001; GREENE; HAIDT, 2002). De fato, em uma impactante meta-análise em 2015 de 51 estudos englobando pesquisas em 5.102 participantes, Justin Landy e Geoffrey Goodman concluíram que o impacto do nojo e alguns outros fatores foram superestimados pelos intuicionistas como relevantes para a causalidade do juízo moral (LANDY; GOODWIN, 2015), o que deixa a tese da supremacia das emoções sobre qualquer racionalidade em maus lençóis.

Neste capítulo, trouxemos a compreensão empírica de dados relacionados à moralidade e à tomada de decisão. Aplicamos conceitos da filosofia experimental e navegamos por diversos tópicos outrora tratados nos capítulos prévios além de novas considerações com base no que foi levantado até então. O intuicionismo social era a barreira final contra qualquer maneira de racionalismo, todavia demonstrou-se insuficiente para lidar com as evidências práticas e sobre nossa funcionalidade, dando peso demais às emoções e intuições e peso de menos à razão. É possível então que a razão tenha, de fato, papel no juízo moral? Vamos debater esta possibilidade frente às evidências e discussões até então no próximo capítulo.

## 7. É POSSÍVEL REFUTAR KANT?

Sim. A teoria de Kant, por mais que possua um teor maior de informação empírica do autor sobre a ação que muitos na tradição e, como vimos com Williams no capítulo 2 e com Greene no capítulo anterior, não pressuponha uma racionalidade perfeita nem procura se apoiar em preceitos absolutamente metafísicos como Platão, ainda assim possui seus limites. Afirmações acerca de nossa cognitividade invariavelmente caem sob o crivo de nosso entendimento contemporâneo sobre o tema, o qual procurei explorar nos últimos capítulos. Todavia, como dito ao princípio, não é o intuito último deste trabalho o projeto negativo de filosofia experimental de procurar criticar ou refutar teorias – até por que uma tarefa deste nível iria necessitar muito mais do que um capítulo, ou possivelmente mesmo uma tese inteira. Nosso objetivo é positivo: aprofundar o entendimento acerca da deliberação moral.

Todavia, recordemo-nos da apresentação de Kant no segundo capítulo. Para ele, os aspectos idiossincráticos da deliberação, ou a apetição “inferior”, embora existentes, devem ser deixados de lados em prol de uma apetição “superior”, uma vontade imparcial, neutra, além de nosso auto-interesse. Em uma primeira instância, facilmente cai por terra se a tese que desenvolvi até então está correta: não há possibilidade de neutralidade ou imparcialidade absoluta nem nos mais idôneos dos humanos deste planeta. Mas nos recordemos das críticas de Prinz aos situacionistas, das autocríticas dos intuicionistas sociais e dos demais locais em que reconhecemos que, apesar desta impossibilidade, existe ainda assim um papel importante que nossas faculdades cognitivas possuem na tomada de decisão. Estaria Kant, então, na verdade, ultimamente correto? Pois segundo ele, a verdadeira liberdade está em escolher pautar-nos pela razão e agir conforme princípios universalizáveis – e não em sermos máquinas automáticas com um acesso direto à lei moral capazes de perfeição. Somos criaturas finitas, afinal. Estaríamos então, após estas mais de 150 páginas estudando os aspectos irracionais da razão na verdade *corroborando* um dos maiores racionalistas que já existiram?

Não. Embora Kant invariavelmente surja nos mais variados debates filosóficos, especialmente em um como este sobre racionalidade e moralidade, ele não é nosso foco, mas o desenvolvimento da resposta para esta pergunta

nos será uma maneira interessante de estabelecermos, afinal, o que todo este trabalho tem a dizer sobre o fenômeno da ação moral.

### **7.1 Como, afinal, ocorre o juízo moral?**

Procuramos desde o princípio deste trabalho entender como ocorre o juízo moral de maneira descritiva. Para isto, versamos sobre concepções filosóficas, paradigmas interdisciplinares, e debatemos evidências empíricas destas últimas e também da filosofia experimental em um contexto mais filosófico. Mas, de maneira resumida, como corre o juízo moral?

Em primeiro lugar, há estímulos. Desde questionários e experimentos controlados até reflexão e a atuação de fato em casos práticos, terminações nervosas informam nossos centros neurais de percepção a ponto de eliciar a ativação dos centros de atenção quando necessário. Esta trajetória chega em nosso sistema límbico e possivelmente passa pela amígdala, que parcialmente pauta o nível de engajamento atento e emocional do sistema nervoso para aquela situação. A partir daí, caso nossa percepção identifique um caso que elicite nossos centros emocionais, entramos em um modo cognitivo de resolução de problemas especializado, no qual temos alguma ciência da situação e temos que escolher uma opção de crença em detrimento de outras. Este processo até então, relacionado ao T1 no modelo cognitivo utilizado até então, é praticamente imediato, intuitivo, medido em milissegundos. Segundo os intuicionistas sociais, o juízo moral forma-se neste instante, e o resto é racionalização. Como no caso do processo justificacional dos participantes do experimento de Milgram ao tentar provêr razões para suas ações desumanas, “podemos nos iludir em pensar que é racional, mas isto é por que as variações situacionais sobrepujam nossa capacidade de realizar decisões decentes”<sup>188</sup> (PRINZ, 2009, p. 123).

De fato, uma série de mecanismos heurísticos e vieses implícitos estão imbuídos neste mecanismo, e em casos de proximidade pessoal do agente é o que Greene estipula como o juízo moral deontológico. Mas caso as áreas emocionais e a amígdala não tenham tamanho estímulo, usualmente trata-se de uma situação que percebemos como “não-moral”: os mecanismos para a tomada

---

<sup>188</sup> No original: “*We may delude ourselves into thinking it is rational, but that it is because the situational variables overwhelm our capacity to make decent decisions*”.

de decisão são idênticos ao da geração de juízos morais, todavia, diferem no seu grau de utilização.

No entanto, como vimos, outras pesquisas sugerem que o processo de ajuizamento moral pode envolver outros passos. Em situações de impessoalidade por exemplo, onde o agente pode julgar de maneira distante, áreas mais associadas com reflexão e pensamento considerado passam a se tornar mais ativas, aliviano as áreas de reação imediata. Em termos de dois processos, áreas de T2 começam processam os estímulos de T1, tornando-se predominantes enquanto a tomada de decisão ocorre, podendo inclusive elicitar novamente áreas de T1 conforme novas conclusões são encontradas. Greene sugere que o juízo moral de cunho consequencialista surge neste caso.

Evidentemente, esta explicação não somente é extremamente simplificada, como não é universal. Inúmeros fatores estão envolvidos no processo deliberativo com combinações quase infindas de tipos diferentes de situações. A saúde física e mental do agente, elementos externos que afetam suas capacidades cognitivas, o fator de tempo, local, tudo pode influenciar de diversas maneiras o processo de ajuizamento a ponto de alterar drasticamente os resultados. Ainda assim, os estudos interdisciplinares e os debates em filosofia experimental nos permitem identificar os principais mecanismos e limitações deste processo, e tal conhecimento pode impactar também na qualidade do juízo moral.

Vimos anteriormente considerações sobre as relações entre crenças e juízos morais. Uma breve análise do debate sobre a cognitividade epistemológica de tais é algo que pode nos ajudar a ter uma ideia mais clara do seu funcionamento. Um dos erros de Kahneman, muito provavelmente fruto de pouca experiência na área filosófica, é de atribuir a crença somente à processos de T2. Aaron Zimmerman estipula esta crítica ao requerer maior integração entre T1 e T2 para adequar-se à sua teoria pragmatista de crença (ZIMMERMAN, 2018, p. 65–66). Embora uma teoria deva explicar os fenômenos, e não esperar que estes se adequem a ela, Zimmerman levanta uma colocação interessante, análoga a esta passagem:

Filósofos da mente analíticos contemporâneos geralmente usam o termo “crença” para referir-se a atitudes que temos, a grosso modo,

sempre que tomamos algo para ser o caso ou o vemos como verdadeiro.<sup>189</sup> (SCHWITZGEBEL, 2015, p. 1)

Crenças não necessariamente pressupõe pensamento, uma vez que qualquer ação de um adulto normal pressupõe uma série de crenças estabelecidas e inconscientes ao interagir com o mundo, desde a data atual até propriedades físicas envolvidas em seu movimento.

Retornando ao exemplo do juízo de perigo proposto por Cline, apesar de vermos forte enfoque na herança humeana do papel secundário da razão na motivação moral, também notamos um problema com seu paradigma não-cognitivista. Se, como esta posição propõe, a motivação possui uma conexão íntima com a normatividade do juízo moral – ou seja, é critério-mor na qualidade daquele juízo de levar à ação de fato –, como é possível um juízo racional como a decisão de andar de avião, puramente elaborado por T2, sobrepõe a motivação emocional da fobia de vôo? Da mesma maneira, há o problema do que os eticistas da virtude chamam do *akrático*, do grego *akrasia* ou “falta da vontade”, em que o agente chega em um juízo moral motivado, porém falha em agir. Não-cognitivistas reduzem estes casos de incongruência entre a motivação e o juízo normativo a “situações adversas” ou “anormais”, apesar de representar um problema real para sua perspectiva. Esta abordagem de processo duplo, no entanto, é capaz de explicar que os processos cognitivos de T2 são capazes de informar os de T1 a fim de gerar motivação a partir de uma crença de uma maneira mais completa do que o não-cognitivismo.

Decerto, podem muito bem surgir novas evidências ou teorias que expliquem este fenômeno do *akrático* de maneira mais eficiente. Todavia, na situação atual, o conjunto da obra parece apontar para a conclusão de Cline, de que o juízo moral é sim uma crença. Assim, temos suporte, a princípio um externalismo cognitivista do juízo moral dentro deste entendimento até então.

Com isto em mente, voltemos aos objetivos secundários propostos no primeiro capítulo: (i) quais fatores externos, ou “não morais”, influenciam a concepção do juízo moral, e como identificá-los e contorná-los, se possível? (ii)

---

<sup>189</sup> No original: “*Contemporary analytic philosophers of mind generally use the term “belief” to refer to the attitude we have, roughly, whenever we take something to be the case or regard it as true*”.

O que caracteriza a deliberação moral em contraponto a uma deliberação não-moral? (iii) é possível depender da razão para a elaboração de juízos morais?

A resposta para (i) vimos paulatinamente desenvolver-se através dos capítulos. Existe um montante de elementos externos à racionalidade que influenciam no produto final do juízo moral: vieses cognitivos, neuropatologias, desconforto, preconceito, sugestionamento etc. Mesmo um ruído alto no momento da decisão pode hipoteticamente alterá-la. Em um contexto prático, é impensável que um agente seja capaz de considerar todos estes elementos para averiguar o quanto estão influenciando seu juízo. Ainda assim, há diversos elementos que podemos reconhecer que podem amenizá-los: estar ciente de vieses implícitos em situações que envolva seus alvos como com o preconceito, almejar a diversidade de seu grupo a fim de contornar câmaras de eco de opiniões semelhantes, e conhecendo sua patologia a fim de compensar suas limitações como no caso dos indivíduos calosotomizados ou com agênese do corpo caloso.

Já (ii) fora amplamente respondido. Emoções são parte fulcral daquilo que caracterizamos como moral em contraste ao não-moral ou amoral, todavia os mecanismos envolvidos em ambos os tipos de deliberação são os mesmos. A herança humeana neste aspecto mantém-se: decidir entre qual refeição fará no café da manhã e julgar se o homicídio de alguém fora justificado utiliza-se majoritariamente dos mesmos processos e das mesmas áreas. A única diferença fulcral é a intensidade em que respostas emocionais e atividades reflexivas se perpassarão. Essas diferenças diminuem mais ainda quando tratamos de juízos morais em um âmbito de distância pessoal alta do agente. Parafraseando Darwin, não há diferença de tipo entre o juízo de gosto, por exemplo, e o juízo moral; apenas de grau de ativação de determinados aspectos dessa estrutura para a tomada de decisão.

A resposta para o item (iii), no entanto, já é mais complexa, simbolizando o cerne desta pesquisa. Considerando sua primazia, tentaremos oferecer uma resposta adequada em uma seção dedicada a seguir.

## **7.2A racionalidade da razão**

É possível, então, depender da razão para a elaboração de juízos morais? De uma maneira bastante geral, não. As intuições advindas de impulsos

emocionais e outros fatores implícitos são em si bastante complexas e refletem de fato o cerne do juízo moral. Todavia, não são suficientes por si só. Então, de uma maneira mais específica, a resposta para a pergunta é “sim, em parte”.

Como vimos ao fim do capítulo 6, apesar da grande maioria da pesquisa acerca do papel das emoções no juízo moral apontar para sua supremacia acima de qualquer resquício de racionalidade, a conclusão é um pouco menos extrema do que isso: ambos a intuição emocional e a avaliação racional em conjunto são parte central na formação de juízos morais. Como diria Damásio: “emoção, sentimento e regulação biológica, todos interpretam um papel na razão humana”<sup>190</sup> (DAMASIO, 1994, p. 13).

Isto, todavia, não compromete com o intuito original deste trabalho. Se seu objetivo fora alcançado até aqui, fica evidente que teorias racionalistas fortes simplesmente não dão conta da miríade de fatores externos envolvidos no processo deliberativo. Como resumem Tooby e Cosmides,

Métodos “racionais” de tomada de decisão [...] lógica, matemática, teoria da probabilidade [...] são computacionalmente fracas: incapazes de resolver os problemas naturais adaptativos que nossos ancestrais tiveram que consistentemente resolver a fim de se reproduzir. [...] Esta performance pobre na maioria de problemas naturais é a razão primária da qual especializações de resolução de problemas foram favorecidas pela seleção natural ao invés de solucionadores de problemas de propósito geral. Apesar de afirmações amplas ao contrário, a mente humana não é pior do que racional [...] mas pode frequentemente ser melhor que racional.<sup>191</sup> (COSMIDES; TOOBY, 1994, p. 324)

O importante, portanto, é estar ciente destes mecanismos, de nossas limitações, e dos motivos pelos quais nossas intuições agem tais como são. A razão é de fato um critério importante para elaborarmos juízos morais, todavia é necessário que se entenda ela dentro do contexto do fenômeno como um todo, afinal

Se falsamente concluirmos que pessoas são irracionais de alguma maneira, podemos desperdiçar nosso esforço em tentar ajuda-las – e

---

<sup>190</sup> No original: “*Emotion, feeling, and biological regulation all play a role in human reason.*”

<sup>191</sup> No original: “*‘Rational’ decision-making methods [...] logic, mathematics, probability theory [...] are computationally weak: incapable of solving the natural adaptive problems our ancestors had to solve reliably in order to reproduce. [...] This poor performance on most natural problems is the primary reason why problem-solving specializations were favored by natural selection over general-purpose problem-solvers. Despite widespread claims to the contrary, the human mind is not worse than rational [...] but may often be better than rational.*”

podemos piorar as coisas ainda. Se falsamente concluirmos que as pessoas são racionais quando não as são, perdemos uma oportunidade para ajuda-las.<sup>192</sup> (BARON, 2007, p. 74, tradução nossa)

Ao início desta pesquisa, parecia claro que a razão e a racionalidade teriam papéis no máximo secundários em comparação com as emoções e intuições. Todavia, como estabelecemos, intuições em geral tendem a estar erradas. De maneira alguma, claro, este trabalho pretende defender um racionalismo puro aos moldes clássicos, ou mesmo contemporâneos: reiteramos novamente que é indiscutível o papel das emoções para a ética. Ainda assim, a razão ainda tem seu papel em identificar emoções, vieses implícitos, analisar o contexto, as consequências e diversos outros fatores a fim de guiar nossa tomada de decisão em um rumo mais bem-informado. E tal nada mais é do que o objetivo da filosofia experimental, como citado ao início deste trabalho: procurar oferecer a maior profundidade explanatória possível sobre o tópico em questão.

Em seu livro “Regras racionais: para uma teoria de aprendizado moral”<sup>193</sup> (NICHOLS, 2021), Nichols procura então conciliar estas evidências com a possibilidade da influência racional na deliberação. Ele postula uma racionalidade baseada em evidências, ou “racionalismo evidencialista”, que consiste em uma compreensão computacional do processo de tomada de decisão onde o aprendizado é baseado em estatística bayesiana através de análise empírica. Há muitas evidências citadas por ele que apoiam esta perspectiva. Por exemplo, mesmo o aprendizado moral dentro de contextos dominados por normas paroquiais, ou seja, normas específicas pra determinados grupos, é passível de influência pelo uso do raciocínio moral como inferência estatística racional, superando vieses baseados em grupos (PARTINGTON; NICHOLS; KUSHNIR, 2023), contradizendo o posicionamento do intuicionismo social.

Esta é, então, a irracionalidade da razão: boa parte dos processos envolvidos na racionalidade são irracionais, inconscientes, automáticos e imediatos. A tradição, em sua maioria esmagadora, falha em compreender a existência e a importância deste aspecto para a deliberação como um todo.

---

<sup>192</sup> No original: “*If we falsely conclude that people are irrational in some way, we may waste our effort in trying to help them — and we may even make them worse. If we falsely conclude that people are rational when they are not, we lose an opportunity to help them*”.

<sup>193</sup> No original: “*Rational Rules: Towards a Theory of Moral Learning*”.

Todavia, isto não implica na derrota completa de qualquer resquício de razão. De fato, reforça sua importância em que compreende nossa capacidade de identificar estes elementos para poder informar nossos poderes de julgamento e de ação. Mas sem sombra de dúvida, não é uma ferramenta infalível, e a frequente consciência acerca desse aspecto é crucial para seu funcionamento e utilidade.

## 8. CONSIDERAÇÕES FINAIS

Ao longo deste trabalho procuramos entender como funciona o processo envolvido na geração de juízos morais em pessoas humanas a fim de estabelecer um critério mínimo para o debate interdisciplinar através da relação entre este processo e o conceito de tomada de decisão. Partimos de um breve panorama geral do papel da razão na racionalidade filosófica tradicional e demonstramos mais aprofundadamente alguns exemplos contemporâneos de sua aplicabilidade como nos realistas propriamente ditos até outras teorias significativamente dependentes de uma noção forte de racionalidade como as de Rawls e Putnam. Estabelecemos então a filosofia experimental como braço do naturalismo filosófico para então compreender as bases das outras ciências mais relevantes para a compreensão da tomada de decisão: a biologia evolutiva, a economia comportamental, a psicologia cognitiva e as neurociências comportamentais. Procedemos, posteriormente, para a aplicação da compreensão destas ciências em dados relevantes em diversos aspectos para a compreensão dos tópicos centrais desta tese, e posteriormente trazendo dados diretos da filosofia experimental assim como trazendo dados anteriores para o debate filosófico.

Durante este trajeto, vimos que o estudo da tomada de decisão não apenas é uma ferramenta crucial para o desenvolvimento de qualquer teoria da ação, como também é um ponto central para o entendimento da consciência, processos de pensamento e estados de conhecimento em geral. Vimos também que o ajuizamento moral e a tomada de decisão possuem íntima relação, e o estudo de um informa também sobre o outro em algum grau.

Emoções, como sugerido desde o primeiro capítulo, possuem papel central na moralidade. Não apenas por toda decisão, menor que seja, envolver mecanismos afetivos em sua consideração, como também no exemplo da psicopatia sua amoralidade é caracterizada primariamente pelo atrofiamento das áreas neurológicas associadas com processos afetivos como a empatia, configurando um desafio a quem pressupõe que a racionalidade por si só pode ser referência maior ou mesmo exclusiva da moralidade.

A razão, no entanto, não é apenas uma ilusão falsa – ao menos não sempre. Embora seja facilmente sobrepujada por miríades de fatores externos à deliberação em si, é de suma importância para a deliberação moral. Nos permite

utilizarmos de evidências e métodos que permitem navegarmos por rotas menos obstruídas no momento de ajuizarmos sobre algo de cunho ético. Mas definitivamente nossa estrutura cognitiva não é um barco fácil de guiar, e devemos sempre estar atentos para que aquele ao leme não seja devorado pelo mar no meio do caminho.

Assim, emoções não são falhas cegas de um cérebro primitivo, mas mecanismos imprescindíveis selecionados por milênios para nossa sobrevivência, sendo capaz de lidar instantaneamente e – muito mais importante – adaptativamente com uma série de cenários que nem conseguimos imaginar antes de ocorrer conosco. A razão também, outrora tomada como antítese das emoções, agora demonstra-se como integrativa de seus processos, tendo papel limitado, porém útil em sua execução.

Iniciamos com um paralelo fictício da série *Star Trek*, onde esta separação igualmente fictícia entre racionalidade e sentimentos manifestam-se na existência dos hiper-rationais *Vulcanos*. Como sugerido, seu foco lógico é considerado sem igual, o que torna um de seus membros mais icônicos, *Spock*, um excelente oficial de ciências capaz de poderes de dedução incríveis. Todavia, como neste trabalho, torna-se muitas vezes uma característica limitadora à situação em questão. Ao segundo episódio da primeira temporada, isto torna-se evidente em que, ao acreditar estar vencendo o capitão de sua nave em um jogo de xadrez 3D<sup>194</sup>, *Spock* recebe um cheque-mate pela “ilogicidade” do curso de ação de seu oponente. De acordo com a figura científica que observamos ao investigar nossa capacidade de decisão, a possibilidade de teorias racionalistas fortes está também sob a mesma afronta.

Os estudos acerca de ambos fatores estes fulcrais para o processo deliberativo, como já vimos, são amplos e diversificados. Além disso, são dotados da característica de que, por tratarem-se de traços cientificamente empíricos, são verificáveis por pesquisas as quais tendem a sempre melhorar tecnologicamente e metodologicamente sobre si mesmas – ainda mais com o já presente histórico rico de material disponível sobre o tópico. Com esperança, este trabalho acrescentará ao progresso destas pesquisas e, de maneira mais importante ainda, ajudará para desobstruir a resistência na comunidade filosófica

---

<sup>194</sup> Um jogo que de fato existe, e que, como o nome sugere, é bastante mais complexo do que o xadrez comum.

de realizar-se pesquisa em diálogo direto com colegas de outras áreas que trabalham com tópicos relevantes para nosso conhecimento.

### **8.1 Limitações da pesquisa**

Um dos objetivos principais deste trabalho era o estabelecimento de critérios mínimos interdisciplinares acerca do debate sobre racionalidade. Desta maneira, por definição, é impossível abranger todas as áreas, evidências, métodos e discussões que impactam este conceito e simultaneamente manter o aspecto didático para filósofos – especialmente tratando-se de uma tese desenvolvida num programa de filosofia – dentro de um trabalho deste escopo. Assim, profundidade e conteúdo foram intencionalmente comprometidos a fim de abrangência (vide bibliografia). Futuramente, pretendo me adentrar na maioria das discussões exibidas brevemente aqui através de pesquisas mais direcionadas em formato de artigos ou capítulos de livro específicos sobre os temas em questão.

E ainda assim, mesmo focando na abrangência, houveram diversos tópicos que foram cortados por motivo de objetividade e de tempo. Era a intenção trabalhar também discussões acerca do debate e validade da metodologia científica contemporânea, trazendo desde estudos sobre a crise da reprodutibilidade na psicologia e o “*p hacking*” até debates sobre metodologia científica e filosofia da ciência em geral como em Feyerabend e van Fraassen dentre outros. Além da epistemologia feminista, a própria epistemologia social que a antecede também poderia trazer uma perspectiva mais ampla acerca da localização do conhecimento enfatizando a irrelevância de um conceito como “neutralidade” e a importância da diversidade para o estudo da verdade e do conhecimento. Estudos etológicos com chimpanzés, bonobos, corvídeos, lobos e outras espécies também serviriam um papel importante em entendermos os nossos próprios processos cognitivos e comportamentais. Há também importantes dados e estudos sobre psicopatologias como depressão e ansiedade, por exemplo, e como afetam nossas decisões, e também como podem surgir com o parto, hábitos de alimentação e sono.

Houveram também inúmeros estudos acerca dos fenômenos de notícias falsas e técnicas de propaganda aplicadas recentemente durante a pandemia, assim como o reforço e surgimento de diversas teorias conspiracionistas como

o movimento antivacina, adeptos da terra plana e defensores da escola de Chicago; todos bastante importantes no entendimento contemporâneo de vieses. Também da influência das mídias sociais em nossos mecanismos cognitivos, além de seu exponencial uso na pandemia e as consequências para os processos de tomada de decisão. E talvez o elemento mais importante que optei por deixar de lado nas fases iniciais da tese, mas que agora é tema central nas discussões é o da polarização política. Desde o aspecto da psicologia de equipe de “nós” contra “eles” até estudos de traços de personalidade de diferentes posicionamentos políticos, tema amplamente trabalhado por Jonathan Haidt, Joshua Knobe e muitos outros, é um debate que tange muito do que foi discutido neste trabalho e possui relevância urgente para a manutenção da democracia e do Estado de direito.

E ainda, a razão não está plenamente a salvo. Como vimos, há um constante combate pelo controle das rédeas dos processos cognitivos entre intuições, emoções e razões para agir. Um dos critérios de análise para uma pesquisa futura, por exemplo, seria investigar o quão confiável é apostar na racionalidade para influenciar os processos intuitivos a longo prazo. Este elemento parece estar ausente do material que trabalhamos. Há também ampla pesquisa em filosofia sobre confiabilidade e intuição como com Alvin Goldman, que poderia informar mais profundamente o papel da intuição na epistemologia moral.

Além disso, o método científico parte do pressuposto de que novas evidências podem – e devem – alterar o paradigma sobre o tópico em questão conforme maior profundidade é encontrada com novas investigações. Assim como a perspectiva fortemente focada no papel da emoção no início dos anos 2000 alterou-se para o que parece sugerir uma importância igualitária entre razão e emoção para a deliberação, isto pode alterar-se no futuro conforme nossa compreensão da deliberação prossegue. O que imagino que tenha ficado claro nesta pesquisa é que não são apenas pessoas da filosofia que possuem interesse em desenvolver o tópico, e esta multitude de pesquisadores, equipes e pesquisas indubitavelmente enriquecerão este campo de maneira exponencialmente crescente conforme o passar do tempo. Espero que esta pesquisa possa ajudar, nem que minimamente, para o progresso desta tarefa.

## REFERÊNCIAS

ABUNDEZ-GUERRA, V. F. How to Deal with Kant's Racism—In and Out of the Classroom. **Teaching Philosophy**, v. 41, n. 2, p. 117–135, 2018.

ALFANO, M.; LOEB, D. Experimental Moral Philosophy. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall ed. Stanford: Metaphysics Research Lab, Stanford University, 2017. p. 34.

ALLGAYER, H. **Significado e função do princípio de seleção natural na Origem das Espécies**. 1. ed. Minas Gerais: Editora Virtual Books, 2013.

AMERICAN PSYCHIATRIC ASSOCIATION (ED.). **Diagnostic and statistical manual of mental disorders: DSM-5**. 5th ed ed. Washington, D.C: American Psychiatric Association, 2013.

ANDERSON, E. Feminist Epistemology: An Interpretation and a Defense. **Hypatia**, v. 10, n. 3, p. 50–84, 1995.

ARBO, J. A Filosofia Feminista encontra a Epistemologia Moral: situando o pensamento de Margaret Urban Walker. **Seara Filosófica**, n. 21, p. 125–135, 2020a.

ARBO, J. B. “Todo ente racional e finito”: o agente moral kantiano encontra a reivindicação de Wollstonecraft. Em: FOFANO, D. et al. (Eds.). **Mulheres, Filosofia e Vivências**. Paraná: Editora CRV, 2020b. p. 107–116.

ASSUMPÇÃO, G. A. O sumo bem e a lei moral. **Aufklärung: journal of philosophy**, v. 4, n. 2, p. 107–120, 3 ago. 2017.

AUTON, A. et al. A global reference for human genetic variation. **Nature**, v. 526, n. 7571, p. 68–74, 1 out. 2015.

AVENANTI, A.; SIRIGU, A.; AGLIOTI, S. M. Racial Bias Reduces Empathic Sensorimotor Resonance with Other-Race Pain. **Current Biology**, v. 20, n. 11, p. 1018–1022, 8 jun. 2010.

BABIÁK, P.; HARE, R. D. **Snakes in suits: when psychopaths go to work**. London: HarperCollins, 2007.

BABILONI, C. et al. Attentional processes and cognitive performance during expectancy of painful galvanic stimulations: a high-resolution EEG study. **Behavioural Brain Research**, v. 152, n. 1, p. 137–147, 4 jun. 2004.

BARBOSA, E. Pode o Naturalismo Reduccionista de David Copp Acomodar o Realismo e o Construtivismo Moral? **Revista Dissertatio de Filosofia**, Dossiê Naturalismo. v. 2, p. 37–64, 1 jan. 2015.

BARON, J. **Thinking and Deciding, 4th Edition**. 4th edition ed. New York: Cambridge University Press, 2007.

BEZERRA, A. R. G.; STANISCUASKI, F.; BARBOSA, M. C. Brazilian research productivity fellows in physics and nursing under a gender lens: 17 years of data. **Cadernos de Gênero e Tecnologia**, v. 15, n. 46, p. 64–79, 2022.

BIANCO, M. et al. Gifted Girls: Gender Bias in Gifted Referrals. **Roeper Review**, v. 33, n. 3, p. 170–181, 1 jul. 2011.

BIZER, G. Y.; SCHINDLER, R. M. Direct evidence of ending-digit drop-off in price information processing. **Psychology & Marketing**, v. 22, n. 10, p. 771–783, 2005.

BLAIR, R. J. R. A cognitive developmental approach to morality: Investigating the psychopath. **Cognition**, v. 57, p. 1–29, 1995.

BLAIR, R. J. R. et al. The psychopathic individual: A lack of responsiveness to distress cues? **Psychophysiology**, v. 34, p. 192–98, 1997.

BLAIR, R. J. R. Understanding the development of the psychopathic individual: An affective cognitive neuroscience approach. Em: BARCH, D. (Ed.). **Cognitive and Affective Neuroscience of Psychopathology**. Maryland: Oxford University Press, 2006.

BORG, J. S.; LIEBERMAN, D.; KIEHL, K. A. Infection, incest, and iniquity: investigating the neural correlates of disgust and morality. **Journal of Cognitive Neuroscience**, v. 20, n. 9, p. 1529–1546, set. 2008.

BRITO, C. et al. Harassment in Brazilian universities: how big is this problem? The Federal University of Rio Grande do Sul (UFRGS) as a case study. **Anais da Academia Brasileira de Ciências**, v. 94, 2022.

BURCH, R.; PARKER, K. A. Charles Sanders Peirce. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Summer 2022 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2022.

CALDER, A. J.; LAWRENCE, A. D.; YOUNG, A. W. Neuropsychology of fear and loathing. **Nature Reviews Neuroscience**, v. 2, n. 5, p. 35072584, 1 maio 2001.

CAMERER, C.; LOEWENSTEIN, G.; PRELEC, D. Neuroeconomics: How Neuroscience Can Inform Economics. **Journal of Economic Literature**, v. 43, n. 1, p. 9–64, fev. 2005.

CHAGAS, F. C. O fato da razão e o sentimento moral enquanto disposição moral do ânimo. **Studia Kantiana**, n. 11, p. 139–161, dez. 2011.

CHEW, A. Aristotle's Functional Theory of the Emotions. **Organon F: Medzinárodný asopis Pre Analytickú Filozofiu**, v. 16, n. 1, p. 5–37, 2009.

CHIODO, L. M. et al. Blood lead levels and specific attention effects in young children. **Neurotoxicology and Teratology**, v. 29, n. 5, p. 538–546, set. 2007.

CHOY, O.; RAINE, A.; SCHUG, R. Larger striatal volume is associated with increased adult psychopathy. **Journal of Psychiatric Research**, v. 149, p. 185–193, 1 maio 2022.

CHU, S.; GEARY, K. Physical stature influences character perception in women. **Personality and Individual Differences**, v. 38, n. 8, p. 1927–1934, 1 jun. 2005.

CHURCHLAND, P. S. **Braintrust: What Neuroscience Tells Us about Morality**. Reprint edition ed. Princeton, NJ: Princeton University Press, 2008.

CLINE, B. Cognitivism, Motivation, and Dual-Process Approaches to Normative Judgment. **Ergo, an Open Access Journal of Philosophy**, v. 4, n. 20180912, 21 set. 2017.

COHON, R. Hume's Moral Philosophy. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall 2010 ed. Stanford: Metaphysics Research Lab, Stanford University, 2010.

COPP, D. Milk, Honey, and the Good Life on Moral Twin Earth. **Synthese**, v. 124, n. 1, p. 113–137, 2000.

COPP, D. Why Naturalism? **Ethical Theory and Moral Practice**, v. 6, p. 179–200, 2003.

COPP, D. Why Naturalism? Em: **Morality in a Natural World: Selected Essays in Metaethics**. Cambridge: Cambridge University Press, 2007. p. 33–54.

CORDEIRO, M. **Breve Histórico da Neurociência**. . Conferência/Slides apresentado em FURG. Rio Grande, 2011.

COSMIDES, L.; TOOBY, J. Better than Rational: Evolutionary Psychology and the Invisible Hand. **The American Economic Review**, v. 84, n. 2, p. 327–332, 1994.

**Cosmos: A Personal Voyage. The Shores of the Cosmic Ocean** Estados Unidos da América PBS, , 28 set. 1980.

CRENSHAW, K. **Demarginalizing the intersection of race and sex: A Black feminist critique of antidiscrimination doctrine**. University of Chicago Legal Forum. **Anais...**1989. Disponível em: <<https://chicagounbound.uchicago.edu/uclf/vol1989/iss1/8>>

DALL'AGNOL, D. **Valor Intrínseco - Metaética, Ética Normativa e Ética Prática em G. E. Moore**. Florianópolis: UFSC, 2005.

DAMASIO, A. **Descartes' Error: Emotion, Reason, and the Human Brain**. Putnam: Avon Books, 1994.

DAMASIO, A. R. **Looking for Spinoza: joy, sorrow and the governance of life**. London: Heinemann, 2003.

DAMASIO, H. et al. The return of Phineas Gage: clues about the brain from the skull of a famous patient. **Science (New York, N.Y.)**, v. 264, n. 5162, p. 1102–1105, 20 maio 1994.

DANIEL KAHNEMAN; SHANE FREDERICK. Representativeness Revisited: Attribute Substitution in Intuitive Judgment. Em: GILOVICH, T.; GRIFFIN, D. W.; KAHNEMAN, D. (Eds.). **Heuristics and biases: the psychology of intuitive judgment**. Cambridge, U.K. ; New York: Cambridge University Press, 2002. p. 49–81.

DANZIGER, S.; LEVAV, J.; AVNAIM-PESSO, L. Extraneous factors in judicial decisions. **Proceedings of the National Academy of Sciences**, v. 108, n. 17, p. 6889–6892, 26 abr. 2011.

DARWALL, S. **The British Moralists and the Internal “Ought”: 1640-1740**. Cambridge: Cambridge University Press, 1995.

DARWIN, C. **The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life**. 6th. ed. London: John Murray, 1876.

DARWIN, C. **The Descent of man, and selection in relation to sex**. 2nd edition, fifteenth thousand ed. London: John Murray, 1889.

DARWIN, C. **The expression of the emotions in man and animals**. 2d. ed. London: John Murray, 1890.

DARWIN, C. The Origin of Species by Means of Natural Selection. Em: **Encyclopedia Britannica - Great Books of the Western World**. London: Encyclopedia Britannica Inc, 1952.

DAVIDSON, R. J.; PUTNAM, K. M.; LARSON, C. L. Dysfunction in the neural circuitry of emotion regulation--a possible prelude to violence. **Science (New York, N.Y.)**, v. 289, n. 5479, p. 591–594, 28 jul. 2000.

DE BONDT, W. F. M. Stock prices: insights from behavioral finance. Em: LEWIS, A. (Ed.). **The Cambridge Handbook of Psychology and Economic Behaviour**. Cambridge Handbooks in Psychology. Cambridge: Cambridge University Press, 2008. p. 64–104.

DE LIMA, F. J. G.; DE OLIVEIRA, N. H. Emoção, juízo moral e razão: Damásio e a crítica do racionalismo ético. **Veritas (Porto Alegre)**, v. 64, n. 2, p. 34220, 15 out. 2019.

DE OLIVEIRA, N. Recasting the naturalism–normativity debate: neuroscience, neurophilosophy, neuroethics. **Princípios: Revista de Filosofia (UFRN)**, v. 20, n. 33, p. 79–103, 2013.

DE WOLFE, T. J. et al. Chasing Ghosts: Race, Racism, and the Future of Microbiome Research. **mSystems**, v. 6, n. 5, p. e00604-21, 26 out. 2021.

DENNETT, D. **Darwin’s Dangerous Idea**. London: Penguin Books, 1995.

DENNETT, D. Darwin's "strange inversion of reasoning". **PNAS: Proceedings of the National Academy of Sciences of the United States of America**, v. 106, p. 100661–10065, 16 jun. 2009.

DENNETT, D. C. **Elbow room: the varieties of free will worth wanting**. Cambridge, Mass: MIT Press, 1984.

DENNETT, D. C. **Freedom Evolves**. Reprint edition ed. London: Penguin Books, 2004.

DOBZHANSKY, T. **Genetics and the Origin of Species**. 3. ed. [s.l.] Columbia University Press, 1959.

DORIS, J. M.; PLAKIAS, A. How to argue about disagreement: Evaluative diversity and moral realism. Em: **Moral psychology, Vol 2: The cognitive science of morality: Intuition and diversity**. Cambridge, MA, US: MIT Press, 2008. p. 303–331.

DUNN, L. C.; DOBZHANSKY, T. Heredity, Race and Society. **Population (French Edition)**, v. 4, n. 4, p. 775, 1949.

EPICURO. **Carta sobre a felicidade (A Meneceu)**. Álvaro Lorencini; Enzo del Carratore. São Paulo: Editora UNESP, 1999.

ESKINE, K. J.; KACINIK, N. A.; PRINZ, J. J. A bad taste in the mouth: gustatory disgust influences moral judgment. **Psychological Science**, v. 22, n. 3, p. 295–299, mar. 2011.

FEINSTEIN, J. S. et al. The Human Amygdala and the Induction and Experience of Fear. **Current Biology**, v. 21, n. 1, p. 34–38, jan. 2011.

FESSLER, D. M. T. Emotions and Cost-benefit Assessment. Em: GIGERENZER, G.; SELTEN, R. (Eds.). **Bounded rationality: the adaptive toolbox**. Dahlem workshop reports. 1. MIT Press paperback ed ed. Cambridge, Mass.: MIT Press, 2002. p. 191–215.

FESTINGER, L. Cognitive Dissonance. **Scientific American**, v. 207, n. 4, p. 93–106, out. 1962.

FESTINGER, L. et al. **When prophecy fails**. New ed. ed. London: Pinter & Martin, 2008.

FINUCANE, M. L. et al. Gender, race, and perceived risk: The "white male" effect. **Health, Risk & Society**, v. 2, n. 2, p. 159–172, 1 jul. 2000.

FLYNN, J.; SLOVIC, P.; MERTZ, C. K. Gender, race, and perception of environmental health risks. **Risk Analysis: An Official Publication of the Society for Risk Analysis**, v. 14, n. 6, p. 1101–1108, dez. 1994.

FURTADO, L. M. F. et al. The Role of Neuroplasticity in Improving the Decision-Making Quality of Individuals With Agenesis of the Corpus Callosum: A Systematic Review. **Cureus**, v. 14, n. 6, p. e26082, jun. 2022.

GANNETT, L. Theodosius Dobzhansky and the genetic race concept. **Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences**, v. 44, n. 3, p. 250–261, set. 2013.

GAZZANIGA, M. S. Forty-five years of split-brain research and still going strong. **Nature Reviews. Neuroscience**, v. 6, n. 8, p. 653–659, ago. 2005.

GIGERENZER, G.; SELTEN, R. (EDS.). **Bounded rationality: the adaptive toolbox**. 1. MIT Press paperback ed ed. Cambridge, Mass.: MIT Press, 2002.

GLIMCHER, P. W.; FEHR, E. **Neuroeconomics: decision making and the brain**. 1. ed. Cambridge: Academic Press, 2014.

GREENE, J. The Secret Joke of Kant's Soul. Em: SINNOT-ARMSTRONG, W. (Ed.). **Moral Psychology, Vol. 3: The Neuroscience of Morality: Emotion, Disease, and Development**. Cambridge, MA: MIT Press, 2007.

GREENE, J. D. et al. An fMRI Investigation of Emotional Engagement in Moral Judgment. **Science**, v. 293, n. 5537, p. 2105–2108, 14 set. 2001.

GREENE, J. D. The rat-a-gorical imperative: Moral intuition and the limits of affective learning. **Cognition, Moral Learning**. v. 167, p. 66–77, 1 out. 2017.

GREENE, J. D.; YOUNG, L. The Cognitive Neuroscience of Moral Judgment and Decision-Making. Em: GAZZANIGA, M. (Ed.). **The Cognitive Neurosciences**. Massachussets: MIT Press, 2020. v. 6.

GREENE, J.; HAIDT, J. How (and where) does moral judgment work? **Trends in Cognitive Sciences**, v. 6, n. 12, p. 517–523, 1 dez. 2002.

GREGORY, P. A. **Quine's Naturalism: Language, Theory, and the Knowing Subject**. London; New York: Continuum, 2008.

GROSS, R. D. **Psychology: the science of mind and behaviour**. Eighth edition ed. London: Hodder Education, 2020.

HAIDT, J. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. **Psychological Review**, v. 108, n. 4, p. 814–834, out. 2001.

HAIDT, J.; BJORKLUND, F. Social intuitionists answer six questions about moral psychology. Em: **Moral psychology, Vol 2: The cognitive science of morality: Intuition and diversity**. Cambridge, MA, US: MIT Press, 2008. p. 181–217.

HARAWAY, D. Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. **Feminist Studies**, v. 14, n. 3, p. 575, 1988.

HARE, R. M. A Philosophical Autobiography. **Utilitas**, v. 14, n. 3, p. 269–305, nov. 2002.

HASELTON, M. G.; NETTLE, D. The Paranoid Optimist: An Integrative Evolutionary Model of Cognitive Biases. **Personality and Social Psychology Review**, v. 10, n. 1, p. 47–66, fev. 2006.

HASELTON, M. G.; NETTLE, D.; ANDREWS, P. W. The evolution of cognitive bias. Em: BUSS, D. M. (Ed.). **The Handbook of Evolutionary Psychology**. Hoboken, NJ, US: John Wiley & Sons Inc, 2005. p. 724–746.

HERCULANO-HOUZEL, S. The human brain in numbers: a linearly scaled-up primate brain. **Frontiers in Human Neuroscience**, v. 3, 2009.

HILDEBRAND, D. John Dewey. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Winter 2021 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2021.

HILLER, R. F. Um Novo Olhar Sobre a Natureza Humana: A Proposta de Filosofia Moral de David Hume. **Revista Seara Filosófica**, n. 18, p. 62–76, 29 ago. 2019.

HUME, D. **A Treatise of Human Nature: Being an Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects**. London: White-Hart, 1739.

HUME, D. **An Enquiry Concerning the Principles of Morals**. London: Leeds Electronic Text Centre, 1777.

HUME, D. **Tratado da natureza humana: uma tentativa de introduzir o método experimental de raciocínio nos assuntos morais**. São Paulo: Editora UNESP, 2009.

IGANSI, L. N. **Normatividade e valor no naturalismo moral**. Dissertação—São Leopoldo: Unisinos, 2014a.

IGANSI, L. N. Naturalismo, ação e normatividade. **Controvérsia**, v. 10, n. 3, p. 131–142, 2014b.

IGANSI, L. N. A falácia naturalista na metaética contemporânea: usos e equívocos. **Fundamento**, v. 1, n. 8, p. 11–31, 2014c.

IGANSI, L. N. Uma Análise da Desigualdade Social sob a Perspectiva do Princípio da Diferença na Teoria da Justiça de John Rawls. **Saberes: Revista Interdisciplinar de Filosofia e Educação**, v. 1, n. 9, 29 maio 2014d.

IGANSI, L. N. Consegue um homem comer um mamute inteiro? Psicologia moral do valor e normatividade. **Controvérsia**, v. 11, n. 1, p. 57–70, 18 out. 2015.

IGANSI, L. N. Objetividade ética e a morte da ontologia em Putnam. **Cognitio: Revista de Filosofia**, v. 21, n. 2, p. 246–259, 28 jan. 2021.

JACKSON, L. et al. Including Vulnerable Populations in the Assessment of Data From Vulnerable Populations. **Frontiers in Big Data**, v. 2, p. 19, 28 jun. 2019.

JOHNSON, R.; CURETON, A. Kant's Moral Philosophy. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Spring 2018 ed. Stanford: Metaphysics Research Lab, Stanford University, 2018.

KAHAN, D. M. et al. Culture and Identity-Protective Cognition: Explaining the White-Male Effect in Risk Perception. **Journal of Empirical Legal Studies**, v. 4, n. 3, p. 465–505, nov. 2007.

KAHANE, G. The armchair and the trolley: an argument for experimental ethics. **Philosophical Studies**, v. 162, n. 2, p. 421–445, 1 jan. 2013.

KAHNEMAN, D. **Thinking, Fast and Slow**. 1st edition ed. New York: Farrar, Straus and Giroux, 2013.

KAHNEMAN, D.; TVERSKY, A. Subjective probability: A judgment of representativeness. **Cognitive Psychology**, v. 3, n. 3, p. 430–454, jul. 1972.

KANDEL, E. R. et al. (EDS.). **Principles of neural science**. Sixth edition ed. New York: McGraw Hill, 2021.

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M. **Essentials of Neural Science and Behavior**. 1 edition ed. Norwalk, CT: Appleton & Lange, 1996.

KANT, I. **Critique of Practical Reason**. Tradução: Mary Gregor. Cambridge: Cambridge University Press, 1997.

KANT, I. **Prolegomena to any future metaphysics that will be able to come forward as science: with selections from the Critique of pure reason**. Rev. ed ed. Cambridge, UK ; New York: Cambridge University Press, 2004.

KARBOWSKI, J. Aristotle on the Deliberative Abilities of Women. **Apeiron**, v. 47, n. 4, 1 jan. 2014.

KATZ, N. **Prosopagnosia: Oliver Sacks' Battle with "Face Blindness"**. Disponível em: <<https://www.cbsnews.com/news/massachusetts-concerning-new-drug-resistant-gonorrhea-strain/>>. Acesso em: 20 jan. 2023.

KHALILI, Z.; MORADI, M. H. **Emotion recognition system using brain and peripheral signals: Using correlation dimension to improve the results of EEG**. 2009 International Joint Conference on Neural Networks. **Anais...** Em: 2009 INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS. jun. 2009.

KIM, J. What is "Naturalized Epistemology?" **Philosophical Perspectives**, v. 2, p. 381–405, 1988.

KLEIN, J.; GIGLIONI, G. Francis Bacon. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall 2020 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2020.

KNOBE, J. Intentional action and side effects in ordinary language. **Analysis**, v. 63, n. 3, p. 190–194, 1 jul. 2003.

KNOBE, J.; NICHOLS, S. **Experimental Philosophy**. Oxford, New York: Oxford University Press, 2008.

KNOBE, J.; NICHOLS, S. *Experimental Philosophy*. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Winter 2017 ed. Stanford: Metaphysics Research Lab, Stanford University, 2017.

KORSGAARD, C. M. Skepticism about Practical Reason. **The Journal of Philosophy**, v. 83, n. 1, p. 5, jan. 1986.

KORSGAARD, C. M. Rationality. Em: GRUEN, L. (Ed.). **Critical Terms for Animal Studies**. 1. ed. Chicago: University of Chicago Press, 2018. p. 448.

KOWALSKI, H. First Individual Diploid Human Genome Published By Researchers at J. Craig Venter Institute: Sequence Reveals that Human to Human Variation is Substantially Greater than Earlier Estimates. **J. Craig Venter Institute Press Release**, 2007.

LANDY, J. F.; GOODWIN, G. P. Does Incidental Disgust Amplify Moral Judgment? A Meta-Analytic Review of Experimental Evidence. **Perspectives on Psychological Science: A Journal of the Association for Psychological Science**, v. 10, n. 4, p. 518–536, jul. 2015.

LE GUIN, U. K. **A Mão Esquerda da Escuridão**. Susana L. de Alexandria. São Paulo, SP: Editora Aleph, 2014.

LENMAN, J. Moral Naturalism. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Summer 2013 ed. Stanford: Metaphysics Research Lab, CSLI, Stanford University, 2013.

LONGINO, H. Epistemologia Feminista. Em: **Compêndio de epistemologia**. 2. ed. São Paulo: Edições Loyola, 2008. p. 505–546.

MACINTYRE, A. C. **After virtue: a study in moral theory**. 3rd ed ed. Notre Dame, Ind: University of Notre Dame Press, 2007.

MANNING, K. C.; SPROTT, D. E.; ARTICLE., J. D. SERVED AS EDITOR AND B. R. SERVED AS ASSOCIATE EDITOR FOR THIS. Price Endings, Left-Digit Effects, and Choice. **Journal of Consumer Research**, v. 36, n. 2, p. 328–335, 2009.

MARGUTTI, P. Pragmatism and decolonial thinking: an analysis of Dewey's ethnocentrism. **Cognitio. Revista de filosofia**, v. 14, n. 1, p. 63–83, 2013.

MAYER, E. A. Gut feelings: the emerging biology of gut–brain communication. **Nature reviews. Neuroscience**, v. 12, n. 8, p. 10.1038/nrn3071, 13 jul. 2011.

MICHEL, C. M.; MURRAY, M. M. Towards the utilization of EEG as a brain imaging tool. **NeuroImage**, v. 61, n. 2, p. 371–385, jun. 2012.

MILLER, M. B. et al. Abnormal Moral Reasoning in Complete and Partial Callosotomy Patients. **Neuropsychologia**, v. 48, n. 7, p. 2215–2220, jun. 2010.

MLODINOW, L. **Subliminar: Como o inconsciente influencia nossas vidas**. Tradução: Claudio Carina. Rio de Janeiro, RJ: Zahar, 2013.

MOORE, G. E. **Principia Ethica**. 2. ed. London: Cambridge University Press, 1993.

NAVIA, R. In Memoriam Hilary Putnam (1926 – 2016): Problemáticas Decisivas y Autocrítica Incansable. **ἔλεγχος**, v. I, n. 1, p. 133–142, maio 2016.

NEEDLEMAN, H. L. Low-Level Lead Exposure and the IQ of Children: A Meta-analysis of Modern Studies. **JAMA**, v. 263, n. 5, p. 673, 2 fev. 1990.

NICHOLS, S. How Psychopaths Threaten Moral Rationalism: Is it Irrational to Be Amoral? **The Monist**, v. 85, n. 2, p. 285–303, 2002.

NICHOLS, S. **The Philosophy And Psychology Of Morality**. , 28 maio 2020. Disponível em: <<https://www.youtube.com/watch?v=OppfzA7oTf0>>. Acesso em: 20 fev. 2023

NICHOLS, S. **Rational rules: towards a theory of moral learning**. First edition ed. Oxford, United Kingdom ; New York, NY, United States: Oxford University Press, 2021.

OHME, R. et al. Analysis of neurophysiological reactions to advertising stimuli by means of EEG and galvanic skin response measures. **Journal of Neuroscience, Psychology, and Economics**, v. 2, n. 1, p. 21–31, 2009.

OLEJNICZAK, P. Neurophysiologic basis of EEG. **Journal of Clinical Neurophysiology: Official Publication of the American Electroencephalographic Society**, v. 23, n. 3, p. 186–189, jun. 2006.

OLOFSSON, A.; RASHID, S. The White (Male) Effect and Risk Perception: Can Equality Make a Difference? **Risk Analysis**, v. 31, n. 6, p. 1016–1032, 2011.

P. H. BARRETT. Early writings of Charles Darwin. Em: GRUBER, H. E. (Ed.). **Darwin on man. A psychological study of scientific creativity; together with Darwin's early and unpublished notebooks**. 1. ed. London: Wildwood House, 1974. p. 479.

PAPINEAU, D. Naturalism. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Summer 2021 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2021.

PARTINGTON, S.; NICHOLS, S.; KUSHNIR, T. Rational learners and parochial norms. **Cognition**, v. 233, p. 105366, 1 abr. 2023.

PEREZ, C. C. **Invisible women: data bias in a world designed for men**. New York: Abrams Press, 2019.

PHELPS, E. A. et al. Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation. **Journal of Cognitive Neuroscience**, v. 12, n. 5, p. 729–738, set. 2000.

PRINZ, J. The emotional basis of moral judgments. **Philosophical Explorations**, v. 9, n. 1, p. 29–43, mar. 2006.

PRINZ, J. **The Emotional Construction of Morals**. Oxford: Oxford University Press, 2007a.

PRINZ, J. Empirical philosophy and experimental philosophy. Em: KNOBE, J.; NICHOLS, S. (Eds.). **Experimental Philosophy**. [s.l.] Oxford University Press, 2007b. p. 189--208.

PRINZ, J. The Normativity Challenge: Cultural Psychology Provides the Real Threat to Virtue Ethics. **The Journal of Ethics**, v. 13, n. 2, p. 117–144, 1 set. 2009.

PULIDO, C.; RYAN, T. A. Synaptic vesicle pools are a major hidden resting metabolic burden of nerve terminals. **Science Advances**, v. 7, n. 49, p. eabi9027, 3 dez. 2021.

PUTNAM, D. et al. Disability and Justice. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall 2019 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2019.

PUTNAM, H. **Reason, Truth, and History**. Cambridgeshire, Cambridge: Cambridge University Press, 1981.

PUTNAM, H. **Ethics without Ontology**. Cambridge, MA: Harvard University Press, 2005.

QUINE, W. VAN O. Epistemology Naturalized. Em: **Ontological Relativity and Other Essays**. New York, NY: Columbia University Press, 1969. p. 69–91.

QUINE, W. VAN O. **Word and object**. New ed ed. Cambridge, Mass: MIT Press, 2013.

RACHLINSKI, J. J. et al. Does Unconscious Racial Bias Affect Trial Judges. **Notre Dame Law Review**, v. 84, p. 1195, 2009 2008.

RAWLS, J. **Justice as Fairness: A Restatement**. 2<sup>a</sup> ed. [s.l.] Belknap Press of Harvard University Press, 2001.

RAWLS, J. **Uma Teoria da Justiça**. Tradução: Jussara Simões. 2. ed. São Paulo: Martins Fontes, 2002.

RAWLS, J. **A Theory of Justice: Original Edition**. Cambridge: The Belknap Press of Harvard University Press, 2005a.

RAWLS, J. **Political Liberalism**. Expanded edition ed. New York: Columbia University Press, 2005b.

REYNA, V. F.; BRAINERD, C. J. Dual processes in decision making and developmental neuroscience: A fuzzy-trace model. **Developmental Review**, p. S0273229711000153, set. 2011.

ROBERTSON, S. I. **Problem solving: perspectives from cognition and neuroscience**. Second edition ed. London: Routledge, 2017.

ROXO, M. R. et al. The Limbic System Conception and Its Historical Evolution. **The Scientific World Journal**, v. 11, p. 2427–2440, 8 dez. 2011.

SAINI, A. **Superior: the return of race science**. Boston: Beacon Press, 2019.

SANTOS, R. O. D. **Estrutura e Funções do Córtex Cerebral**. Monografia—Brasília: Centro Universitário de Brasília, 2002.

SCHNALL, S. et al. Disgust as Embodied Moral Judgment. **Personality & social psychology bulletin**, v. 34, n. 8, p. 1096–1109, ago. 2008.

SCHROEDER, S. R. et al. Separating the effects of lead and social factors on IQ. **Environmental Research**, v. 38, n. 1, p. 144–154, out. 1985.

SCHWITZGEBEL, E. Belief. Em: **The Stanford Encyclopedia of Philosophy**. Summer 2015 ed. Stanford: Metaphysics Research Lab, 2015. p. 53.

SEN, A. The Formulation of Rational Choice. **The American Economic Review**, v. 84, n. 2, p. 385–390, 1994.

SHARMA, M. K. The impact on consumer buying behaviour: Cognitive dissonance. **Global Journal of Finance and Management**, v. 6, n. 9, p. 833–840, 2014.

SHERMER, M. **The believing brain: from ghosts and gods to politics and conspiracies--how we construct beliefs and reinforce them as truths**. 1st ed ed. New York: Times Books, 2011.

SILVA, R. P. A. D. et al. Misogyny in Brazilian Federal Government Agencies for Science and High-Education. **Anais da Academia Brasileira de Ciências**, v. 93, 2021.

SILVEIRA, M. DE M. **O papel da simpatia nas distinções morais: uma leitura humeana numa perspectiva evolucionista**. Dissertação—São Leopoldo: Unisinos, 28 abr. 2010.

SILVEIRA, M. DE M. Deconstructing Prinz's moral theory. **Controvérsia**, v. 12, n. 3, p. 188–192, 12 dez. 2016.

SILVEIRA, M. DE M.; ALLGAYER, H.; HILLER, R. F. Distinções entre a teleologia aristotélica e a darwinista e sua superação pela teoria da síntese estendida. **Argumentos - Revista de Filosofia**, n. 16, p. 37–49, 2016.

SIMON, H. A Behavioral Model of Rational Choice. Em: **Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting**. New York: Wiley, 1957.

SIMON, H. A. Bounded rationality in social science: Today and tomorrow. **Mind & Society**, v. 1, n. 1, p. 25–39, 1 mar. 2000.

SINNOTT-ARMSTRONG, W.; MILLER, C. B. (EDS.). **Virtue and character**. Cambridge, Mass. London, England: The MIT Press, 2017.

SKINNER, B. F. **Sobre o behaviorismo**. Tradução: Maria da Penha Villalobos. 9. ed ed. São Paulo: Cultrix, 2004.

SMITH, A. **An inquiry into the nature and causes of The wealth of nations**. Petersfield: Harriman House, 2007.

SMITH, J. E. H. **Irrationality: a history of the dark side of reason**. Princeton: Princeton University Press, 2019.

SOLSO, R. L.; MACLIN, O. H.; MACLIN, M. K. **Cognitive psychology**. Harlow: Pearson Education Limited, 2014.

STURGEON, N. L. Ethical Naturalism. Em: COPP, D. (Ed.). **The Oxford Handbook of Ethical Theory**. New York: Oxford University Press, 2007. p. 91–121.

THALER, R. H.; SUNSTEIN, C. R. **Nudge: improving decisions about health, wealth, and happiness**. New Haven: Yale University Press, 2008.

TIMMONS, M. Putnam's Moral Objectivism. **Erkenntnis (1975-)**, v. 34, n. 3, p. 371–399, 1991.

TVERSKY, A.; KAHNEMAN, D. Judgement under Uncertainty: Heuristics and Biases. Em: KAHNEMAN, D. (Ed.). **Thinking, Fast and Slow**. 1. ed. New York: Farrar, Straus and Giroux, 2013. p. 419–432.

VECCHIATO, G. et al. Changes in Brain Activity During the Observation of TV Commercials by Using EEG, GSR and HR Measurements. **Brain Topography**, v. 23, n. 2, p. 165–179, 1 jun. 2010.

VERANO, J. W.; FINGER, S. Ancient trepanation. Em: **Handbook of Clinical Neurology**. 3. Edinburgh: Elsevier, 2009. v. 95p. 3–14.

WEBER, E. U.; JOHNSON, E. J. Decisions Under Uncertainty: Psychological, Economic, and Neuroeconomic Explanations of Risk Preference. Em: GLIMCHER, P. W. et al. (Eds.). **Neuroeconomics: Decision-Making and the Brain**. 1. ed. Cambridge, MA: Academic Press, 2008. p. 556.

WEINSTEIN, D. Herbert Spencer. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall 2019 ed. [s.l.] Metaphysics Research Lab, Stanford University, 2019.

WHEELER, G. Bounded Rationality. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Fall 2020 ed. Stanford: Metaphysics Research Lab, Stanford University, 2020.

WILLIAMS, G. Kant's Account of Reason. Em: ZALTA, E. N. (Ed.). **The Stanford Encyclopedia of Philosophy**. Summer 2018 Edition ed. Stanford: Stanford University Press, 2018.

WITELSON, S. F.; KIGAR, D. L.; HARVEY, T. The exceptional brain of Albert Einstein. **The Lancet**, v. 353, n. 9170, p. 2149–2153, jun. 1999.

XU, H.; PLATANIOTIS, K. N. **Affect recognition using EEG signal**. 2012 IEEE 14th International Workshop on Multimedia Signal Processing (MMSP). **Anais...** Em: 2012 IEEE 14TH INTERNATIONAL WORKSHOP ON MULTIMEDIA SIGNAL PROCESSING (MMSP 2012). Banff, AB: IEEE, set. 2012. Disponível em: <<http://ieeexplore.ieee.org/document/6343458/>>. Acesso em: 12 mar. 2019

YETER, D.; BANKS, E. C.; ASCHNER, M. Disparity in Risk Factor Severity for Early Childhood Blood Lead among Predominantly African-American Black Children: The 1999 to 2010 US NHANES. **International Journal of Environmental Research and Public Health**, v. 17, n. 5, p. 1552, 28 fev. 2020.

ZIMMERMAN, A. **Moral epistemology**. New York: Routledge, 2010.

ZIMMERMAN, A. Z. **Belief: A Pragmatic Picture**. Oxford: Oxford University Press, 2018.